LEHRBUCH

DER

ALGEBRA

VON

HEINRICH WEBER

PROFESSOR DER MATHEMATIK AN DER UNIVERSITÄT STRASSBURG

KLEINE AUSGABE IN EINEM BANDE

ZWEITER UNVERÄNDERTER ABDRUCK

SPRINGER FACHMEDIEN WIESBADEN GMBH

ISBN 978-3-663-06369-8 ISBN 978-3-663-07282-9 (eBook)
DOI 10.1007/978-3-663-07282-9

VORWORT.

Als mir vor Jahren von der Verlagsbuchhandlung der Vorschlag zu einer neuen Auflage des ersten Bandes meines Lehrbuches der Algebra gemacht wurde, glaubte ich aus zwei Gründen nicht darauf eingehen zu können. Einmal, weil ich das Lehrbuch der Algebra, das als ein Ganzes geplant war, nicht durch besondere Ausgabe des ersten Bandes in zwei Teile zerreißen wollte, dann auch, weil ich zu einer Neubearbeitung des Ganzen in meinem Alter nicht mehr die Kraft fühlte. Ich war daher der Verlagshandlung sehr dankbar, daß sie auf meinen Vorschlag einer kleinen Ausgabe des Lehrbuches als ein in sich abgeschlossenes Ganze einging. So ist das vorliegende kleine Lehrbuch entstanden. Ich wollte mich dabei aber nicht auf das ganz Elementare und überall zu Findende beschränken, und so ist denn auch von den schwierigeren Partien der Algebra einiges mit aufgenommen worden, um auch dem kleinen Lehrbuch mehr selbständiges Interesse zu geben. Daß dabei die Grundzüge der Theorie der algebraischen Zahlen nicht fehlen durften, wird der Leser verstehen. Auch manche Verbesserungen und Erweiterungen hat der Text erfahren.

Es bleibt mir noch übrig, den Fachgenossen, die mich durch sachkundigen Rat bei der Korrektur unterstützt haben, besonders den Herren Löwy, Epstein und Levi, meinen Dank zu sagen.

INHALTSVERZEICHNIS.

Vierter Abschnitt: **Wurzeln.**

Fünfter Abschnitt: **Kubische und biquadratische Gleichungen.**

Sechster Abschnitt: **Der Sturmsche Lehrsatz.**

Siebenter Abschnitt: **Genäherte Berechnung der Wurzeln.**

Achter Abschnitt: **Gruppen.**

Neunter Abschnitt: **Die Galoissche Theorie.**

Zehnter Abschnitt: **Zyklische Gleichungen.**

Elfter Abschnitt: **Kreisteilung.**

Zwölfter Abschnitt: **Auflösung der Kreisteilungsgleichung.**

Dreizehnter Abschnitt: **Algebraische Auflösung von Gleichungen.**

Vierzehnter Abschnitt:
Zahlen und Funktionale eines algebraischen Körpers.

Erster Abschnitt.

Determinanten.

§ 1.

Permutationen.

Eins der kräftigsten Hilfsmittel der Algebra sind die Determinanten. Begriff und einfachste Rechenregeln gehören heute wohl zu den allgemein bekannten Elementen und die tiefer eindringende Theorie hat in Lehrbüchern eingehende Darstellungen gefunden, unter denen aus älterer Zeit Baltzer (Theorie und Anwendung der Determinanten) und unter den neueren Kowalewski (Einführung in die Determinantentheorie) genannt seien. Es wird daher genügen, wenn wir uns hier auf eine kurze Darlegung und Ableitung der Sätze beschränken, die wir in der Folge gebrauchen.

Das System der n Ziffern

(1) $$1, 2, 3 \ldots n$$

läßt sich bekanntlich auf $n! = 1.2.3 \ldots n$ verschiedene Arten anordnen, diese Anordnungen heißen Permutationen. Sind $\alpha_1, \alpha_2 \ldots \alpha_n$ dieselben Ziffern wie (1), in irgend einer Reihenfolge, so ist

$$\mathfrak{A} = \alpha_1, \alpha_2, \alpha_3 \ldots \alpha_n$$

eine solche Permutation.

I. Man kann jede Permutation $\mathfrak{A}$ aus jeder anderen auf unendlich viele Arten durch wiederholte Vertauschung zweier Ziffern oder Transpositionen herleiten.

Man erkennt dies, wenn man zuerst eine Ziffer α_ν, die nicht an νter Stelle steht, mit ν vertauscht und so fortfährt, bis alle Ziffern die gewünschte Stelle haben.

Wir wählen nun n beliebige, aber untereinander verschiedene Zahlenwerte

(2) $$a_1, a_2, a_3 \ldots a_n$$

und bilden das Produkt aller Differenzen je zweier dieser Zahlen:

$$P = (a_1 - a_2)\,(a_1 - a_3) \ldots (a_1 - a_n)\,(a_2 - a_3) \ldots (a_2 - a_n) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a_{n-1} - a_n) \tag{3}$$

Das Produkt

$$P' = (a_{\alpha_1} - a_{\alpha_2})\,(a_{\alpha_1} - a_{\alpha_3}) \ldots (a_{\alpha_1} - a_{\alpha_n})\,(a_{\alpha_2} - a_{\alpha_3}) \ldots (a_{\alpha_2} - a_{\alpha_n}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a_{\alpha_{n-1}} - a_{\alpha_n}) \tag{4}$$

hat denselben absoluten Wert wie P, aber das entgegengesetzte Vorzeichen, wenn unter den Faktoren von P' eine ungerade Anzahl das entgegengesetzte Vorzeichen wie die entsprechende Differenz von P hat.

II. Man unterscheidet hiernach die Permutationen $\mathfrak{A}$ in zwei Arten:

Erste Art solche Permutationen, für die das Vorzeichen von P' dasselbe ist, wie das von P; zweite Art solche Permutationen, für die P und P' das entgegengesetzte Vorzeichen haben.

Da P durch eine Transposition (z. B. durch Vertauschung von 1 mit 2) das Vorzeichen ändert, so folgt: Die Permutationen der ersten Art bestehen aus einer geraden, die der zweiten Art aus einer ungeraden Anzahl von Transpositionen.

Hieraus ergibt sich, daß durch eine Transposition jede Permutation der ersten Art in eine der zweiten Art und jede der zweiten Art in eine der ersten Art übergeht, und folglich:

III. Es gibt ebensoviel Permutationen erster, wie zweiter Art, nämlich $\frac{1}{2}\,n!$.

§ 2.

Determinanten.

Wir betrachten jetzt ein System von n^2 beliebigen Größen, mit denen die rationalen Rechenoperationen ausgeführt werden können. Zu einer einfachen Bezeichnung dieser Größen wählen wir einen Buchstaben mit einem doppelten Index $a_i^{(k)}$, worin i sowohl als k die Reihe der Ziffern 1, 2, 3 ... n durchlaufen soll.

Zur besseren Übersicht ordnen wir diese Größen in ein Quadrat, so daß alle a mit demselben oberen Index in einer Horizontalreihe, alle a mit demselben unteren Index in einer Vertikalreihe stehen, und bezeichnen dies Quadrat mit Δ, also:

$$\Delta = \begin{vmatrix} a_1^{(1)}, & a_2^{(1)}, & a_3^{(1)} & \dots & a_n^{(1)} \\ a_1^{(2)}, & a_2^{(2)}, & a_3^{(2)} & \dots & a_n^{(2)} \\ a_1^{(3)}, & a_2^{(3)}, & a_3^{(3)} & \dots & a_n^{(3)} \\ \dots & \dots & \dots & \dots & \dots \\ a_1^{(n)}, & a_2^{(n)}, & a_3^{(n)} & \dots & a_n^{(n)} \end{vmatrix} \tag{1}$$

Die Horizontalreihen heißen auch Zeilen, die Vertikalreihen Kolonnen oder Spalten. Die Größen $a_i^{(k)}$ heißen die Elemente des Systems Δ.

Wir wollen aber unter dem zwischen vertikalen Strichen eingeschlossenen Quadrat nicht nur den Komplex der Größen a verstehen, sondern eine bestimmte arithmetische Verbindung dieser Größen, die sich ausrechnen läßt, sobald die a numerisch gegeben sind, und die wir jetzt beschreiben wollen.

Man bilde das Produkt aller Elemente a, die in der von links oben nach rechts unten gehenden Diagonale stehen:

$$M = a_1^{(1)}\, a_2^{(2)}\, a_3^{(3)} \dots a_n^{(n)}, \tag{2}$$

leite daraus $n!$ Produkte M, M', M'' ... her, indem man die unteren Indices permutiert, und gebe jedem so entstandenen Produkt das positive oder negative Zeichen, je nachdem die angewandte Permutation zur ersten oder zur zweiten Art gehört:

$$M' = \pm a_{\alpha_1}^{(1)}\, a_{\alpha_2}^{(2)} \dots a_{\alpha_n}^{(n)}. \tag{3}$$

Die Summe aus diesen mit Vorzeichen versehenen Produkten

$$M + M' + M'' + \dots = \Sigma M$$

soll Δ sein. Δ wird die Determinante der n^2 Elemente $a_i^{(k)}$ genannt, und zwar, wenn die Unterscheidung notwendig ist, eine n-reihige Determinante (auch Determinante $n$ten Grades oder $n$ter Ordnung). Das Glied M dieser Summe, d. h. also das Produkt aller in der Diagonale des Quadrates stehenden Elemente, wird das Hauptglied genannt.

Die Bezeichnung (1) ist in vielen Fällen zu umständlich; es sind daher noch andere, kürzere Zeichen im Gebrauch. So setzt

Jacobi, indem er nur das Hauptglied der entwickelten Determinante ausführlich schreibt:

$$\Delta = \Sigma \pm a_1^{(1)} a_2^{(2)} \dots a_n^{(n)}, \tag{4}$$

und Kronecker noch kürzer:

$$\Delta = | a_i^{(k)} |. \tag{5}$$

Beide Bezeichnungen sind aber nur dann ganz deutlich, wenn die Elemente in der hier vorausgesetzten Weise durch zwei Indices bezeichnet sind, und durchaus unanwendbar, wenn die Elemente z. B. numerisch gegeben sind.

Es kommen bisweilen Determinanten vor, bei denen

$$a_i^{(k)} = a_k^{(i)}$$

ist, bei denen also in (1) die symmetrisch zur Diagonale des Quadrates stehenden Elemente einander gleich sind. Wir werden in diesen Fällen gewöhnlich beide Indices (um ihre Gleichwertigkeit anzudeuten) unten hinsetzen, also

$$a_{i,k} = a_{k,i}$$

setzen. Solche Determinanten heißen symmetrisch.

Wenn wir in dem Produkt

$$M' = \pm a_{\alpha_1}^{(1)} a_{\alpha_2}^{(2)} \dots a_{\alpha_n}^{(n)} \tag{6}$$

die Faktoren umstellen, so ändert sich sein Wert nicht. Wir können also die Faktoren auch so anordnen, daß die unteren Indices in ihrer natürlichen Reihenfolge 1, 2 ... n erscheinen. Dabei werden dann die oberen Indices in einer gewissen Weise permutiert erscheinen, also M' die Form erhalten:

$$\pm a_1^{(\beta_1)} a_2^{(\beta_2)} \dots a_n^{(\beta_n)}, \tag{7}$$

worin

$$(\beta_1, \beta_2 \dots \beta_n) = \mathfrak{B},$$

ebenso wie

$$(\alpha_1, \alpha_2 \dots \alpha_n) = \mathfrak{A}$$

eine Permutation der Ziffern 1, 2 ... n bedeutet. Man kann die Anordnung $\mathfrak{B}$ dadurch erhalten, daß man in den Faktoren von M' die Transpositionen, die zu $\mathfrak{A}$ geführt haben, von der letzten anfangend, rückgängig macht, um in der Reihe der unteren Indices wieder die ursprüngliche Anordnung zu erhalten. Die dabei sich ergebende Reihenfolge der oberen Indices ist dann die Permutation $\mathfrak{B}$. Es folgt daraus, daß $\mathfrak{B}$ zur ersten oder zur zweiten Art gehört, je nachdem $\mathfrak{A}$ zur ersten oder zur zweiten

Art gehört, da beide durch die gleiche Anzahl von Transpositionen entstehen. Die Gesamtheit der $\mathfrak{B}$ stellt ebenso wie die Gesamtheit der $\mathfrak{A}$ alle Permutationen der n Elemente dar, da zwei verschiedene $\mathfrak{A}$ niemals zu demselben $\mathfrak{B}$ führen können. Damit ist bewiesen:

IV. Die Determinante Δ kann auch dadurch gebildet werden, daß man in dem Hauptgliede $a_1^{(1)} a_2^{(2)} \dots a_n^{(n)}$ die oberen Indices auf alle möglichen Arten permutiert, jedem der so gebildeten Produkte das positive oder negative Zeichen gibt, je nachdem die angewandte Permutation zur ersten oder zweiten Art gehört, und dann die Summe aller dieser Produkte nimmt.

In der Darstellung (1) von Δ werden durch die oberen Indices die Zeilen, durch die unteren Indices die Kolonnen gekennzeichnet, und demnach können wir diesem Satze auch den folgenden Ausdruck geben:

V. Eine Determinante ändert sich nicht, wenn die Zeilen zu Kolonnen und die Kolonnen zu Zeilen gemacht werden.

Wenn wir in den sämtlichen Anordnungen $\mathfrak{A}$, $\mathfrak{A}'$, $\mathfrak{A}''$ ... der n Elemente irgend zwei Elemente miteinander vertauschen, so bleibt die Gesamtheit dieser Anordnungen ungeändert, aber es geht jede Anordnung erster Art in eine Anordnung zweiter Art über und umgekehrt. Wenn wir also in den Gliedern M, M', M'' ..., aus denen Δ zusammengesetzt ist, irgend zwei untere Indices vertauschen, so geht jedes Glied mit positivem Zeichen in ein anderes über, das in Δ mit dem negativen Zeichen behaftet war und umgekehrt, also ändert Δ sein Vorzeichen. Daraus folgt mit Hilfe von V. der Satz:

VI. Wenn man zwei Zeilen oder zwei Kolonnen miteinander vertauscht, so ändert die Determinante nur ihr Vorzeichen,

und daraus allgemeiner:

VII. Wenn in einer Determinante die Zeilen oder die Kolonnen permutiert werden, so ändert sich der absolute Wert nicht, und das Vorzeichen ändert sich nicht oder geht in das entgegengesetzte

über, je nachdem die angewandte Permutation zur ersten oder zweiten Art gehört.

Aus VI. erhält man den folgenden Fundamentalsatz:

VIII. Wenn in zwei Zeilen oder in zwei Kolonnen die an gleicher Stelle stehenden Glieder einander gleich sind (kürzer ausgedrückt: wenn zwei Reihen einander gleich sind), so hat die Determinante den Wert Null.

Denn die Vertauschung der zwei Reihen ändert nach VI. das Zeichen, kann aber andererseits, da beide Reihen identisch sind, nichts ändern, so daß für Δ nur der Wert Null übrig bleibt.

§ 3.
Unterdeterminanten.

In jedem Gliede der entwickelten Determinante

$$\Sigma \pm a_1^{(1)} a_2^{(2)} \dots a_n^{(n)},$$

deren Wert wir jetzt mit A bezeichnen wollen, kommt jede der Zahlen $1, 2 \dots n$ ein- und nur einmal als unterer Index vor. Es wird also ein gewisser Komplex von Gliedern den Faktor $a_1^{(1)}$ enthalten, ein anderer Komplex den Faktor $a_1^{(2)}$ usf., endlich ein Komplex den Faktor $a_1^{(n)}$; jedes Glied der Determinante kommt in einem und nur in einem dieser Komplexe vor.

Bezeichnen wir also den ersten dieser Komplexe mit $a_1^{(1)} A_1^{(1)}$, den zweiten mit $a_1^{(2)} A_1^{(2)}$, den letzten mit $a_1^{(n)} A_1^{(n)}$, so können wir die Determinante folgendermaßen darstellen:

$$A = a_1^{(1)} A_1^{(1)} + a_1^{(2)} A_1^{(2)} + \dots + a_1^{(n)} A_1^{(n)}. \tag{1}$$

An Stelle des unteren Index 1 hätten wir ebensogut jeden anderen, ν, herausgreifen und daher

$$A = a_\nu^{(1)} A_\nu^{(1)} + a_\nu^{(2)} A_\nu^{(2)} + \dots + a_\nu^{(n)} A_\nu^{(n)} \tag{2}$$

setzen können. Darin bedeutet das Produkt $a_\nu^{(\mu)} A_\nu^{(\mu)}$ den Komplex aller Glieder der Determinante, die den Faktor $a_\nu^{(\mu)}$ enthalten.

Da dieselben Regeln wie für die unteren, so auch für die oberen Indices gelten, so kann man die Determinante auch noch in der folgenden Weise schreiben:

$$A = a_1^{(\mu)} A_1^{(\mu)} + a_2^{(\mu)} A_2^{(\mu)} + \dots + a_n^{(\mu)} A_n^{(\mu)}, \tag{3}$$

worin μ gleichfalls jeden der Indices $1, 2 \dots n$ bedeuten kann.

Die hierdurch vollständig definierten Größen $A_\nu^{(\mu)}$ heißen die Unterdeterminanten der Determinante A. Um ihre Bildungsweise genau kennen zu lernen, betrachten wir zunächst den Komplex $a_1^{(1)} A_1^{(1)}$. Man erhält ihn, wenn man in dem Produkt

$$a_1^{(1)} a_2^{(2)} \dots a_n^{(n)}$$

den unteren Index 1 ungeändert läßt und nur die übrigen Indices 2, 3 ... n auf alle Arten permutiert und die Summe der entstandenen Glieder mit Rücksicht auf die Zeichenregel bildet, d. h. es ist $A_1^{(1)}$ die $(n-1)$-reihige Determinante:

$$A_1^{(1)} = \begin{vmatrix} a_2^{(2)}, & a_3^{(2)} \dots & a_n^{(2)} \\ a_2^{(3)}, & a_3^{(3)} \dots & a_n^{(3)} \\ \dots\dots\dots & & \\ a_2^{(n)}, & a_3^{(n)} \dots & a_n^{(n)} \end{vmatrix} \tag{4}$$

oder die Determinante, die man aus A erhält, wenn man in dem A darstellenden Quadrat [§ 2, (1)] die erste Zeile und die erste Kolonne wegläßt.

Daraus ergibt sich leicht die Bedeutung von $A_\nu^{(\mu)}$; man kann, indem man $\nu - 1$ Zeilenvertauschungen vornimmt, die νte Zeile zur ersten machen, und wenn man noch $\mu - 1$ Vertauschungen der Kolonnen hinzunimmt, die μte Kolonne zur ersten; im übrigen bleiben die Reihen in ihrer Aufeinanderfolge ungeändert. Die Determinante selbst hat den Faktor $(-1)^{\mu+\nu}$ angenommen und ist dem absoluten Werte nach ungeändert geblieben. In der so umgeänderten Reihenfolge ist aber das Element $a_\nu^{(\mu)}$ an die Stelle des Elementes $a_1^{(1)}$ getreten, und daraus schließt man auf folgendes Bildungsgesetz:

Man erhält die Unterdeterminante $A_\nu^{(\mu)}$ dadurch, daß man in dem die Determinante darstellenden Quadrat die beiden Reihen wegläßt, die sich in $a_\nu^{(\mu)}$ kreuzen, und den Faktor $(-1)^{\nu+\mu}$ hinzufügt.

Wenn die Elemente a_ν^μ unabhängige Variable sind, so ergibt sich durch Differentiation von (1) oder (2) nach $a_\nu^{(\mu)}$:

$$A_\nu^{(\mu)} = \frac{\partial A}{\partial a_\nu^{(\mu)}}. \tag{5}$$

Da der untere Index ν in $A_\nu^{(\mu)}$ gar nicht vorkommt, so ändert sich $A_\nu^{(\mu)}$ nicht, wenn der untere Index ν durch einen anderen ersetzt wird. Dann aber verschwindet nach § 2, VII. die Determinante. Wir erhalten demnach aus (2) die folgende

wichtige Relation, in der μ, ν irgend zwei voneinander verschiedene Ziffern 1, 2 ... n sein können:

$$0 = a_\mu^{(1)} A_\nu^{(1)} + a_\mu^{(2)} A_\nu^{(2)} + \cdots + a_\mu^{(n)} A_\nu^{(n)}, \tag{6}$$

und ebenso bekommt man aus (3):

$$0 = a_1^{(\mu)} A_1^{(\nu)} + a_2^{(\mu)} A_2^{(\nu)} + \cdots + a_n^{(\mu)} A_n^{(\nu)}. \tag{7}$$

Wenn wir die Relation (6) mit einem beliebigen Faktor λ multiplizieren und zu (2) addieren, so erhalten wir die Formel:

$$A = (a_\nu^{(1)} + \lambda a_\mu^{(1)}) A_\nu^{(1)} + (a_\nu^{(2)} + \lambda a_\mu^{(2)}) A_\nu^{(2)} + \cdots + (a_\nu^{(n)} + \lambda a_\mu^{(n)}) A_\nu^{(n)}, \tag{8}$$

die uns den folgenden Satz ausdrückt:

IX. Die Determinante ändert ihren Wert nicht, wenn man zu den Elementen einer Zeile die mit einem beliebigen gemeinschaftlichen Faktor multiplizierten entsprechenden Elemente einer anderen Zeile addiert.

Wenn die oberen Indices wirkliche Potenzexponenten bedeuten, so ist der Wert der Determinante:

$$\Delta = \begin{vmatrix} 1, & a_1, & a_1^2 \ldots & a_1^{n-1} \\ 1, & a_2, & a_2^2 \ldots & a_2^{n-1} \\ \ldots\ldots\ldots & & \ldots\ldots\ldots & \\ 1, & a_n, & a_n^2 \ldots & a_n^{n-1} \end{vmatrix} \tag{9}$$

gleich

$$(a_2 - a_1)(a_3 - a_1) \ldots (a_n - a_1)$$
$$(a_3 - a_2) \ldots (a_n - a_2)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(a_n - a_{n-1}).$$

Man beweist dies nach IX., wenn man jede der $(n-1)$ ersten Spalten mit a_1 multipliziert von der folgenden subtrahiert und so fortfährt.

Wir suchen alle Glieder der entwickelten Determinante A auf, die den Faktor

$$a_1^{(1)} a_2^{(2)} \ldots a_\nu^{(\nu)}$$

enthalten, worin ν eine beliebige Zahl unter n sein kann.

Diese Glieder erhalten wir aus dem Hauptgliede

$$a_1^{(1)} a_2^{(2)} \ldots a_\nu^{(\nu)} a_{\nu+1}^{(\nu+1)} \ldots a_n^{(n)},$$

wenn wir bei der Permutation der unteren Indices 1, 2 ... ν ungeändert lassen und nur $\nu + 1$ ... n auf alle Arten permutieren, unter Berücksichtigung der Vorzeichenregel.

Demnach ist der Inbegriff der gesuchten Glieder

$$(10) \qquad a_1^{(1)} a_2^{(2)} \dots a_\nu^{(\nu)} \begin{vmatrix} a_{\nu+1}^{(\nu+1)} & \dots & a_n^{(\nu+1)} \\ \dots\dots\dots\dots & & \dots\dots \\ a_{\nu+1}^{(n)} & \dots & a_n^{(n)} \end{vmatrix}.$$

X. Die hier als Faktor auftretende Determinante von $n - \nu$ Reihen, die wir mit $A_{1,2,\dots,\nu}^{1,2,\dots,\nu}$ bezeichnen, entsteht aus A durch Weglassen der ν ersten Zeilen und Kolonnen.

Dieses Resultat wollen wir nun auf folgende Art verallgemeinern:

Wir wählen irgend ν Elemente

$$a_{\alpha_1}^{(\beta_1)}, \; a_{\alpha_2}^{(\beta_2)} \dots a_{\alpha_\nu}^{(\beta_\nu)}$$

aus, jedoch so, daß nicht zwei Elemente in derselben Zeile oder in derselben Kolonne vorkommen, d. h. so, daß nicht zweimal derselbe untere oder derselbe obere Index vorkommt, und bezeichnen den Inbegriff der Glieder der Determinante, die das Produkt dieser Elemente als Faktor enthalten, mit

$$(11) \qquad a_{\alpha_1}^{(\beta_1)} a_{\alpha_2}^{(\beta_2)} \dots a_{\alpha_\nu}^{(\beta_\nu)} A_{\alpha_1, \alpha_2 \dots \alpha_\nu}^{\beta_1, \beta_2 \dots \beta_\nu}.$$

Man kann durch Umstellen von Zeilen und Kolonnen, wodurch höchstens das Zeichen der Determinante geändert wird, immer erreichen, daß die Elemente

$$(12) \qquad a_{\alpha_1}^{(\beta_1)}, \; a_{\alpha_2}^{(\beta_2)} \dots a_{\alpha_\nu}^{(\beta_\nu)}$$

an die Stelle der Elemente

$$a_1^{(1)}, \; a_2^{(2)} \dots a_\nu^{(\nu)}$$

gelangen; dann aber läßt sich die Regel X auf die Bestimmung von $A_{\alpha_1, \alpha_2 \dots \alpha_\nu}^{\beta_1, \beta_2 \dots \beta_\nu}$ anwenden, und es ergibt sich:

XI. Man erhält (vom Vorzeichen abgesehen) $A_{\alpha_1, \alpha_2 \dots \alpha_\nu}^{\beta_1, \beta_2 \dots \beta_\nu}$ als $(n - \nu)$-reihige Determinante, wenn man in A alle Zeilen und Kolonnen wegläßt, die sich in einem der Elemente (12) schneiden, und die übrig bleibenden Zeilen und Kolonnen in ihrer Reihenfolge stehen läßt.

Für die Zeichenbestimmung aber ergibt sich folgende Vorschrift.

Man ordne die unteren und die oberen Indices $1, 2 \dots n$ in der Weise:

$$(13) \qquad \begin{matrix} \alpha_1, \alpha_2 \dots \alpha_\nu, \alpha_{\nu+1} \dots \alpha_n \\ \beta_1, \beta_2 \dots \beta_\nu, \beta_{\nu+1} \dots \beta_n, \end{matrix}$$

indem man $\alpha_{\nu+1} \dots \alpha_n$ und ebenso $\beta_{\nu+1} \dots \beta_n$ der Größe nach aufeinander folgend annimmt.

XII. Die in XI. beschriebene $(n-\nu)$-reihige Determinante erhält das positive oder negative Zeichen, je nachdem die beiden Anordnungen (13) der Ziffern $1, 2 \dots n$ beide zu derselben oder zu verschiedenen Arten gehören.

Denn die Determinante ändert ihr Zeichen durch jede Vertauschung zweier unterer oder zweier oberer Indices. Um den allgemeinen Fall (11) auf den besonderen Fall (10) zurückzuführen, hat man so viele Transpositionen oberer und unterer Indices vorzunehmen, daß die Permutationen (13) beide in die ursprüngliche Anordnung $1, 2, 3 \dots n$ übergehen, und ebenso viele Zeichenwechsel haben stattgefunden.

Die so definierten Größen

$$A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}$$

heißen die νten Unterdeterminanten oder Unterdeterminanten νter Ordnung. Sie sind dargestellt durch $(n-\nu)$-reihige Determinanten.

Aus XII. folgt in bezug auf diese Unterdeterminanten der Satz:

XIII. Die Unterdeterminante $A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}$ ändert nur ihr Vorzeichen, wenn zwei ihrer unteren oder zwei ihrer oberen Indices vertauscht werden, oder allgemeiner: sie bleibt dem absoluten Werte nach ungeändert, wenn die Anordnung der Indices $\alpha_1, \alpha_2 \dots \alpha_\nu$ durch irgend eine andere Anordnung ersetzt wird und ändert das Zeichen oder nicht, je nachdem diese Permutation zur zweiten oder zur ersten Art gehört.

Bezeichnen wir aber mit $\alpha'_1, \alpha'_2 \dots \alpha'_\nu$ irgend eine Anordnung der $\alpha_1, \alpha_2 \dots \alpha_\nu$, so enthält die Determinante A auch den Komplex der Glieder

$$\pm a^{(\beta_1)}_{\alpha'_1} a^{(\beta_2)}_{\alpha'_2} \dots a^{(\beta_\nu)}_{\alpha'_\nu} A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu},$$

und wenn wir also alle diese Glieder sammeln, so erhalten wir den Komplex:

$$A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu} \quad \Sigma \pm a^{(\beta_1)}_{\alpha_1} a^{(\beta_2)}_{\alpha_2} \dots a^{(\beta_\nu)}_{\alpha_\nu}. \tag{14}$$

Die hier auftretende ν-reihige Determinante

$$\Sigma \pm a^{(\beta_1)}_{\alpha_1} a^{(\beta_2)}_{\alpha_2} \dots a^{(\beta_\nu)}_{\alpha_\nu} = \begin{vmatrix} a^{(\beta_1)}_{\alpha_1}, & a^{(\beta_1)}_{\alpha_2} & \dots & a^{(\beta_1)}_{\alpha_\nu} \\ a^{(\beta_2)}_{\alpha_1}, & a^{(\beta_2)}_{\alpha_2} & \dots & a^{(\beta_2)}_{\alpha_\nu} \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ a^{(\beta_\nu)}_{\alpha_1}, & a^{(\beta_\nu)}_{\alpha_2} & \dots & a^{(\beta_\nu)}_{\alpha_\nu} \end{vmatrix}$$

wollen wir die zu $A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}$ komplementäre Unterdeterminante nennen und mit $B^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}$ bezeichnen. Sie enthält genau die Zeilen und Kolonnen, die in $A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}$ fehlen und stimmt, abgesehen vom Vorzeichen, mit der Unterdeterminante $(n - \nu)$ter Ordnung

$$A^{\beta_{\nu+1} \dots \beta_n}_{\alpha_{\nu+1} \dots \alpha_n}$$

überein. Der Komplex der Glieder (14) wird also bezeichnet mit

$$A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu} \quad B^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}. \tag{15}$$

Wählen wir nun für $\alpha_1, \alpha_2 \dots \alpha_\nu$ jede Kombination von ν der Ziffern $1, 2 \dots n$, so erhalten wir, indem wir $\beta_1, \beta_2 \dots \beta_\nu$ festhalten, ebenso viele Komplexe der Form (15) und jedes Glied der Determinante A kommt in einem und nur in einem dieser Komplexe vor. Demnach erhalten wir, wenn wir alle Ausdrücke (15) summieren, den Satz von Laplace.

XIV. Die Determinante A ist:

$$A = \sum^{\alpha} A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu} \quad B^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu}. \tag{16}$$

Selbstverständlich kann man auch die Kombination der α festhalten und in bezug auf die β summieren.

Dies ist der Satz von Laplace.

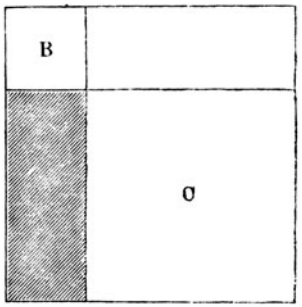
Als häufig vorkommender spezieller Fall mag der erwähnt werden, wo $\beta_1, \beta_2 \dots \beta_\nu = 1, 2 \dots \nu$ und wo $a^{(s)}_r = 0$ ist, wenn r einen der Werte $1, 2 \dots \nu$ und gleichzeitig s einen der Werte $\nu + 1 \dots n$ hat. Dann werden alle $A^{1, 2 \dots \nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu} = 0$, mit Ausnahme von $A^{1, 2 \dots \nu}_{1, 2 \dots \nu}$, und man erhält

$$A = A^{1, 2 \dots \nu}_{1, 2 \dots \nu} \quad B^{1, 2 \dots \nu}_{1, 2 \dots \nu}.$$

Der Satz wird durch folgendes Diagramm veranschaulicht, in dem das ganze Quadrat die Determinante A vorstellt, und wo in dem schraffierten Rechteck lauter Nullen stehen. Sind dann B und C die in den beiden kleinen Quadraten stehenden Determinanten von ν und $n - \nu$ Reihen, so ist

(17) $$A = BC.$$

Fig. 1.



Noch eine andere Darstellung der Determinante A durch die ersten und zweiten Unterdeterminanten erhält man auf folgende Weise.

Man wähle in A irgend zwei Reihen aus, die sich in einem Elemente, etwa in $a_\nu^{(\mu)}$, schneiden. In jedem Gliede von A kommt ein Element mit dem unteren Index ν und ein Element mit dem oberen Index μ vor. Wir haben also zunächst in A den Komplex $a_\nu^{(\mu)} A_\nu^{(\mu)}$ und ferner die verschiedenen Komplexe $a_\nu^{(i)} a_k^{(\mu)} A_{\nu,k}^{i,\mu}$, worin i jeden von μ verschiedenen und k jeden von ν verschiedenen Index bedeuten kann.

XV. Wir können daher setzen:

$$A = a_\nu^{(\mu)} A_\nu^{(\mu)} + \sum^{i,k} a_\nu^{(i)} a_k^{(\mu)} A_{\nu,k}^{i,\mu}$$

oder nach XIII.

(18) $$A = a_\nu^{(\mu)} A_\nu^{(\mu)} - \sum^{i,k} a_\nu^{(i)} a_k^{(\mu)} A_{\nu,k}^{\mu,i}.$$

Wir bemerken zu diesem Satze noch, daß $A_{\nu,k}^{\mu,i}$ die dem Elemente $a_k^{(i)}$ entsprechende erste Unterdeterminante der $(n-1)$-reihigen Determinante $A_\nu^{(\mu)}$ ist; denn $A_{\nu,k}^{\mu,i}$ ist der Koeffizient von $a_\nu^{(\mu)} a_k^{(i)}$ in der Entwickelung von A und $A_\nu^{(\mu)}$ der Koeffizient von $a_\nu^{(\mu)}$, folglich $A_{\nu,k}^{\mu,i}$ der Koeffizient von $a_k^{(i)}$ in der Determinante $A_\nu^{(\mu)}$.

Man kann nach diesem Satze die sogenannte geränderte Determinante

(19) $$U = \begin{vmatrix} a_1^{(1)}, & a_2^{(1)} & \dots & a_n^{(1)}, & u_1 \\ a_1^{(2)}, & a_2^{(2)} & \dots & a_n^{(2)}, & u_2 \\ \dots\dots\dots\dots & & & \dots\dots & \\ a_1^{(n)}, & a_2^{(n)} & \dots & a_n^{(n)}, & u_n \\ v_2, & v_2 & \dots & v_n, & q \end{vmatrix}$$

nach den Elementen der letzten Zeile und Kolonne entwickeln und erhält

$$U = qA - \sum^{i} \sum^{k} u_i v_k A_k^{(i)}. \tag{20}$$

Man erhält diese Gleichung aus (18), wenn man n in $n + 1$ verwandelt, und die Elemente der letzten Zeile und Kolonne durch eine andere Bezeichnung auszeichnet.

Ersetzt man q durch $q - u$, worin u wieder eine beliebige Größe ist, so ergibt sich die oft angewandte Formel:

$$\begin{vmatrix} a_1^{(1)}, & a_2^{(1)} & \dots & a_n^{(1)}, & u_1 \\ a_1^{(2)}, & a_2^{(2)} & \dots & a_n^{(2)}, & u_2 \\ \dots\dots & \dots\dots & \dots & \dots\dots & \dots\dots \\ a_1^{(n)}, & a_2^{(n)} & \dots & a_n^{(n)}, & u_n \\ v_1, & v_2 & \dots & v_n, & q - u \end{vmatrix} = U - Au, \tag{21}$$

worin U die Bedeutung (19) hat.

Auch bei den höheren Unterdeterminanten ist bisweilen die Bezeichnung durch Differentialquotienten zweckmäßig, so daß z. B.

$$A_{\nu,k}^{i,\mu} = \frac{\partial^2 A}{\partial a_\nu^{(i)} \partial a_k^{(\mu)}} \tag{22}$$

gesetzt wird.

§ 4.

Lineare homogene Gleichungen.

Die hauptsächlichste Anwendung der Determinanten ist die Auflösung linearer Gleichungen.

Wir wollen hier die Aufgabe gleich in allgemeinster Weise in Angriff nehmen, da die spezielle Form kaum eine Vereinfachung ist und sich nachher leicht aus dem allgemeinen Resultate ableiten läßt.

Wir betrachten ein System von m Gleichungen ersten Grades, in denen n Unbekannte $x_1, x_2 \dots x_n$ homogen vorkommen:

$$\begin{aligned} a_1^{(1)} x_1 + a_2^{(1)} x_2 + \dots + a_n^{(1)} x_n &= 0 \\ a_1^{(2)} x_1 + a_2^{(2)} x_2 + \dots + a_n^{(2)} x_n &= 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ a_1^{(m)} x_1 + a_2^{(m)} x_2 + \dots + a_n^{(m)} x_n &= 0, \end{aligned} \tag{1}$$

worin die Koeffizienten $a_i^{(k)}$ als gegebene Größen betrachtet werden. Über die Zahlen m, n wollen wir vorläufig noch gar

keine Voraussetzung machen, sondern uns allgemein die Aufgabe stellen, alle Wertsysteme der $x_1, x_2 \ldots x_n$ zu ermitteln, die den Gleichungen (1) genügen.

Eine Lösung der Gleichungen (1) können wir sofort angeben: sie sind nämlich, was auch die Koeffizienten $a_i^{(k)}$ sein mögen, erfüllt, wenn

$$x_1 = 0,\ x_2 = 0,\ \ldots x_n = 0. \tag{2}$$

Einen anderen extremen Fall können wir noch erwähnen; wenn nämlich die Koeffizienten $a_i^{(k)}$ sämtlich den Wert Null haben, dann sind die Gleichungen (1) für beliebige Werte von $x_1, x_2 \ldots x_n$ befriedigt.

Der allgemeinen Beantwortung der Frage schicken wir folgende Bemerkungen voraus.

Wir schreiben das System der Koeffizienten von (1) in Form eines Rechtecks

$$\begin{array}{l} a_1^{(1)},\ a_2^{(1)} \ldots a_n^{(1)} \\ a_1^{(2)},\ a_2^{(2)} \ldots a_n^{(2)} \\ \ldots\ldots\ldots\ldots \\ a_1^{(m)},\ a_2^{(m)} \ldots a_n^{(m)}. \end{array} \tag{3}$$

Ein solches Schema, das für sich noch keine numerische Bedeutung hat, heißt eine Matrix und ist die Quelle einer größeren Anzahl von Determinanten.

Die der Matrix entstammenden Determinanten erhält man, wenn man beliebige Zeilen und Kolonnen wegläßt, in beliebiger, nur insoweit bestimmter Anzahl, daß die übrig bleibenden Elemente ein Quadrat bilden, und dieses Quadrat als Determinante auffaßt.

So erhält man aus der Matrix einreihige, zweireihige usf. Determinanten. Die höchsten Determinanten sind n- oder m-reihig, je nachdem n oder m die kleinere Zahl ist (oder n-reihig, wenn $n = m$ ist).

Wir machen nun die Annahme, daß unter den ν-reihigen Determinanten der Matrix wenigstens eine von Null verschieden sei, während die $(\nu+1)$-reihigen und folglich auch die höheren Determinanten, falls solche vorhanden sind, alle verschwinden sollen. ν kann jede Zahl sein, die nicht größer als die kleinere der beiden Zahlen n oder m ist (oder falls $n = m$ ist, diesen gemeinschaftlichen Wert nicht übertrifft).

Eine solche Zahl ν wird sich immer finden lassen, wenn wir den schon erledigten, ganz interesselosen Fall ausschließen, daß alle Koeffizienten $a_i^{(k)}$ verschwinden.

Wir können, ohne die Allgemeinheit zu beschränken, zur Vereinfachung der Bezeichnung annehmen, die nicht verschwindende ν-reihige Determinante sei

$$A = \begin{vmatrix} a_1^{(1)}, & a_2^{(1)} & \dots & a_\nu^{(1)} \\ a_1^{(2)}, & a_2^{(2)} & \dots & a_\nu^{(2)} \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ a_1^{(\nu)}, & a_2^{(\nu)} & \dots & a_\nu^{(\nu)} \end{vmatrix}. \tag{4}$$

Denn offenbar steht es uns frei, das Gleichungssystem (1) in beliebiger Weise anzuordnen, und ferner können wir die Bezeichnung der Unbekannten x so wählen, daß irgend ν von ihnen die ν ersten sind.

Die Unterdeterminanten von A bezeichnen wir wie früher mit $A_i^{(k)}$, worin i, k von 1 bis ν gehen.

Wenn nun zunächst $\nu = n$ ist, was voraussetzt, daß m nicht kleiner als n ist, so haben die Gleichungen (1) keine andere Lösung, als die in den Gleichungen (2) enthaltene.

Denn greifen wir die n ersten der Gleichungen (1) heraus:

$$\begin{aligned} a_1^{(1)} x_1 + a_2^{(1)} x_2 + \dots + a_n^{(1)} x_n &= 0 \\ a_1^{(2)} x_1 + a_2^{(2)} x_2 + \dots + a_n^{(2)} x_n &= 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ a_1^{(n)} x_1 + a_2^{(n)} x_2 + \dots + a_n^{(n)} x_n &= 0, \end{aligned} \tag{5}$$

multiplizieren diese der Reihe nach mit $A_\mu^{(1)}, A_\mu^{(2)} \dots A_\mu^{(n)}$, worin μ jeder der Indices $1, 2 \dots n$ sein kann, und addieren sie, so folgt, weil nach § 3, (2) und (6)

$$\sum_{1,\nu}^{i} A_\mu^{(i)} a_\lambda^{(i)} = 0 \text{ oder } = A$$

ist, je nachdem λ von μ verschieden ist oder nicht,

$$A x_\mu = 0,$$

und da nach unserer Voraussetzung A von Null verschieden ist,

$$x_\mu = 0.$$

Damit ist bewiesen:

XVI. Wenn ein System von m linearen homogenen Gleichungen mit n Unbekannten eine Lösung hat, bei der nicht alle Unbekannte verschwinden, so müssen, wenn $m \gtreqless n$ ist, sämtliche n-reihige

Unterdeterminanten der Matrix (3) der Koeffizienten verschwinden.

Wir heben den am meisten angewendeten besonderen Fall $m = n$ hervor und geben dem Satze für diesen Fall den folgenden Ausdruck:

XVII. Wenn ein System von n linearen homogenen Gleichungen mit n Unbekannten eine von Null verschiedene Determinante hat, so haben sämtliche Unbekannte den Wert Null, oder:

Wenn ein System von n linearen Gleichungen mit ebenso vielen homogen vorkommenden Unbekannten eine Lösung hat, bei der nicht alle Unbekannten verschwinden, so verschwindet die Determinante des Systems.

Unter der Determinante eines Systems von n linearen homogenen Gleichungen mit n Unbekannten ist hier die Determinante aus den n^2 Koeffizienten dieser Gleichungen verstanden.

Wir betrachten ferner den Fall, daß ν kleiner als n ist. Da m gleich oder größer als ν sein muß, so wählen wir die ν ersten Gleichungen des Systems (1) und schreiben sie so:

$$(6)\quad \begin{aligned} a_1^{(1)}x_1 + a_2^{(1)}x_2 + \cdots + a_\nu^{(1)}x_\nu &= -a_{\nu+1}^{(1)}x_{\nu+1} - \cdots - a_n^{(1)}x_n \\ a_1^{(2)}x_1 + a_2^{(2)}x_2 + \cdots + a_\nu^{(2)}x_\nu &= -a_{\nu+1}^{(2)}x_{\nu+1} - \cdots - a_n^{(2)}x_n \\ &\cdots\cdots\cdots\cdots \\ a_1^{(\nu)}x_1 + a_2^{(\nu)}x_2 + \cdots + a_\nu^{(\nu)}x_\nu &= -a_{\nu+1}^{(\nu)}x_{\nu+1} - \cdots - a_n^{(\nu)}x_n. \end{aligned}$$

Wir bezeichnen wieder mit μ einen der Indices $1, 2 \ldots \nu$, multiplizieren dann die Gleichungen (6) der Reihe nach mit $A_\mu^{(1)}, A_\mu^{(2)} \ldots A_\mu^{(\nu)}$ und addieren sie. Daraus folgt, wie vorhin, mit Benutzung von § 3, (2), (6):

$$(7)\quad A x_\mu = -C_{\nu+1,\mu}x_{\nu+1} - \cdots - C_{n,\mu}x_n,$$

wenn zur Abkürzung gesetzt ist:

$$(8)\quad C_{\lambda,\mu} = \sum_{1,\nu}^{i} a_\lambda^{(i)} A_\mu^{(i)}, \qquad \lambda = \nu+1, \nu+2 \ldots n.$$

Nach § 3, (2) ist $C_{\lambda,\mu}$ die Determinante, die aus der durch (4) definierten Determinante dadurch hervorgeht, daß man die Elemente der μten Kolonne $a_\mu^{(1)}, a_\mu^{(2)} \ldots a_\mu^{(\nu)}$ durch $a_\lambda^{(1)}, a_\lambda^{(2)} \ldots a_\lambda^{(\nu)}$ ersetzt.

Durch (7) sind nun, da A von Null verschieden ist, die $x_1, x_2, \ldots x_\nu$ linear ausgedrückt durch $x_{\nu+1}, \ldots x_n$ und durch die bekannten Größen.

Es ist nun noch zu zeigen, daß durch die Ausdrücke (7) die Gleichungen (1) befriedigt sind, welche Werte auch $x_{\nu+1}, \ldots x_n$ haben mögen, daß also $n - \nu$ von den Unbekannten willkürlich bleiben, von denen die übrigen ν nach (7) abhängig sind.

Um die Wahrheit dieses Satzes einzusehen, haben wir nur die Ausdrücke (7) in die Gleichungen (1) einzusetzen. Man vereinfacht die Rechnung sehr durch Anwendung eines Summenzeichens Σ, bei dem wir, wie schon oben, die Summationsbuchstaben oben, die Grenzen unten anhängen. Zunächst können wir dann die Gleichungen (7) so schreiben:

$$(9) \qquad A x_\mu = -\sum_{\nu+1,n}^{h} \sum_{1,\nu}^{i} a_h^{(i)} A_\mu^{(i)} x_h, \qquad \mu = 1, 2, \ldots \nu.$$

Wir multiplizieren, wenn k irgend eine der Ziffern $1, 2, \ldots m$ bedeutet, mit $a_\mu^{(k)}$ und summieren in bezug auf μ:

$$(10) \qquad A \sum_{1,\nu}^{\mu} a_\mu^{(k)} x_\mu = -\sum_{\nu+1,n}^{h} x_h \sum_{1,\nu}^{i} \sum_{1,\nu}^{\mu} a_h^{(i)} a_\mu^{(k)} A_\mu^{(i)}, \qquad k = 1, 2, \ldots m.$$

Dazu addieren wir beiderseits die Summe

$$A \sum_{\nu+1,n}^{h} a_h^{(k)} x_h$$

und erhalten

$$(11) \qquad A \sum_{1,n}^{\mu} a_\mu^{(k)} x_\mu = \sum_{\nu+1,n}^{h} x_h \left(A a_h^{(k)} - \sum_{1,\nu}^{i} \sum_{1,\nu}^{\mu} a_h^{(i)} a_\mu^{(k)} A_\mu^{(i)} \right).$$

Der Faktor von x_h in der Summe auf der rechten Seite ist nach § 3, (20) die Determinante

$$(12) \qquad \begin{vmatrix} a_1^{(1)}, & a_2^{(1)} \ldots a_\nu^{(1)}, & a_h^{(1)} \\ a_1^{(2)}, & a_2^{(2)} \ldots a_\nu^{(2)}, & a_h^{(2)} \\ \cdots\cdots\cdots\cdots & & \\ a_1^{(\nu)}, & a_2^{(\nu)} \ldots a_\nu^{(\nu)}, & a_h^{(\nu)} \\ a_1^{(k)}, & a_2^{(k)} \ldots a_\nu^{(k)}, & a_h^{(k)} \end{vmatrix},$$

und verschwindet daher, wenn $k \lessgtr \nu$ ist, dann, weil zwei Zeilen übereinstimmen, wenn aber $k > \nu$ ist, nach der Voraussetzung, weil dann (12) eine $\nu + 1$-reihige Determinante der Matrix (3) ist. Wir bekommen also aus (11), da A von Null verschieden ist,

$$(13) \qquad \sum_{1,n}^{\mu} a_{\mu}^{(k)} x_{\mu} = 0, \qquad k = 1, 2 \ldots m,$$

d. h. das System der Gleichungen (1) ist durch (7) befriedigt.

Damit ist bewiesen:

XVIII. Wenn in einem System von m linearen homogenen Gleichungen mit n Unbekannten alle $\nu+1$-reihigen Unterdeterminanten der Matrix der Koeffizienten verschwinden, so hat das System eine Lösung, in der mindestens $n - \nu$ von der Unbekannten willkürlich bleiben. Ist $m < n$, so gibt es immer eine Lösung, in der mindestens $n - m$ der Unbekannten willkürlich bleiben;

und hiervon ist ein häufig vorkommender spezieller Fall:

XIX. Wenn die Determinante eines Systems von n linearen homogenen Gleichungen mit ebensoviel Unbekannten verschwindet, so können die Gleichungen so befriedigt werden, daß nicht alle Unbekannte verschwinden.

Wir wollen von dem so bewiesenen Satze noch den anderen Fall hervorheben, daß $m = n - 1$ und $\nu = n - 1$ ist. In diesem Falle bleibt nur eine der Unbekannten beliebig, und die Verhältnisse der Unbekannten sind völlig bestimmt. Wir können diesem Resultate folgenden Ausdruck geben:

Bezeichnen wir die $(n-1)$-reihigen Determinanten der Matrix

$$(14) \qquad \begin{matrix} a_1^{(1)}, & a_2^{(1)} & \ldots a_n^{(1)} \\ a_1^{(2)}, & a_2^{(2)} & \ldots a_n^{(2)} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_1^{(n-1)}, & a_2^{(n-1} & \ldots a_n^{(n-1)}, \end{matrix}$$

mit abwechselndem Vorzeichen genommen durch

$$A_1, A_2, \ldots A_n$$

und nehmen an, daß wenigstens eine von diesen Größen von Null verschieden sei, so ist die Lösung des Systems:

$$(15) \qquad \begin{matrix} a_1^{(1)} x_1 & + a_2^{(1)} x_2 & + \cdots + a_n^{(1)} x_n & = 0 \\ a_1^{(2)} x_1 & + a_2^{(2)} x_2 & + \cdots + a_n^{(2)} x_n & = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_1^{(n-1)} x_1 + & a_2^{(n-1)} x_2 & + \cdots + a_n^{(n-1)} x_n & = 0 \end{matrix}$$

gegeben durch die Verhältnisse

$$x_1 : x_2 : \cdots : x_n = A_1 : A_2 : \cdots : A_n. \tag{16}$$

Von den Sätzen ist eine Anwendung die, daß ein Ausdruck der Form

$$F(x) = A_0 x^{n-1} + A_1 x^{n-2} + \cdots + A_{n-2} x + A_{n-1}$$

nicht für n verschiedene Werte $a_1, a_2, \ldots a_n$ von x verschwinden kann, außer wenn die Koeffizienten $A_0, A_1, \ldots A_n$ alle gleich Null sind.

Denn setzen wir die n Gleichungen an:

$$F(a_1) = 0, \quad F(a_2) = 0, \quad \ldots F(a_n) = 0,$$

so ist deren Determinante nach § 3 (9) nicht $= 0$, und daraus folgt, daß die $A_0, A_1, \ldots A_{n-1} = 0$ sein müssen.

§ 5.

Lineare Substitutionen.

Ein Ausdruck von der Form

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

heißt eine lineare Funktion von $x_1, x_2, \ldots x_n$; m solcher Funktionen

$$y_k = \sum_{1,n}^{i} a_i^{(k)} x_i, \qquad k = 1, 2 \ldots m \tag{1}$$

heißen linear unabhängig, wenn eine identische, d. h. für variable x gültige Relation

$$\lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_m y_m = 0 \tag{2}$$

mit konstantem λ nicht anders als für $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_m = 0$ bestehen kann, im anderen Falle linear abhängig. Die Gleichung (2) fordert aber für die $a_i^{(k)}$ die Gleichungen

$$\sum_{1,m}^{k} \lambda_k a_i^{(k)} = 0, \qquad i = 1, 2 \ldots n. \tag{3}$$

Die Sätze XVI, XVII, § 4 (in denen m und n zu vertauschen sind) geben uns also die Bedingungen für die lineare Abhängigkeit oder Unabhängigkeit. Insbesondere erhält man die Sätze:

XX. Ist $m > n$, so sind die m Funktionen (1) immer linear abhängig. Ist $m = n$, so sind diese Funk-

tionen dann und nur dann linear abhängig, wenn die Determinante

$$A = \Sigma \pm a_1^{(1)} a_2^{(2)} \dots a_n^{(n)}$$

verschwindet. Sind die Funktionen also linear unabhängig, so folgt aus § 3, (2), (6)

$$A x_i = \sum_{1,n}^{k} y_k A_i^{(k)} \tag{4}$$

und die x_i sind also auch lineare Funktionen von $y_1, y_2 \dots y_n$.

Sind die x_i Variable, so werden auch die y_i Variable, und ein System von n unabhängigen Gleichungen (1) bildet eine lineare Substitution, durch die die Veränderlichen y_i durch die Veränderlichen x_i ausgedrückt werden. Die Gleichungen (4) drücken dann umgekehrt die Variablen x_i durch die Variablen y_i aus.

Aus dem Koeffizientensystem einer solchen Substitution kann man eine Matrix bilden, die man so bezeichnet:

$$A = \begin{pmatrix} a_1^{(1)}, & a_2^{(1)} \dots & a_n^{(1)} \\ a_1^{(2)}, & a_2^{(2)} \dots & a_n^{(2)} \\ \dots\dots\dots & & \\ a_1^{(n)}, & a_2^{(n)} \dots & a_n^{(n)} \end{pmatrix} \tag{5}$$

oder kürzer mit $(a_i^{(k)})$, und dann wird die lineare Substitution auch so bezeichnet:

$$(y_1, y_2 \dots y_n) = A(x_1, x_2 \dots x_n), \tag{6}$$

oder kürzer:

$$(y) = A(x). \tag{7}$$

Die Gleichungen:

$$y_1 = x_1, y_2 = x_2 \dots y_n = x_n$$

bilden demnach eine lineare Substitution, deren Matrix

$$J = \begin{pmatrix} 1, 0 \dots 0 \\ 0, 1 \dots 0 \\ \dots\dots\dots \\ 0, 0 \dots 1 \end{pmatrix} \tag{8}$$

ist und die die identische Substitution genannt wird.

Wenn $n = 1, 2, 3, 4$ ist, so heißt die Substitution unär, binär, ternär, quaternär. Im allgemeinen heißt n die Dimension der Substitution.

Nehmen wir zwei lineare Substitutionen von folgender Form:

$$y_k = \sum_{1,n}^{s} a_s^{(k)} x_s, \qquad (y) = A(x), \tag{9}$$

$$z_i = \sum_{1,n}^{s} b_s^{(i)} y_s, \qquad (z) = B(y), \tag{10}$$

worin i und k ebenso wie s von 1 bis n gehen, und führen für y_s aus (9) die Werte in (10) ein, so werden die z durch die x ausgedrückt mittels einer neuen linearen Substitution C, die wir so bezeichnen können:

$$z = C(x) = BA(x). \tag{11}$$

Schreiben wir diese ausführlicher:

$$z_k = \sum_{1,n}^{h} c_h^{(k)} x_h,$$

so ist:

$$c_h^{(k)} = \sum_{1,n}^{s} b_s^{(k)} a_h^{(s)} \tag{12}$$

und die Substitution $C = BA$ heißt aus B und A zusammengesetzt oder komponiert. Obwohl wir hier das Zeichen der Multiplikation brauchen, so ist doch zwischen BA und AB zu unterscheiden, und in dem besonderen Falle, wo $AB = BA$ ist, heißen die Substitutionen kommutativ oder vertauschbar.

Um drei Substitutionen derselben Dimension C, B, A zusammenzusetzen, muß man die Ausdrücke (11) für z in eine neue lineare Substitution

$$(u) = C(z) \tag{13}$$

einführen und die Variablen u durch die x ausdrücken:

$$(u) = CBA(x). \tag{14}$$

Da es offenbar gleichgültig ist, ob man die Ausdrücke (11) in (13) einführt, oder ob man zuerst z nach (10) durch y und dann y nach (9) durch x ausdrückt, so gilt für diese Komposition das assoziative Gesetz:

$$C(BA) = (CB)A = CBA, \tag{15}$$

d. h. man kann, ohne eine Undeutlichkeit befürchten zu müssen, die Klammer ganz weglassen. Dies läßt sich auch leicht durch Rechnung bestätigen, wenn man nach (12) die Elemente von $C(BA)$ und $(CB)A$ bildet. Man findet für beide den Ausdruck:

$$\sum^{i} \sum^{h} c_i^{(k)} b_h^{(i)} a_l^{(h)}.$$

Wir sprechen also den Satz aus:

XXI. Bei der Zusammensetzung der Matrizen gilt das assoziative, aber nicht immer das kommutative Gesetz.

§ 6.

Multiplikation der Determinanten.

Wenn wir die Determinante der zusammengesetzten Substitution bilden, so kommen wir auf das Multiplikationsgesetz der Determinanten, das wir wohl am einfachsten nach Hesse ableiten.

Wir bilden das Hauptglied der Determinante C, indem wir mit $s_1, s_2 \dots s_n$ voneinander unabhängige Summationsbuchstaben bezeichnen, die von 1 bis n laufen:

$$c_1^{(1)} c_2^{(2)} \dots c_n^{(n)} = \sum^{s_1} b_{s_1}^{(1)} a_1^{(s_1)} \sum^{s_2} b_{s_2}^{(2)} a_2^{(s_2)} \dots \sum^{s_n} b_{s_n}^{(n)} a_n^{(s_n)}$$

$$= \sum^{s_1, s_2, \dots s_n} b_{s_1}^{(1)} b_{s_2}^{(2)} \dots b_{s_n}^{(n)} a_1^{(s_1)} a_2^{(s_2)} \dots a_n^{(s_n)}.$$

Die Permutation der unteren Indices der c entspricht der Permutation der unteren Indices der a, und wenn man also mit Rücksicht auf die Vorzeichenregel die Determinante $|C|$ bildet, so ergibt sich:

$$\Sigma \pm c_1^{(1)} c_2^{(2)} \dots c_n^{(n)}$$

$$= \sum^{s_1, s_2, \dots s_n} b_{s_1}^{(1)} b_{s_2}^{(2)} \dots b_{s_n}^{(n)} \Sigma \pm a_1^{(s_1)} a_2^{(s_2)} \dots a_n^{(s_n)}. \tag{1}$$

Nun ist

$$\Sigma \pm a_1^{(s_1)} a_2^{(s_2)} \dots a_n^{(s_n)}$$

nach VII und VIII, § 2, immer dann gleich Null, wenn unter den $s_1, s_2 \dots s_n$ zweimal dieselbe Ziffer vorkommt, und gleich $+A$ oder gleich $-A$, je nachdem $s_1, s_2, \dots s_n$ eine Anordnung erster oder zweiter Art der Ziffern $1, 2 \dots n$ bildet.

Demnach wird die rechte Seite von (1):

$$A \Sigma \pm b_{s_1}^{(1)} b_{s_2}^{(2)} \dots b_{s_n}^{(n)},$$

woraus sich ergibt:

$$C = AB, \tag{2}$$

also das Multiplikationsgesetz der Determinanten.

XXII. Das Produkt von zwei Determinanten desselben Grades ist eine Determinante desselben Grades, deren Elemente man findet, indem man die Elemente einer Zeile des ersten Faktors mit den entsprechenden Elementen einer Kolonne des zweiten Faktors multipliziert und die Produkte addiert.

Hierdurch ist auch das Gesetz für die Komposition der Substitutionen oder, was dasselbe ist, der Matrizen ausgedrückt. Während aber bei den Matrizen die Reihenfolge der Faktoren wesentlich ist, so ist diese bei den Determinantenprodukten gleichgültig, und da man in den Determinanten die Zeilen zu Kolonnen und die Kolonnen zu Zeilen machen kann, so läßt sich ein Determinantenprodukt auf vier Arten bilden.

Auf dem gleichen Wege ergibt sich eine Verallgemeinerung des Multiplikationstheorems.

Ist $m > n$, so setzen wir:

$$c_h^{(k)} = \sum_{1,m}^{s} b_s^{(k)} a_h^{(s)}. \tag{3}$$

Wir lassen die Summationsbuchstaben $s_1, s_2, \ldots s_n$ von 1 bis m laufen und bilden die Determinante C wie oben:

$$\Sigma \pm c_1^{(1)} c_2^{(2)} \ldots c_n^{(n)}$$

$$= \sum^{s_1, s_2, \ldots s_n} b_{s_1}^{(1)} b_{s_2}^{(2)} \ldots b_{s_n}^{(n)} \Sigma \pm a_1^{(s_1)} a_2^{(s_2)} \ldots a_n^{(s_n)}. \tag{4}$$

Wählen wir unter den Ziffern $1, 2 \ldots m$ irgend eine Kombination von n verschiedenen Ziffern $\alpha_1, \alpha_2 \ldots \alpha_n$ aus, so enthält die Summe (4) alle Produkte von der Form:

$$\Sigma + b_{\alpha_1}^{(1)} b_{\alpha_2}^{(2)} \ldots b_{\alpha_n}^{(n)} \cdot \Sigma \pm a_1^{(\alpha_1)} a_2^{(\alpha_2)} \ldots a_n^{(\alpha_n)},$$

aber keine anderen Glieder, und die Determinante C ist also gleich der Summe aller dieser Produkte. Da wir überdies die Zeilen mit den Kolonnen, d. h. die oberen mit den unteren Indices vertauschen können, so haben wir den Satz:

XXIII. Wenn die Elemente der Determinante

$$C = \Sigma + c_1^{(1)} c_2^{(2)} \ldots c_n^{(n)}$$

die Form (3) haben, worin $m > n$ ist, so wähle man unter den m Ziffern $1, 2 \ldots m$ auf alle möglichen Arten n verschiedene aus, die in einer be-

liebigen Reihenfolge mit $\alpha_1, \alpha_2, \ldots \alpha_n$ bezeichnet werden. Ist dann

$$A_\alpha = \Sigma \pm a_1^{(\alpha_1)} a_2^{(\alpha_2)} \ldots a_n^{(\alpha_n)}$$
$$B_\alpha = \Sigma \pm b_1^{(\alpha_1)} b_2^{(\alpha_2)} \ldots b_n^{(\alpha_n)}, \tag{5}$$

so ist

$$C = \Sigma A_\alpha B_\alpha, \tag{6}$$

wenn sich die Summe auf alle möglichen Kombinationen der α erstreckt.

Die Anzahl der Glieder dieser Summe ist gleich der Anzahl der Kombinationen ohne Wiederholung von m Elementen zur $n$ten Klasse, also:

$$\frac{m(m-1)(m-2)\ldots(m-n+1)}{1.2.3\ldots n}.$$

§ 7.

Determinanten der Unterdeterminanten.

Wir machen hier gleich eine Anwendung von dem Multiplikationsgesetz der Determinanten.

Es sei A eine Determinante:

$$A = \begin{vmatrix} a_1^{(1)}, & a_2^{(1)} & \ldots & a_n^{(1)} \\ a_1^{(2)}, & a_2^{(2)} & \ldots & a_n^{(2)} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ a_1^{(n)}, & a_2^{(n)} & \ldots & a_n^{(n)} \end{vmatrix} \tag{1}$$

und

$$\begin{matrix} A_1^{(1)}, & A_2^{(1)} & \ldots & A_n^{(1)} \\ A_1^{(2)}, & A_2^{(2)} & \ldots & A_n^{(2)} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ A_1^{(n)}, & A_2^{(n)} & \ldots & A_n^{(n)} \end{matrix} \tag{2}$$

das System der Unterdeterminanten. Bilden wir aus (2) die Determinante, die wir mit Δ bezeichnen wollen, so können wir auf das Produkt $A\Delta$ die Multiplikationsregel anwenden. Dies gibt aber nach § 3 (3) und (7):

$$A\Delta = \begin{vmatrix} A, & 0 & \ldots & 0 \\ 0, & A & \ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ 0, & 0 & \ldots & A \end{vmatrix} = A^n,$$

und daraus durch Division mit A:

$$\Delta = A^{n-1}. \tag{3}$$

Es ist also Δ die $(n-1)$te Potenz von A. Bei dieser Ableitung ist allerdings zunächst vorausgesetzt, daß A von Null verschieden sei. Da aber (3) in bezug auf die Elemente $a_i^{(h)}$ eine Identität ist, d. h. auch dann gilt, wenn diese Größen unabhängige Variable sind, so folgt, daß auch noch in diesem Ausnahmefall die Formel (3) gilt, d. h. daß, wenn A verschwindet, auch Δ verschwindet.

Dies Ergebnis ist ein spezieller Fall eines allgemeineren Satzes, nach dem jede beliebige Determinante der Matrix (2) gebildet werden kann. Betrachten wir die ν-reihige Unterdeterminante

$$\Delta_\nu = \begin{vmatrix} A_1^{(1)}, & A_2^{(1)} & \dots & A_\nu^{(1)} \\ A_1^{(2)}, & A_2^{(2)} & \dots & A_\nu^{(2)} \\ \dots & \dots & \dots & \dots \\ A_1^{(\nu)}, & A_2^{(\nu)} & \dots & A_\nu^{(\nu)} \end{vmatrix}, \tag{4}$$

aus der man durch Permutation der oberen und unteren Indices alle anderen ν-reihigen Unterdeterminanten ableiten kann, so kann man die Multiplikationsregel anwenden, indem man Δ_ν in eine n-reihige Determinante verwandelt:

$$\Delta_\nu = \begin{vmatrix} A_1^{(1)}, & A_2^{(1)} & \dots & A_\nu^{(1)}, & A_{\nu+1}^{(1)} & \dots & A_n^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_1^{(\nu)}, & A_2^{(\nu)} & \dots & A_\nu^{(\nu)}, & A_{\nu+1}^{(\nu)} & \dots & A_n^{(\nu)} \\ 0, & 0 & \dots & 0, & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0, & 0 & \dots & 0, & 0 & \dots & 1 \end{vmatrix}, \tag{5}$$

wobei $n-1$ Zeilen und Kolonnen beigefügt sind, von denen die $(n-\nu)$ letzten außer in den Diagonalgliedern lauter Nullen haben.

Bildet man jetzt das Produkt $A\Delta_\nu$, so folgt

$$A\Delta_\nu = \begin{vmatrix} A, & 0 \dots 0, & a_{\nu+1}^{(1)} & \dots & a_n^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0, & 0 \dots A, & a_{\nu+1}^{(\nu)} & \dots & a_n^{(\nu)} \\ 0, & 0 \dots 0, & a_{\nu+1}^{(\nu+1)} & \dots & a_n^{(\nu+1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0, & 0 \dots 0, & a_{\nu+1}^{(n)} & \dots & a_n^{(n)} \end{vmatrix}, \tag{6}$$

und dies ist nach § 3

$$= A^{\nu} \begin{vmatrix} a_{\nu+1}^{(\nu+1)} \dots a_n^{(\nu+1)} \\ \dots\dots\dots\dots \\ a_{\nu+1}^{(n)} \dots a_n^{(n)} \end{vmatrix}.$$

Dividiert man hier durch A und wendet die Bezeichnung des § 3 an, so folgt

$$\Delta_{\nu} = A^{\nu-1} A_{1,2\dots\nu}^{1,2\dots\nu}. \tag{7}$$

Für $\nu = 2$ ergibt sich das spezielle Resultat

$$A_1^{(1)} A_2^{(2)} - A_1^{(2)} A_2^{(1)} = A A_{1,2}^{1,2}. \tag{8}$$

In der Bezeichnung durch die Differentialquotienten läßt sich diese Formel so darstellen:

$$A \frac{\partial^2 A}{\partial a_1^{(1)} \partial a_2^{(2)}} = \frac{\partial A}{\partial a_1^{(1)}} \frac{\partial A}{\partial a_2^{(2)}} - \frac{\partial A}{\partial a_1^{(2)}} \frac{\partial A}{\partial a_2^{(1)}}. \tag{9}$$

Besonders wichtig ist sie in dem Falle, wo A eine symmetrische Determinante ist, wo also $a_i^{(k)} = a_k^{(i)}$ ist; dann ist auch

$$\frac{\partial A}{\partial a_1^{(2)}} = \frac{\partial A}{\partial a_2^{(1)}},$$

worin bei der Differentiation nicht Rücksicht genommen ist auf die Abhängigkeit $a_i^{(k)} = a_k^{(i)}$; dann wird die Formel (9)

$$A \frac{\partial^2 A}{\partial a_1^{(1)} \partial a_2^{(2)}} = \frac{\partial A}{\partial a_1^{(1)}} \frac{\partial A}{\partial a_2^{(2)}} - \left(\frac{\partial A}{\partial a_1^{(2)}}\right)^2. \tag{10}$$

Eine andere Anwendung des Multiplikationsgesetzes führt zu einem Determinantensatz von Sylvester, den wir hier nach dem Vorgange von Frobenius beweisen wollen.

Es sei

$$A = \Sigma \pm a_1^{(1)} a_2^{(2)} \dots a_n^{(n)}$$

eine Determinante, r irgend eine Zahl $< n$ und

$$A_r = \Sigma \pm a_1^{(1)} a_2^{(2)} \dots a_r^{(r)}$$

von Null verschieden. Mit $A_h^{(k)}$ bezeichnen wir jetzt die Unterdeterminanten von A_r (nicht von A) und bilden das Produkt der beiden Determinanten

$$\begin{vmatrix} a_1^{(1)} \dots a_1^{(r)} & a_1^{(r+1)} \dots a_n^{(n)} \\ \dots\dots\dots\dots\dots\dots\dots \\ a_r^{(1)} \dots a_r^{(r)} & a_r^{(r+1)} \dots a_r^{(n)} \\ a_{r+1}^{(1)} \dots a_{r+1}^{(r)} & a_{r+1}^{(r+1)} \dots a_{r+1}^{(n)} \\ \dots\dots\dots\dots\dots\dots\dots \\ a_n^{(1)} \dots a_n^{(r)} & a_n^{(r+1)} \dots a_n^{(n)} \end{vmatrix} \cdot \begin{vmatrix} A_1^{(1)} \dots A_1^{(r)} & 0 \dots 0 \\ \dots\dots\dots\dots\dots \\ A_r^{(1)} \dots A_r^{(r)} & 0 \dots 0 \\ Y_{r+1}^{(1)} \dots Y_{r+1}^{(r)} & A_r \dots 0 \\ \dots\dots\dots\dots\dots \\ Y_n^{(1)} \dots Y_n^{(r)} & 0 \dots A_r \end{vmatrix}, \tag{11}$$

dessen Wert, was auch die Y sein mögen, den Wert

$$A A_r^{n-1} \tag{12}$$

hat. Bestimmen wir die Y und eine weitere Reihe von Größen $B_\alpha^{(\beta)}$ aus den Gleichungen

$$\begin{aligned}
&a_1^{(1)} Y_\beta^{(1)} + \cdots + a_1^{(r)} Y_\beta^{(r)} + a_1^{(\beta)} A_r = 0 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&a_r^{(1)} Y_\beta^{(1)} + \cdots + a_r^{(r)} Y_\beta^{(r)} + a_r^{(\beta)} A_r = 0 \\
&a_\alpha^{(1)} Y_\beta^{(1)} + \cdots + a_\alpha^{(r)} Y_\beta^{(r)} + a_\alpha^{(\beta)} A_r = B_\alpha^{(\beta)}
\end{aligned}$$

$$\begin{matrix} \alpha = r+1 \ldots n \\ \beta = r+1 \ldots n, \end{matrix}$$

so folgt durch Elimination von $Y_\beta^{(1)}, \ldots Y_\beta^{(r)}$:

$$B_\alpha^{(\beta)} = \begin{vmatrix} a_1^{(1)} \ldots a_1^{(r)}, & a_1^{(\beta)} \\ \cdots\cdots\cdots\cdots & \\ a_r^{(1)} \ldots a_r^{(r)}, & a_r^{(\beta)} \\ a_\alpha^{(1)} \ldots a_\alpha^{(r)}, & a_\alpha^{(\beta)} \end{vmatrix}, \tag{13}$$

und die Multiplikation der beiden Determinanten (11) ergibt:

$$\begin{vmatrix} A_r \ldots\ldots 0, & 0 \ldots\ldots 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ 0 \ldots\ldots A_r, & 0 \ldots\ldots 0 \\ X_{r+1}^{(1)} \ldots X_{r+1}^{(r)}, & B_{r+1}^{(r+1)} \ldots B_{r+1}^{(n)} \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ X_n^{(1)} \ldots X_{n+1}^{(r)}, & B_n^{(r+1)} \ldots B_n^{(n)} \end{vmatrix},$$

worin die X gewisse r-reihige Determinanten sind, auf deren Kenntnis es nicht ankommt. Der Wert dieser Determinante ist aber nach § 3

$$A_r^r \Sigma \pm B_{r+1}^{(r+1)} \ldots B_n^{(n)},$$

und hieraus und aus (12) ergibt sich die Relation von Sylvester:

$$\Sigma \pm B_{r+1}^{(r+1)} \ldots B_n^{(n)} = A_r^{n-r-1} A, \tag{14}$$

wo die $B_\alpha^{(\beta)}$ die Bedeutung (13) haben und die für $r = n - 2$ in die Relation (9) übergeht[1]).

---

[1]) Sylvester, Philosophical Magazine 1851 (coll. math. papers I). Frobenius, Crelles Journal, Bd. 86, S. 53. Sitzungsberichte d. Berliner Akademie 1894, XII, XXIII.

§ 8.

Orthogonale Substitution.

Die Substitution

$$y_\alpha = \sum_{1,n}^{s} a_s^{(\alpha)} x_s \tag{1}$$

heißt orthogonal, wenn sie die Gleichung

$$\sum_{1,n}^{s} y_s^2 = \sum_{1,n}^{s} x_s^2 \tag{2}$$

zu einer identischen macht. Die aus den Koeffizienten gebildete Matrix

$$A = \begin{pmatrix} a_1^{(1)}, a_2^{(1)} \dots a_n^{(1)} \\ a_1^{(2)}, a_2^{(2)} \dots a_n^{(2)} \\ \dots\dots\dots\dots\dots \\ a_1^{(n)}, a_2^{(n)} \dots a_n^{(n)} \end{pmatrix} \tag{3}$$

heißt dann auch eine orthogonale Matrix.

Substituiert man (1) in (2), so ergeben sich die folgenden Relationen:

$$\begin{aligned} \sum^{s} a_\alpha^{(s)} a_\beta^{(s)} &= 0 \qquad \alpha \neq \beta \\ &= 1 \qquad \alpha = \beta \end{aligned} \tag{4}$$

Dafür schreiben wir auch, wenn wir $(\alpha, \beta) = 0$ oder $= 1$ setzen, je nachdem $\alpha \neq \beta$ oder $\alpha = \beta$ ist:

$$\sum^{s} a_\alpha^{(s)} a_\beta^{(s)} = (\alpha, \beta). \tag{5}$$

Mit Hilfe dieser Gleichungen läßt sich die Substitution (1) sehr einfach auflösen, indem man (1) mit $a_t^{(\alpha)}$ multipliziert und nach α summiert:

$$x_t = \sum_{1,n}^{\alpha} a_t^{(\alpha)} y_\alpha \tag{6}$$

und wenn man dies in (2) einsetzt, so ergibt sich eine zweite Form der Relationen (5):

$$\sum_{1,n}^{s} a_s^{(\alpha)} a_s^{(\beta)} = (\alpha, \beta). \tag{7}$$

Wenn wir nach der Multiplikationsregel das Quadrat der Determinante A bilden, so ergibt sich mit Benutzung von (5) oder (7):

$$A^2 = 1,$$

also

$$A = \pm 1, \tag{8}$$

und es gibt also zwei Arten orthogonaler Substitutionen, je nachdem hier das obere oder das untere Zeichen steht.

Setzt man zwei oder mehrere orthogonale Substitutionen zusammen, so entsteht eine neue orthogonale Substitution. Die Zusammensetzung zweier Substitutionen der ersten oder zweier der zweiten Art ergibt eine Substitution der ersten Art ($A = +1$). Die Zusammensetzung zweier verschiedenartiger Substitutionen ergibt eine Substitution der zweiten Art.

Multiplizieren wir die Relation (5) mit der Unterdeterminante $A_\beta^{(t)}$ und summieren in bezug auf β, so ergibt sich nach § 3 (3), (7), wenn wir $A = +1$ nehmen und α für t setzen:

$$A_\beta^{(\alpha)} = a_\beta^{(\alpha)} \qquad \begin{matrix} \alpha = 1, 2 \dots n \\ \beta = 1, 2 \dots n. \end{matrix} \tag{9}$$

Dies läßt sich sehr verallgemeinern, wenn man die Formel (7), (5) § 7 anwendet. Bezeichnet man mit

$$\begin{matrix} \alpha_1, & \alpha_2 \dots \alpha_\nu, & \alpha_{\nu+1} \dots \alpha_n \\ \beta_1, & \beta_2 \dots \beta_\nu, & \beta_{\nu+1} \dots \beta_n \end{matrix}$$

zwei Anordnungen der Ziffern 1, 2, 3 ... n, so ist bei einer orthogonalen Substitution der Determinante $+1$:

$$A^{\beta_1, \beta_2 \dots \beta_\nu}_{\alpha_1, \alpha_2 \dots \alpha_\nu} = A^{\beta_1+1, \beta_\nu+2 \dots \beta_n}_{\alpha_\nu+1, \alpha_\nu+2 \dots \alpha_n} \tag{10}$$

oder in Worten ausgedrückt:

XXIV. Bei einer orthogonalen Matrix mit der Determinante $+1$ ist jede Unterdeterminante beliebiger Ordnung gleich ihrer komplementären Unterdeterminante.

Die ternären orthogonalen Substitutionen sind seit lange in der Geometrie bekannt, wo sie die rechtwinklige Koordinatentransformation darstellen, die quaternären haben in jüngster Zeit für die Elektrodynamik in der sogenannten Relativitätstheorie eine neue Bedeutung gewonnen, wobei besonders die Formel (10) nützlich ist[1]).

§ 9.

Quadratische Formen.

Die linearen Substitutionen dienen dazu, ganze homogene Funktionen umzuformen. Eine solche Funktion wird auch eine

[1]) Vgl. Riemann-Weber, Die partiellen Differentialgleichungen der mathematischen Physik, 5. Aufl., Bd. II, § 156 (1911).

Form genannt, und wir beschäftigen uns hier zunächst mit den quadratischen Formen. Wir bezeichnen die n Variablen, von denen eine solche Form abhängt, mit

$$x_1, x_2, \dots x_n$$

und die Form mit

$$\varphi(x_1, x_2 \dots x_n) = \varphi(x) = \sum_{i,k}^{i,k} a_{i,k} x_i x_k, \tag{1}$$

worin i und k, von einander unabhängig, von 1 bis n laufen und die Koeffizienten der Bedingung

$$a_{ik} = a_{ki} \tag{2}$$

genügen, so daß z. B. für $n = 3$

$$\varphi(x_1, x_2, x_3) = a_{11} x_1^2 + a_{22} x_2^2 + a_{33} x_3^2 + 2 a_{23} x_2 x_3 + 2 a_{31} x_3 x_1 + 2 a_{12} x_1 x_2$$

ist. Wir setzen die Derivierten

$$\tfrac{1}{2} \varphi'(x_k) = \sum_{1,n}^{i} a_{ik} x_i \tag{3}$$

und nennen

$$R = \begin{vmatrix} a_{11}, a_{12} \dots a_{1n} \\ a_{22}, a_{22} \dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1}, a_{n2} \dots a_{nn} \end{vmatrix}, \tag{4}$$

was wegen (2) eine symmetrische Determinante ist, die Determinante der Form φ. Sie kann auch aufgefaßt werden als die Determinante der zweiten partiellen Ableitungen von φ (geteilt durch 2^n) und heißt als solche die Hessesche Determinante von φ.

Wir machen nun in φ für die Variable x eine lineare Substitution

$$x_k = \sum_{1,n}^{i} \alpha_k^{(i)} x_i', \tag{5}$$

deren Substitutionsdeterminante

$$r = \begin{vmatrix} \alpha_1^{(1)}, \alpha_1^{(2)} \dots \alpha_1^{(n)} \\ \alpha_2^{(1)}, \alpha_2^{(2)} \dots \alpha_2^{(n)} \\ \dots\dots\dots\dots \\ \alpha_n^{(1)}, \alpha_n^{(2)} \dots \alpha_n^{(n)} \end{vmatrix} \tag{6}$$

von Null verschieden ist.

Dadurch geht die Form $\varphi(x)$ in eine ganz ebenso gebaute Form $\psi(x')$ über, die wir so bezeichnen:

$$\psi(x_1', x_2' \dots x_n') = \psi(x') = \sum_{i,k}^{i,k} a_{ik}' x_i' x_k', \tag{7}$$

und durch (5) wird

$$\varphi(x) = \psi(x') \tag{8}$$

zu einer identischen Gleichung.

Zwischen den Determinanten R und R' der Formen φ und ψ besteht nun eine fundamentale Relation, die sich auf folgende Weise ergibt. Differenziiert man die identische Gleichung $\varphi = \psi$, so folgt aus (5):

$$\frac{\partial \psi}{\partial x_i'} = \sum^{k} \frac{\partial \varphi}{\partial x_k} \alpha_k^{(i)}$$

$$\frac{\partial^2 \psi}{\partial x_i' \partial x_h'} = \sum^{k} \frac{\partial^2 \varphi}{\partial x_k \partial x_h'} \alpha_k^{(i)},$$

und folglich nach dem Multiplikationsgesetz:

$$\sum \pm \frac{\partial^2 \psi}{\partial x_1' \partial x_1'} \frac{\partial^2 \psi}{\partial x_2' \partial x_2'} \cdots \frac{\partial^2 \psi}{\partial x_n' \partial x_n'}$$

$$= r \sum \pm \frac{\partial^2 \varphi}{\partial x_1 \partial x_1'} \frac{\partial^2 \varphi}{\partial x_2 \partial x_2'} \cdots \frac{\partial^2 \varphi}{\partial x_n \partial x_n'}.$$

Substituiert man hier wieder:

$$\frac{\partial^2 \varphi}{\partial x_h \partial x_h'} = \sum^{k} \frac{\partial^2 \varphi}{\partial x_h \partial x_k} \alpha_k^{(i)}$$

und wendet nochmals die Multiplikationsregel an, so erhält man nun, da

$$\frac{\partial^2 \varphi}{\partial x_i \partial x_k} = a_{ik}, \quad \frac{\partial^2 \psi}{\partial x_i' \partial x_k'} = a_{ik}'$$

ist:

$$R' = r^2 R. \tag{9}$$

Wenn also R verschwindet, so verschwindet auch R', und wir beweisen den folgenden Satz:

XXV. Das Verschwinden der Determinante R hat die Bedeutung, daß die Funktion $\varphi(x)$ sich ausdrücken läßt durch weniger als n linear unabhängige Funktionen der x.

Bildet man nämlich aus (5) und (7) die Koeffizienten a_{st}', so erhält man:

$$a_{st}' = \sum^{i,k} a_{i,k} \alpha_i^{(s)} \alpha_k^{(t)},$$

und wenn man also die Substitutionskoeffizienten $\alpha_1^{(1)}, \alpha_2^{(1)} \ldots \alpha_n^{(1)}$ aus den Bedingungen bestimmt:

(10) $$\sum_{k}^{k} a_{ik}\alpha_k^{(1)} = 0, \qquad i = 1, 2 \ldots n,$$

so erhält man

$$a'_{11} = 0, \quad a'_{21} = 0, \quad \ldots a'_{n,1} = 0$$

und dies heißt, daß die Funktion ψ von x'_1 unabhängig ist.

Die Gleichungen (10) können aber nach § 4 dann und nur dann durch ein System der Unbekannten $\alpha_k^{(1)}$, die nicht alle verschwinden, befriedigt werden, wenn die Determinante R verschwindet.

Ist z. B. $\alpha_1^{(1)}$ von Null verschieden, so kann man die Substitutionsmatrix

$$\begin{array}{cccc} \alpha_1^{(1)}, & \alpha_2^{(1)}, & \ldots & \alpha_n^{(1)} \\ 0, & 1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0, & 0 & \ldots & 1 \end{array}$$

nehmen, und dann sind die x' linear unabhängige Funktionen von den x.

§ 10.

Trägheitsgesetz der quadratischen Formen.

XXVI. Eine quadratische Form von n Variablen läßt sich darstellen als eine Summe von höchstens n Quadraten unabhängiger linearer Funktionen der Variablen.

Um ihn zu beweisen, nehmen wir zunächst an, es sei einer der Koeffizienten a_{ii}, etwa a_{11} von Null verschieden. Dann können wir setzen:

(1) $$a_{11}\varphi(x) = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^2 + \varphi_1(x),$$

worin $\varphi_1(x)$ nur von den Variablen $x_2, x_3 \ldots x_n$, nicht mehr von x_1 abhängig ist, und damit ist der Fall n auf den Fall $n - 1$ zurückgeführt.

Wenn aber die Koeffizienten $a_{11}, a_{22} \ldots a_{nn}$ alle verschwinden, so muß doch einer der anderen Koeffizienten, etwa a_{12}, von Null verschieden sein. Dann setze man

$$x_1 = x'_1 + x'_2, \qquad x_2 = x'_1 - x'_2$$

und erhält:

(2) $$\varphi(x) = 2a_{12}(x'^2_1 - x'^2_2) + \psi(x'_1, x'_2, x_3 \ldots x_n),$$

worin ψ die Glieder mit x'^2_1, x'^2_2 nicht enthält.

Damit ist dieser Fall auf den vorigen zurückgeführt und der Satz XXVI bewiesen.

Ist die Funktion $\varphi(x)$ eine reelle Funktion, so können auch die Quadrate der linearen Bestandteile, in die man $\varphi(x)$ zerlegt, nach (1) reell angenommen werden, sind aber positiv oder negativ, und diese Zerlegung kann auf unendlich viele verschiedene Arten geschehen. Darüber gilt nun der folgende Satz, der von Sylvester den Namen des Gesetzes der Trägheit der quadratischen Form erhalten hat.

XXVII. Wie man auch eine reelle quadratische Funktion $\varphi(x)$ in die Summe von positiven und negativen Quadraten linearer Funktionen zerlegen mag, die Anzahl der positiven und negativen dieser Quadrate und also auch ihre Gesamtzahl ist immer dieselbe, vorausgesetzt, daß zwischen diesen linearen Funktionen keine lineare Abhängigkeit besteht.

Zum Beweise dieses Satzes seien

$$(3)\quad \begin{aligned}\varphi(x) &= Y_1^2 + Y_2^2 + \cdots Y_\nu^2 - Y_1'^2 - Y_2'^2 - \cdots - Y_{\nu'}'^2\\ &= Z_1^2 + Z_2^2 + \cdots Z_\mu^2 - Z_1'^2 - Z_2'^2 - \cdots - Z_{\mu'}'^2\end{aligned}$$

zwei Zerlegungen von $\varphi(x)$ in Quadrate. Es seien also $Y_1, Y_2 \ldots Y_\nu, Y_1', Y_2' \ldots Y_{\nu'}'$ homogene lineare Funktionen von x, zwischen denen keine lineare Relation mit konstanten Koeffizienten besteht und also $\nu + \nu' \gtreqless n$; dieselben Voraussetzungen werden über die Funktionen $Z_1, Z_2 \ldots Z_\mu, Z_1', Z_2' \ldots Z_{\mu'}'$ gemacht, so daß auch $\mu + \mu' \gtreqless n$ ist.

Angenommen, es sei

$$\nu < \mu,$$

dann können wir die Variablen x den linearen Gleichungen

$$(4)\quad \begin{aligned} Y_1 &= 0, \quad Y_2 = 0 \ldots Y_\nu = 0\\ Z_1' &= 0, \quad Z_2' = 0 \ldots Z_{\mu'}' = 0\end{aligned}$$

unterwerfen, deren Anzahl $\nu + \mu'$ kleiner als n ist.

Wenn nun die sämtlichen Funktionen $Z_1, Z_2 \ldots Z_\mu$ linear abhängig wären von den Funktionen $Y_1, Y_2 \ldots Y_\nu, Z_1', Z_2' \ldots Z_{\mu'}'$, so könnte man aus diesen μ Gleichungen durch Elimination der ν Variablen $Y_1, Y_2 \ldots Y_\nu$ eine lineare Gleichung zwischen den $Z_1, Z_2 \ldots Z_\mu, Z_1', Z_2' \ldots Z_{\mu'}'$ herleiten, die nach Voraussetzung

nicht bestehen soll. Es ist also das Verschwinden sämtlicher $Z_1, Z_2 \ldots Z_u$ nicht eine notwendige Folge der Gleichungen (4), und wir können zu den Gleichungen (4) noch eine nicht homogene hinzufügen, etwa

$$Z_1 = 1;$$

dann haben wir für die n Unbekannten x ein System von n oder weniger Gleichungen, die voneinander unabhängig sind, sich also nicht widersprechen.

Dann ergibt aber die zweite Darstellung (3) für $\varphi(x)$ einen positiven Wert, während die erste einen negativen oder verschwindenden Wert gibt, worin ein Widerspruch liegt. Es folgt also

$$\nu \gtreqless \mu,$$

und da man ebenso schließt

$$\mu \gtreqless \nu,$$

so bleibt nur $\nu = \mu$ übrig. In gleicher Weise kann man beweisen, daß $\nu' = \mu'$ ist, wodurch unser Satz vollständig bewiesen ist.

Bezeichnen wir mit π die Anzahl der positiven und mit ν die Anzahl der negativen Quadrate, so ist $\pi + \nu$ höchstens gleich n, kann aber auch kleiner als n sein.

Wenn wir den Unterschied $n - \pi - \nu$ mit ϱ bezeichnen, so kann, wenn wir die allgemeine quadratische Form von n Variablen als Summe von n Quadraten darstellen, ϱ als die Anzahl der Quadrate mit verschwindenden Koeffizienten bezeichnet werden.

Die drei Zahlen π, ν, ϱ, von denen keine negativ sein kann, und deren Summe gleich der Anzahl der Variablen ist:

$$\pi + \nu + \varrho = n, \tag{5}$$

sind also für eine bestimmte quadratische Form unveränderlich. Wir werden sie die charakteristischen Zahlen der quadratischen Formen nennen[1]).

Wir haben nach Mitteln zu suchen, um aus den Koeffizienten der quadratischen Form die Zahlen π, ν, ϱ zu ermitteln.

---

[1]) Nach Frobenius heißt $\pi + \nu$ der Rang, $\pi - \nu$ die Signatur der quadratischen Form („Über das Trägheitsgesetz der quadratischen Formen"; Sitzungsberichte der Berliner Akademie 1894, auch in Crelles Journal, Bd. 114); ϱ wird auch der Defekt der quadratischen Form genannt.

Es sei also, wie in § 9

$$(6) \qquad \varphi(x_1, x_2 \dots x_n) = \sum_{1,n}^{i,k} a_{i,k}\, x_i\, x_k, \qquad a_{i,k} = a_{k,i}$$

eine quadratische Form von n Veränderlichen mit reellen Koeffizienten $a_{i,k}$.

Es sei wieder R die Determinante der quadratischen Form $\varphi(x)$, deren Koeffizienten jetzt reell vorausgesetzt sind. Unter den Haupt-Unterdeterminanten verstehen wir solche, die man erhält, wenn man Zeilen und Kolonnen (in beliebiger Anzahl) ausstreicht, die sich in Diagonalgliedern schneiden, die also, wenn $\alpha, \beta, \gamma \dots$ irgend welche unter den Indices $1, 2 \dots n$ bedeuten, durch

$$(7) \qquad \Sigma \pm a_{\alpha,\alpha}\, a_{\beta,\beta}\, a_{\gamma,\gamma} \dots$$

zu bezeichnen sind.

Die Indices $1, 2, 3 \dots n$ lassen sich auf $n! = 1.2.3 \dots n$ verschiedene Arten anordnen; wenn wir mit irgend einer dieser Anordnungen die Determinante bilden:

$$\begin{vmatrix} a_{1,1}, & a_{1,2}, & a_{1,3} & \dots & a_{1,n} \\ a_{2,1}, & a_{2,2}, & a_{2,3} & \dots & a_{2,n} \\ a_{3,1}, & a_{3,2}, & a_{3,3} & \dots & a_{3,n} \\ \dots\dots & \dots\dots & \dots\dots & \dots & \dots\dots \\ a_{n,1}, & a_{n,2}, & a_{n,3} & \dots & a_{n,n} \end{vmatrix},$$

so läßt sich eine Reihe von Haupt-Unterdeterminanten daraus ableiten, wie es durch die punktierten Striche angedeutet ist, d. h. so, daß man jede vorhergehende aus der nachfolgenden erhält, indem man die letzte Zeile und Kolonne wegläßt; es ist also

$$R_0 = 1, \quad R_1 = a_{1,1}, \quad R_2 = a_{1,1}\, a_{2,2} - a_{1,2}^2,$$

$$R_3 = \begin{vmatrix} a_{1,1}, & a_{1,2}, & a_{1,3} \\ a_{2,1}, & a_{2,2}, & a_{2,3} \\ a_{3,1}, & a_{3,2}, & a_{3,3} \end{vmatrix}, \dots\dots R_n = R.$$

Ein solches System soll, wenn es in absteigender Reihe

$$R_n, \; R_{n-1} \dots R_1, \; R_0$$

geordnet ist, eine Kette von Haupt-Unterdeterminanten heißen. Solcher Ketten lassen sich $n!$ verschiedene bilden, in denen allen das erste Glied R, das letzte Glied 1 ist.

Wir betrachten zunächst den Fall, daß die Determinante R verschwindet.

§ 11.

Verschwindende Determinante.

Das Verschwinden der Determinante

$$R = \Sigma \pm a_{1,1}\, a_{2,2} \dots a_{n,n}$$

zeigt an, daß φ durch lineare Transformation in eine Funktion von weniger als n Variablen transformiert werden kann, daß also ϱ größer als Null ist.

Wenn sich unter den Haupt-Unterdeterminanten von R eine k-reihige findet, die von Null verschieden ist, etwa

$$R_k = \Sigma \pm a_{1,1}\, a_{2,2} \dots a_{k,k},$$

so ist, wenn wir

$$x_{k+1} = 0, \quad x_{k+2} = 0, \quad \dots x_n = 0$$

setzen,

$$\varphi(x_1, x_2 \dots x_k, 0, 0 \dots 0)$$

eine quadratische Form von k Variablen $x_1, x_2 \dots x_k$, deren Determinante R_k von Null verschieden ist, die sich sonach nicht durch weniger als k Variable ausdrücken läßt. Um so weniger kann also bei unbeschränkt veränderlichen $x_1, x_2 \dots x_n$ die Funktion φ von weniger als k Variablen abhängen, und es folgt, wenn, wie im vorigen Paragraphen, $n - \varrho$ die kleinstmögliche Zahl von linearen Funktionen der x ist, durch die sich $\varphi(x)$ darstellen läßt:

$$\varrho \gtreqless n - k. \tag{1}$$

Wir nehmen jetzt an, die Form $\varphi(x_1, x_2 \dots x_n)$ lasse sich durch k voneinander unabhängige lineare Formen von x ausdrücken, die wir mit

$$\begin{aligned}
z_1 &= \beta_1^{(1)} x_1 + \beta_2^{(1)} x_2 + \dots + \beta_k^{(1)} x_k + \beta_{k+1}^{(1)} x_{k+1} + \dots + \beta_n^{(1)} x_n \\
z_2 &= \beta_1^{(2)} x_1 + \beta_2^{(2)} x_2 + \dots + \beta_k^{(2)} x_k + \beta_{k+1}^{(2)} x_{k+1} + \dots + \beta_n^{(2)} x_n \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
z_k &= \beta_1^{(k)} x_1 + \beta_2^{(k)} x_2 + \dots + \beta_k^{(k)} x_k + \beta_{k+1}^{(k)} x_{k+1} + \dots + \beta_n^{(k)} x_n
\end{aligned} \tag{2}$$

bezeichnen wollen.

Da $z_1, z_2 \dots z_k$ voneinander linear unabhängig sein sollen, so muß unter den aus der Matrix

$$\begin{pmatrix}
\beta_1^{(1)}, & \beta_2^{(1)} & \dots & \beta_n^{(1)} \\
\beta_1^{(2)}, & \beta_2^{(2)} & \dots & \beta_n^{(2)} \\
\dots\dots & \dots\dots & \dots & \dots\dots \\
\beta_1^{(k)}, & \beta_2^{(k)} & \dots & \beta_n^{(k)}
\end{pmatrix}$$

zu bildenden k-reihigen Determinanten wenigstens eine von Null verschieden sein. Denn wären sie alle gleich Null, so ließen sich die Unbekannten $h_1, h_2 \dots h_k$, die nicht alle verschwinden sollen, aus den Gleichungen

$$h_1 \beta_s^{(1)} + h_2 \beta_s^{(2)} + \cdots + h_k \beta_s^{(k)} = 0, \quad s = 1, 2, 3 \dots n$$

bestimmen, und die $z_1, z_2 \dots z_k$ würden einer Gleichung

$$h_1 z_1 + h_2 z_2 + \cdots + h_k z_k = 0$$

genügen, also nicht unabhängig sein.

Wir können, ohne die Allgemeinheit zu beschränken, annehmen, daß die Determinante

$$\Sigma \pm \beta_1^{(1)} \beta_2^{(2)} \dots \beta_k^{(k)}$$

von Null verschieden sei, und dann können wir die Gleichungen (2) in bezug auf die Variablen $x_1, x_2 \dots x_k$ auflösen.

Diesen Auflösungen geben wir die Form:

$$\begin{aligned} x_1 &= y_1 + \alpha_{k+1}^{(1)} x_{k+1} + \cdots + \alpha_n^{(1)} x_n \\ x_2 &= y_2 + \alpha_{k+1}^{(2)} x_{k+1} + \cdots + \alpha_n^{(2)} x_n \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_k &= y_k + \alpha_{k+1}^{(k)} x_{k+1} + \cdots + \alpha_n^{(k)} x_n, \end{aligned} \tag{3}$$

worin die y_s lineare homogene Verbindungen der z_s sind, die aus den Gleichungen

$$\begin{aligned} \beta_1^{(1)} y_1 + \beta_2^{(1)} y_2 + \cdots + \beta_k^{(1)} y_k &= z_1, \\ \beta_1^{(2)} y_1 + \beta_2^{(2)} y_2 + \cdots + \beta_k^{(2)} y_k &= z_2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_1^{(k)} y_1 + \beta_2^{(k)} y_2 + \cdots + \beta_k^{(k)} y_2 &= z_k \end{aligned}$$

bestimmt werden, während die α neue Koeffizienten sind, die in einer leicht zu übersehenden Weise aus den β abgeleitet werden, auf deren Bildung es uns hier nicht weiter ankommt.

Nun läßt sich nach der Voraussetzung die Funktion $\varphi(x_1, x_2 \dots x_n)$ durch $z_1, z_2 \dots z_k$, also auch durch $y_1, y_2 \dots y_k$ allein ausdrücken. Bezeichnen wir diesen Ausdruck mit $\Phi(y_1, y_2 \dots y_k)$, so haben wir [durch Vermittelung von (3)] die Identität

$$\varphi(x_1, x_2 \dots x_k, \quad x_{k+1} \dots x_n) = \Phi(y_1, y_2 \dots y_k), \tag{4}$$

und wenn wir darin $x_{k+1}, x_{k+2} \dots x_n = 0$ setzen:

$$(5) \qquad \varphi(x_1, x_2 \dots x_k, 0, 0 \dots 0) = \Phi(x_1, x_2 \dots x_k),$$

wodurch, da es auf die Bezeichnung der Variablen nicht ankommt, Φ vollständig bestimmt ist: wir können daher setzen:

$$(6) \qquad \begin{aligned} \Phi(y_1, y_2 \dots y_k) &= \varphi(y_1, y_2 \dots y_k, 0, 0 \dots 0) \\ &= \varphi(x_1, x_2 \dots x_n); \end{aligned}$$

Φ entsteht also aus φ dadurch, daß man die $n - k$ letzten Variablen (bei der hier gewählten Anordnung) gleich Null setzt:

$$(7) \qquad \Phi(x_1, \dots x_k) = \sum_{1,k}^{r,s} a_{r,s} x_r x_s.$$

Die Determinante von Φ ist eine der Haupt-Unterdeterminanten von R, und zwar erhält man sie, indem man in R die $n - k$ letzten Zeilen und Kolonnen wegstreicht.

Nehmen wir an, daß φ sich nicht durch noch weniger als k Variable ausdrücken läßt, daß also $k = n - \varrho$ sei, so kann die Determinante dieser Funktion Φ nach XXV, § 9 nicht verschwinden.

Hieraus und aus dem unter (1) Bewiesenen ergibt sich nun der Satz:

XXVIII. Wenn alle Haupt-Unterdeterminanten von R von mehr als k Reihen verschwinden, während unter den Haupt-Unterdeterminanten von k Reihen wenigstens eine von Null verschieden ist, so ist

$$\varrho = n - k.$$

Denn aus (7) folgt, daß, wenn $\varrho = n - k$ ist, wenigstens eine k-reihige Haupt-Unterdeterminante von Null verschieden sein muß, und aus (1), daß, wenn auch noch eine Haupt-Unterdeterminante von mehr als k Reihen von Null verschieden ist,

$$\varrho < n - k$$

ist. Wenn also $\varrho = n - k$ ist, so kann keine Unterdeterminante von mehr als k Reihen von Null verschieden sein.

Sind π, ν, ϱ die charakteristischen Zahlen der Form φ, so sind π, ν die Anzahlen der positiven und negativen Quadrate, in die sich die durch (7) bestimmte Form Φ zerlegen läßt, und die Bestimmung der Zahlen π, ν braucht also nur noch für letztere Funktion, deren Determinante von Null verschieden ist, durchgeführt zu werden.

§ 12.

Nicht verschwindende Determinante.

Bei der Untersuchung der Formen $\varphi(x_1, x_2 \dots x_n)$ mit nicht verschwindender Determinante machen wir von dem im § 7 bewiesenen Determinantensatz (8) Gebrauch:

Ist A eine Determinante von n Reihen, und sind $A_i^{(k)}$ ihre ersten, $A_{i,i'}^{k,k'}$ ihre zweiten Unterdeterminanten, so ist

$$A_1^{(1)} A_i^{(i)} - A_1^{(i)} A_i^{(1)} = A A_{1,i}^{1,i}. \tag{1}$$

Ist A eine symmetrische Determinante, so ist $A_1^{(i)} = A_i^{(1)}$, und wir können die Formel (1) so schreiben:

$$A_1^{(1)} A_i^{(i)} - A_1^{(i)2} = A A_{1,i}^{1,i}. \tag{2}$$

Darin kann i jeden der Indices 2, 3 ... n bedeuten. Wenn $A_1^{(1)} = 0$ und $A_{1,i}^{1,i} = 0$ ist, so folgt hieraus, daß auch $A_1^{(i)} = 0$ sein muß, und daraus schließen wir, daß nicht zugleich

$$A_1^{(1)},\ A_{1,2}^{1,2},\ A_{1,3}^{1,3} \dots A_{1,n}^{1,n}$$

verschwinden können, da sonst auch

$$A_1^{(2)},\ A_1^{(3)} \dots A_1^{(n)},$$

also auch A verschwinden würde.

Wenden wir dies auf die jetzt von Null verschieden angenommene Determinante $R = R_n$ unserer Funktion φ an, so folgt, daß zwar die erste Haupt-Unterdeterminante R_{n-1}, dann aber nicht alle zweiten Haupt-Unterdeterminanten R_{n-2} verschwinden können. Dieselbe Schlußweise läßt sich anwenden, wenn wir R durch ein nicht verschwindendes R_{n-1}, R_{n-2} ... ersetzen, und wir gelangen also zu folgendem wichtigen Satz:

XXIX. Man kann, wenn R von Null verschieden ist, die Indices 1, 2 ... n so anordnen, daß in der Kette der Haupt-Unterdeterminanten

$$R_n,\ R_{n-1} \dots R_1,\ R_0 \tag{3}$$

nicht zwei aufeinander folgende Glieder verschwinden.

Die Determinantenrelation (2) ergibt, wenn man

$$A = R_{k+1}, \qquad k = 1, 2, \dots, n-1$$

setzt, eine Gleichung zwischen drei aufeinander folgenden Gliedern der Kette (3), die wir so schreiben können:

$$R_k S_k - T_k^2 = R_{k-1} R_{k+1}, \tag{4}$$

worin S_k und T_k gewisse Unterdeterminanten von R, also ganze rationale Funktionen der Koeffizienten $a_{i,k}$ sind.

Näher bezeichnet, sind R_k, S_k, T_k erste Unterdeterminanten von R_{k+1}, und zwar ist

$$R_k = \frac{\partial R_{k+1}}{\partial a_{k+1,k+1}}, \qquad S_k = \frac{\partial R_{k+1}}{\partial a_{k,k}}, \qquad T_k = \frac{\partial R_{k+1}}{\partial a_{k,k+1}}.$$

Eine dem Satz XXIX entsprechende Anordnung der Indices wollen wir für die Folge als gewählt voraussetzen. Dann ist, wenn R_k verschwindet, R_{k-1} und R_{k+1} von Null verschieden, und (4) zeigt, daß sie entgegengesetzte Vorzeichen haben, also:

XXX. Wenn ein inneres Glied einer Kette von Haupt-Unterdeterminanten verschwindet, so haben die beiden angrenzenden Glieder entgegengesetzte Vorzeichen.

Um die Anzahl der positiven und negativen Quadrate einer Form mit nicht verschwindender Determinante zu bestimmen, nehmen wir zunächst an, daß in der Kette der Haupt-Unterdeterminanten

$$R_n,\ R_{n-1},\ R_{n-2} \ldots R_1,\ R_0 = 1$$

kein Glied verschwinde. Wir können dann den Koeffizienten λ so bestimmen, daß die Determinante der Form

$$\psi = \varphi(x_1, x_2, x_3 \ldots x_n) - \lambda x_n^2 \tag{5}$$

verschwindet. Die Determinante ist nämlich

$$\begin{vmatrix} a_{1,1}, & a_{1,2} \ldots & a_{1,n} \\ a_{2,1}, & a_{2,2} \ldots & a_{2,n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{n,1}, & a_{n,2} \ldots & a_{n,n} - \lambda \end{vmatrix} = R_n - \lambda R_{n-1},$$

und sie verschwindet, wenn

$$\lambda = \frac{R_n}{R_{n-1}}$$

gesetzt wird.

Die Funktion ψ läßt sich dann durch $n - 1$ Variable y ausdrücken, und man erhält nach der Formel (6), § 11

$$\psi(x_1, x_2 \ldots x_n) = \psi(y_1, y_2 \ldots y_{n-1}, 0),$$

oder nach (5)

$$\varphi(x_1, x_2 \ldots x_n) = \frac{R_n}{R_{n-1}} x_n^2 + \varphi(y_1, y_2 \ldots y_{n-1}, 0), \tag{6}$$

und die Untersuchung der Funktion φ von n Variablen ist dadurch auf die Untersuchung von $\varphi_1(y) = \varphi(y_1, y_2 \ldots y_{n-1}, 0)$ von

$n - 1$ Variablen zurückgeführt, deren Determinante gleich R_{n-1}, also von Null verschieden ist. Je nachdem $R_n : R_{n-1}$ positiv oder negativ ist, wird die Funktion $\varphi(x)$ ein positives oder ein negatives Quadrat mehr haben als $\varphi_1(y)$.

Durch Anwendung des gleichen Verfahrens auf $\varphi_1(y)$ und die folgenden Funktionen ergibt sich der Satz:

Die Anzahl der positiven und negativen Glieder der Reihe

$$\frac{R_n}{R_{n-1}}, \quad \frac{R_{n-1}}{R_{n-2}} \quad \cdots \quad \frac{R_1}{R_0}$$

stimmt überein mit der Anzahl der positiven und negativen Quadrate, in die sich die Funktion $\varphi(x)$ zerlegen läßt.

Wir wollen diesem Satze noch einen etwas anderen Ausdruck geben, schicken aber folgende Erklärung voraus:

Wenn eine Reihe von Null verschiedener reeller Zahlen in bestimmter Anordnung vorliegt, so können die Vorzeichen dieser Größen in mannigfaltiger Weise wechseln; folgen zwei Größen von gleichem Zeichen aufeinander, so zählt man eine Zeichenfolge (Permanenz), folgt aber auf eine Größe eine andere von entgegengesetztem Zeichen, so haben wir einen Zeichenwechsel (Variation) zu zählen.

Betrachten wir nun von diesem Gesichtspunkte die Reihe der Größen

$$R_n, \; R_{n-1}, \; R_{n-2} \; \ldots \; R_1, \; R_0,$$

so findet beim Übergange von R_k zu R_{k-1} eine Zeichenfolge oder ein Zeichenwechsel statt, je nachdem der Quotient $R_k : R_{k-1}$ positiv oder negativ ist.

Wir können also auch den folgenden Satz aussprechen:

XXXI. Ist π die Anzahl der positiven, ν die der negativen Quadrate von φ, so ist π gleich der Anzahl der Zeichenfolgen, ν gleich der Anzahl der Zeichenwechsel in der Kette

$$R_n, \; R_{n-1}, \; R_{n-2} \; \ldots \; R_1, \; R_0. \tag{7}$$

Diese Fassung des Satzes hat den Vorzug, daß sie sich auf den Fall übertragen läßt, daß in der Reihe (7) einzelne innere Glieder verschwinden, wenn nur nicht zwei aufeinander folgende Glieder Null sind. Wenn nämlich $R_k = 0$ und R_{k-1}, R_{k+1} von Null verschieden sind, so haben R_{k-1} und R_{k+1} verschiedene Vorzeichen. In der Reihe

$$R_{k+1}, \quad R_k, \quad R_{k-1}$$

findet also ein Zeichenwechsel und eine Zeichenfolge statt, gleichviel ob wir das verschwindende R_k durch eine positive oder eine negative Größe ersetzen.

Wenn wir nun die Kette (7)

$$R_n,\ R_{n-1},\ \dots\ R_1,\ R_0,$$

in der einzelne Glieder verschwinden, durch eine andere ersetzen,

(8) $$R'_n,\ R'_{n-1},\ \dots\ R'_1,\ R'_0,$$

in der kein Glied verschwindet, und in der den nicht verschwindenden Gliedern der Reihe (7) Glieder von demselben Vorzeichen entsprechen, so haben die Reihen (7) und (8) gleich viele Zeichenwechsel, welches Zeichen auch die den verschwindenden R entsprechenden R' haben mögen.

Es sei nun $\psi(x)$ eine zweite beliebige quadratische Form der Variablen x, mit der wir die Form

(9) $$\varphi' = \varphi + \varepsilon\psi$$

bilden, worin ε ein noch unbestimmter Koeffizient ist. Wir werden nun sogleich zeigen, daß wir ε so wählen können, daß φ' und φ dieselbe Zahl von positiven und negativen Quadraten haben, daß aber in der Kette der Haupt-Unterdeterminanten R'_k der Form φ' keine verschwindenden Glieder vorkommen, und daß endlich einem nicht verschwindenden R_k ein R'_k von demselben Vorzeichen entspricht. Dann können die Zahlen π, ν für φ und für φ' sowohl aus der Reihe (7), als auch aus der Reihe (8) ermittelt werden, und die Anzahl der Zeichenwechsel, die in beiden gleich ist, gibt die Anzahl ν der negativen Quadrate.

Um nun den Nachweis zu führen, daß der Koeffizient ε in der angegebenen Weise bestimmt werden kann, nehmen wir an, es sei φ irgendwie in eine Summe von Quadraten verwandelt

(10) $$\varphi = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2,$$

und ψ, in den Variablen $y_1,\ y_2\ \dots\ y_n$ dargestellt, habe den Ausdruck

$$\psi = \Sigma\, \beta_{i,k}\, y_i\, y_k.$$

Die Zahlen π, ν für die Funktion φ' werden dann aus der Kette der Haupt-Unterdeterminanten von

(11)
$$\begin{vmatrix} \lambda_1 + \varepsilon\beta_{1,1}, & \varepsilon\beta_{1,2}, \dots & \varepsilon\beta_{1,n} \\ \varepsilon\beta_{2,1}, & \lambda_2 + \varepsilon\beta_{2,2}, \dots & \varepsilon\beta_{2,n} \\ \dots\dots & \dots\dots & \dots\dots \\ \varepsilon\beta_{n,1}, & \varepsilon\beta_{n,2}, \dots & \lambda + \varepsilon\beta_{n,n} \end{vmatrix}$$

nach dem Satze XXXI bestimmt.

Nun kann man aber ε so klein annehmen, daß diese Haupt-Unterdeterminanten dem Zeichen nach übereinstimmen mit den Produkten

$$\lambda_1 \lambda_2 \lambda_3 \dots \lambda_n, \quad \lambda_1 \lambda_2 \lambda_3 \dots \lambda_{n-1}, \quad \dots \lambda_1 \lambda_2, \quad \lambda_1, \quad 1,$$

und dann ist die Anzahl der positiven und negativen Quadrate von φ' gleich der Anzahl der positiven und negativen unter den Koeffizienten λ, von denen keiner verschwindet, d. h. die Zahlen π und ν sind für φ und φ' dieselben.

Sind nun die Koeffizienten von ψ, in den ursprünglichen Variablen ausgedrückt, $b_{i,k}$, also

$$\psi = \Sigma b_{i,k} x_i x_k,$$

so ist eine der Haupt-Unterdeterminanten von φ'

$$R_k' = \begin{vmatrix} a_{1,1} + \varepsilon b_{1,1} \dots a_{1,k} + \varepsilon b_{1,k} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k,1} + \varepsilon b_{k,1} \dots a_{k,k} + \varepsilon b_{k,k} \end{vmatrix},$$

und ist also eine ganze rationale Funktion k^{ten} Grades von ε,

$$R_k' = R_k + \varepsilon M_1 + \varepsilon^2 M_2 + \dots + \varepsilon^k M_k,$$

worin die M_1, $M_2 \dots M_k$ rational von den $a_{i,k}$, $b_{i,k}$ abhängen. Insbesondere ist

$$M_k = \Sigma \pm b_{1,1} b_{2,2} \dots b_{k,k},$$

und man kann die $b_{i,k}$ immer so annehmen, daß M_k von Null verschieden ist. Man kann dann also ε so annehmen, daß, wenn R_k von Null verschieden ist, R_k' dasselbe Zeichen hat wie R_k, und wenn R_k verschwindet, R_k' nicht verschwindet.

Wir können das hierdurch Bewiesene mit den Ergebnissen des § 11 in eine allgemeine Regel zur Bestimmung der Anzahl der negativen Quadrate, auch für den Fall verschwindender Determinante, zusammenfassen:

XXXII. Um die charakteristischen Zahlen π, ν, ϱ der Funktion $\varphi(x)$ zu bestimmen, ordne man die Variablen x_1, $x_2 \dots x_n$ so an, daß in der Kette der Haupt-Unterdeterminanten

(12) $$R_n, R_{n-1} \dots R_1, R_0$$

eine möglichst kleine Anzahl von Anfangsgliedern verschwindet, und daß von den folgenden Gliedern nicht zwei nebeneinander stehende

verschwinden; ϱ ist dann die Anzahl der verschwindenden Anfangsglieder, ν die Anzahl der Zeichenwechsel und $\pi = n - \nu - \varrho$ [1]).

Zu bemerken ist noch, daß man die Anzahl der Zeichenwechsel in der Reihe (12) auch von rechts nach links abzählen kann, d. h. daß man dieselbe Anzahl von Zeichenwechseln findet, wenn man die Reihe in umgekehrter Ordnung schreibt.

Bezeichnen wir mit π, ν, ϱ die charakteristischen Zahlen der Form φ und bilden die Form

$$\varphi' = \varphi + \varepsilon\psi,$$

so können wir ψ so bestimmen, daß die Determinante von φ' nicht identisch für jedes ε verschwindet. Dann sind π', ν', 0 die charakteristischen Zahlen von φ' und

$$\pi' + \nu' = \pi + \nu + \varrho. \tag{13}$$

Es ist dann in (10)

$$\lambda_n = 0, \quad \lambda_{n-1} = 0, \quad \ldots \quad \lambda_{n-\varrho+1} = 0,$$
$$\lambda_{n-\varrho}, \quad \lambda_{n-\varrho-1}, \quad \ldots \quad \lambda_1 \text{ von Null verschieden.}$$

In der Kette der Haupt-Unterdeterminanten von (11) stimmen dann bei hinlänglich kleinem ε die $n - \varrho$ Endglieder im Vorzeichen überein mit

$$\lambda_1\lambda_2 \ldots \lambda_{n-\varrho}, \quad \lambda_1\lambda_2 \ldots \lambda_{n-\varrho-1}, \ldots, \quad \lambda_1\lambda_2, \quad \lambda_1, \quad 1$$

und die Anzahl der in dieser Kette vorkommenden Zeichenfolgen ist daher mindestens gleich der Anzahl der positiven λ und die Anzahl der Zeichenwechsel mindestens gleich der Anzahl der negativen λ. Es ergibt sich hieraus:

$$\pi' \gtreqless \pi, \qquad \nu' \gtreqless \nu. \tag{14}$$

Nennen wir eine Form φ, deren charakteristische Zahlen $\pi = n$, $\nu = 0$, $\varrho = 0$ sind, die also als Summe von n positiven Quadraten voneinander unabhängiger linearer Funktionen darstellbar ist, eine positive Form, so gilt der Satz, daß für eine

[1]) Hierüber ist zu vergleichen Gundelfingers Zusatz zur dritten Auflage von Hesses analytischer Geometrie des Raumes. Frobenius, „Über das Trägheitsgesetz der quadratischen Formen". Sitzungsber. d. Berliner Akademie 1894; auch in Crelles Journal, Bd. 114. Frobenius gibt ein Verfahren an, um die Signatur, also die Differenz $\pi - \nu$, in gewissen Fällen auch dann aus der Kette (12) zu bestimmen, wenn darin beliebige Glieder verschwinden, so daß ein vorheriges Ordnen der Indices entweder ganz vermieden oder wenigstens eingeschränkt wird.

positive Form keine der Haupt-Unterdeterminanten verschwinden kann. Denn eine positive Form kann für kein reelles Wertsystem der Variablen $x_1, x_2 \ldots x_n$ verschwinden, außer wenn diese Variablen alle gleich Null sind. Wenn aber die Haupt-Unterdeterminante $R_k = 0$ ist, so ist die Funktion von k Variablen

$$\varphi(x_1, x_2 \ldots x_k, 0, \ldots 0),$$

deren Determinante R_k ist, durch weniger als k lineare Funktionen der $x_1, x_2 \ldots x_k$, etwa $y_1, y_2 \ldots y_{k-1}$, darstellbar, und die Funktion φ verschwindet daher, wenn

$$y_1 = 0, \quad \ldots y_{k-1} = 0, \quad x_{k+1} = 0, \quad \ldots x_n = 0$$

ist. Dies ist ein System linearer Gleichungen, das durch nicht verschwindende x befriedigt werden kann. Daraus ergibt sich:

XXXIII. Die notwendige und hinreichende Bedingung für eine positive Form ist die, daß die Glieder der Kette (12) alle positiv sind.

Den entsprechenden Satz für eine negative Form erhält man, wenn man XXXIII auf die Form $-\varphi$ anwendet. Positive und negative Formen werden auch unter dem Namen der definiten Formen zusammengefaßt.

Zweiter Abschnitt.

Zahlen und ganze Funktionen.

§ 13.

Zahlen und Zahlkörper.

Die Algebra hat es zunächst zu tun mit den natürlichen (ganzen positiven) Zahlen, mit denen in bekannter Weise gerechnet wird. Man bezeichnet die Rechnungsarten, Addition, Subtraktion, Multiplikation, Division, als die vier Spezies oder als rationale Operationen. Um diese Operationen allgemeiner ausführen zu können, erweitert man das Zahlenreich durch die Null, die negativen Zahlen und die Brüche, und es zeigt sich, daß die genannten Operationen in diesem Bereiche allgemein ausführbar sind, mit Ausnahme der Division durch Null.

Aber die Algebra braucht auch noch andere Zahlen, wie z. B. die Quadratwurzeln $\sqrt{2}$, $\sqrt{3}$, dann die imaginäre Einheit $i = \sqrt{-1}$. Im Bereich dieser erweiterten Zahlen sind gleichfalls die vier Spezies anwendbar. Hieraus ist ein Begriff abgeleitet, den Kronecker als Rationalitätsbereich, Dedekind als Zahlkörper bezeichnet hat. Zu einem Rationalitätsbereich oder Zahlkörper gehören alle die Zahlen, die aus einem gegebenen System von Zahlen durch Anwendung der vier Spezies erreichbar sind.

Jeder Zahlkörper, zu dem außer der Null noch mindestens eine Zahl a gehört, enthält die Zahl 1 (Division von a durch a), ferner alle ganzen Zahlen (Addition und Subtraktion von Zahlen 1), alle Brüche (Division zweier ganzen Zahlen).

Die ganzen und gebrochenen Zahlen heißen rationale Zahlen. Der Bereich aller rationalen Zahlen ist ein Zahlkörper; er ist in jedem Zahlkörper enthalten und heißt der absolute Rationalitätsbereich.

Ein anderer Körper entspringt aus den komplexen Zahlen $x + yi$, wenn x und y rationale Zahlen sind.

Die ganzen positiven und negativen Zahlen gestatten uneingeschränkt die Addition, die Subtraktion und die Multiplikation; die Division aber nur ausnahmsweise, nämlich dann, wenn der Dividend durch den Divisor teilbar ist.

Wir verstehen daher unter einem Integritätsbereich ein Zahlengebiet, in dem die drei Operationen der Addition, Subtraktion und Multiplikation unbegrenzt ausführbar sind, und nennen eine Zahl a eines solchen Bereichs durch eine Zahl b desselben Bereichs teilbar, wenn es in dem Bereich eine Zahl x gibt, die der Bedingung $bx = a$ genügt.

§ 14.

Variable und Funktionen.

Aus der Analysis ist man gewohnt, unter einer „Variablen" ein Zeichen zu verstehen, das nach und nach verschiedene Werte annimmt. Die Algebra gebraucht das Wort Variable gleichfalls, aber in einem anderen Sinne. Es sind hier lediglich Rechnungssymbole, mit denen man nach den Regeln der Buchstabenrechnung operiert, und mit deren Hilfe man aus den Ergebnissen dieser Rechnungen gewisse Resultate über Zahlen gewinnt oder einfach ausdrückt und zusammenfaßt. Es ist damit nicht ausgeschlossen, daß man für diese Zeichen, für die man gleichfalls Buchstaben wählt, in irgend einem Stadium der Rechnung irgend welche numerische Werte setzt.

Ist x eine solche Variable, während $a_0, a_1, a_2 \ldots$ Zahlen sind, so heißt ein Ausdruck wie

$$(1) \qquad f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$$

eine Funktion von x (ganze rationale Funktion von x); die Zahlen $a_0, a_1 \cdots a_n$ heißen die Koeffizienten. Wenn die Koeffizienten einem Zahlkörper angehören, so heißt $f(x)$ eine Funktion in diesem Körper. Der höchste Exponent n von x heißt der Grad dieser Funktion (wenn a_0 nicht gleich Null ist).

Ganze Funktionen mehrerer Variablen $x, y, z \ldots$ erhält man, wenn man in (1) die Koeffizienten a_i durch ganze Funktionen einer neuen Variablen y ersetzt, in dieser wieder die Koeffizienten durch ganze Funktionen einer dritten Variablen z usf. Man erhält dann Ausdrücke, die man kurz so schreibt:

$$(2) \qquad f(x, y, z \ldots) = \sum^{\alpha, \beta, \gamma \ldots} a_{\alpha, \beta, \gamma} \ldots x^\alpha y^\beta z^\gamma \ldots,$$

worin die Exponenten $\alpha, \beta, \gamma \ldots$ ganze, nicht negative Zahlwerte durchlaufen.

Eine ganze Funktion wird nur dann gleich Null gesetzt, wenn in dem geordneten Ausdruck alle Koeffizienten einzeln verschwinden, und es gilt dann der Satz:

> Sind $\Phi_1(x, y, z \ldots)$, $\Phi_2(x, y, z \ldots)$, $\Phi_3(x, y, z \ldots) \ldots$ nicht verschwindende ganze Funktionen mit numerischen Koeffizienten, so kann man den Veränderlichen $x, y, z \ldots$ auf unendlich viele Arten solche rationale Werte beilegen, daß von den Funktionen $\Phi_1, \Phi_2, \Phi_3 \ldots$ keine den Wert Null erhält.

Der Satz ist zunächst evident, wenn die Funktionen $\Phi_1, \Phi_2, \Phi_3 \ldots$ nur von einer Variablen abhängen; denn dann gibt es überhaupt nur eine endliche Anzahl von Zahlwerten für diese Veränderliche, die eine dieser Funktionen zum Verschwinden bringen (§ 4).

Dann aber können wir die Richtigkeit des Satzes für Funktionen von $n + 1$ Veränderlichen leicht einsehen, falls wir ihn für Funktionen von n Veränderlichen als erwiesen betrachten. Denn ordnen wir die Funktionen nach der $(n + 1)$ten Veränderlichen t, so können wir für die übrigen n Veränderlichen nach Voraussetzung solche Werte setzen, daß in keiner der Funktionen die Koeffizienten aller Potenzen von t verschwinden; dann haben wir Funktionen der einen Veränderlichen t und können für diese einen solchen Wert setzen, daß keine der Funktionen verschwindet. Damit ist der Satz bewiesen, und man sieht, daß er auch noch richtig bleibt, wenn für die absoluten Werte der Zahlen, die für die Variablen zu setzen sind, beliebige obere Grenzen festgesetzt sind, oder wenn gefordert wird, daß nur rationale Zahlen für die Variablen zu setzen seien.

§ 15.

Teilung ganzer Funktionen.

Ist

$$(1) \qquad \Phi(x) = A_0 x^m + A_1 x^{m-1} + \cdots + A_{m-1} x + A_m$$

eine ganze Funktion $m$ten Grades von x und

$$(2) \qquad f(x) = x^n + a_n x^{n-1} + \cdots + a_{n-1} x + a_n$$

eine ganze Funktion $n$ten Grades, in der der Koeffizient von x^n gleich 1 ist, und ist $n \gtrless m$, so ist

$$\Phi(x) - A_0 x^{m-n} f(x) = \Phi_1(x) = A_0' x^{m_1} + \cdots \tag{3}$$

eine ganze Funktion, deren Grad m_1 kleiner als m ist.

Ist m_1 noch nicht kleiner als n, so kann man

$$\Phi_1(x) - A_0' x^{m_1-n} f(x) = \Phi_2(x) \tag{4}$$

setzen und erhält eine Funktion $\Phi_2(x)$ von noch niedrigerem Grade m_2, und so fortfahrend kann man eine Reihe von Funktionen $\Phi_1, \Phi_2, \Phi_3 \ldots$ von abnehmenden Graden $m_1, m_2, m_3 \ldots$ bilden, deren letzte von niedrigerem Grade als n ist.

Faßt man dies zusammen, so erhält man das gewöhnliche Divisionsverfahren von ganzen Funktionen durcheinander. Man kommt zuletzt auf eine Formel von der Gestalt:

$$\Phi(x) = Q(x) f(x) + \varphi(x), \tag{5}$$

worin $\varphi(x)$ höchstens vom Grade $n - 1$ und

$$Q(x) = A_0 x^{m-n} + A_0' x^{m_1-n} + \cdots$$

vom Grade $m - n$ ist.

Worauf es uns hier wesentlich ankommt, ist der Satz, der sich aus der Bildungsweise der Formeln (3), (4) unmittelbar ergibt:

I. Die Funktionen $Q(x)$, $\varphi(x)$ der Formel (5) sind nicht nur ganze Funktionen von x, sondern auch der Koeffizienten $A_0, A_1 \ldots A_m, a_1, a_2 \ldots a_n$.

Die Formel (5) bleibt auch noch gültig, wenn der Koeffizient von x^n in $f(x)$ nicht $= 1$, sondern $= a_0$ ist; nur tritt dann eine Potenz von a_0 im Nenner auf.

Auf diesem Divisionsverfahren beruht die Ermittelung des größten gemeinschaftlichen Teilers zweier ganzer Funktionen $A(x)$, $A'(x)$. Nach dem Satze I läßt sich setzen:

$$\begin{aligned} A &= Q'A' &&+ A'' \\ A' &= Q''A'' &&+ A''' \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A^{(\nu-3)} &= Q^{(\nu-2)} A^{(\nu-2)} &&+ A^{(\nu-1)} \\ A^{(\nu-2)} &= Q^{(\nu-1)} A^{(\nu-1)} &&+ A^{(\nu)}. \end{aligned} \tag{6}$$

Da die Grade der Funktionen A', A'', A''' ... stets abnehmen, so kann man mit dieser Division so lange fortfahren, bis der Grad von $A^{(\nu)}$ gleich Null, also A von x unabhängig geworden ist.

Wenn nun A, A' irgend einen gemeinsamen Teiler haben, so ist dieser, wie der Anblick der Gleichungen (6) lehrt, auch Teiler von A'', A''', A'''' usf. bis $A^{(\nu)}$. Ist $A^{(\nu)}$ eine von Null verschiedene Konstante, so kann also kein von x abhängiger Teiler von A und A' existieren. Solche Funktionen heißen teilerfremd oder relativ prim. Man sagt auch, indem man nur die von x abhängigen Teiler berücksichtigt, die Funktionen haben keinen gemeinsamen Teiler.

Die Bedingung, daß die beiden Funktionen einen gemeinsamen Teiler haben, ist also:

$$A^{(\nu)} = 0. \tag{7}$$

In diesem Falle ist $A^{(\nu-1)}$ Teiler von $A^{(\nu-2)}$, wie die letzte Gleichung (6) zeigt, und nach der vorletzten dieser Gleichungen Teiler von $A^{(\nu-3)}$ usf., also auch Teiler von A und A'. Und da umgekehrt jeder gemeinsame Teiler von A, A' auch Teiler von $A^{(\nu-1)}$ ist, so heißt $A^{(\nu-1)}$ der größte gemeinsame Teiler von A und A'. Der Euklidische Algorithmus (6) zeigt, daß man $A^{(\nu)}$ und $A^{(\nu-1)}$ aus den Koeffizienten von A und A' durch die rationalen Rechenoperationen ableiten kann, und zwar so, daß immer nur durch die Koeffizienten der höchsten Potenzen von x in den Funktionen A, A', A'' ..., die von Null verschieden sind, dividiert wird.

Wir wollen diese Betrachtungen auf ein Beispiel anwenden, das auf die Diskriminante der kubischen Formen führt.

Wir setzen:

$$\begin{aligned} f(x) &= a_0x^3 + a_1x^2 + a_2x + a_3 = A \\ f'(x) &= 3a_0x^2 + 2a_1x + a_2 = B, \end{aligned} \tag{8}$$

und setzen a_0 von Null verschieden voraus, dann haben wir:

$$\begin{aligned} A &= QB + C, \\ C &= c_0x + c_1, \end{aligned} \tag{9}$$

worin

$$Q = -\frac{3a_0x + a_1}{9a_0},$$

$$c_0 = \frac{6a_0a_2 - 2a_1^2}{9a_0}, \qquad c_1 = \frac{9a_0a_3 - a_1a_2}{9a_0}. \tag{10}$$

Wenn c_0 gleich Null ist, so ist hiermit der Algorithmus schon geschlossen; wenn c_1 von Null verschieden ist, dann haben A und B keinen gemeinsamen Teiler; ist aber C gleich Null, so ist B selbst der größte gemeinsame Teiler von A und B, d. h. A ist durch B teilbar. Die Bedingungen hierfür sind also:

$$3a_0a_2 - a_1^2 = 0, \qquad 9a_0a_3 - a_1a_2 = 0. \tag{11}$$

Ist c_0 nicht gleich Null, so gehen wir einen Schritt weiter und setzen:

$$B = PC + D, \tag{12}$$

worin D konstant wird und den Ausdruck erhält:

$$D = \frac{a_2 c_0^2 - 2 a_1 c_0 c_1 + 3 a_0 c_1^2}{c_0^2}. \tag{13}$$

Ist dieser Ausdruck von Null verschieden, so sind A und B teilerfremd, ist er gleich Null, so haben A und B den größten gemeinschaftlichen Teiler C. Setzen wir für c_0, c_1 die Werte (10) ein, lassen den Nenner weg, heben noch den von Null verschiedenen Faktor $9 a_0$ heraus und kehren das Vorzeichen um, so erhält diese Bedingung nach einfacher Rechnung die Gestalt:

$$a_1^2 a_2^2 + 18 a_0 a_1 a_2 a_3 - 4 a_0 a_2^3 - 4 a_1^3 a_3 - 27 a_0^2 a_3^2 = 0. \tag{14}$$

Sie ist, wie man leicht durch Rechnung oder auch aus (13) sieht, auch dann erfüllt, wenn die Bedingungen (11) bestehen, und ist also die notwendige und hinreichende Bedingung dafür, daß $f(x)$ und $f'(x)$ einen gemeinsamen Teiler haben. Die linke Seite von (14), die eine ganze rationale und homogene Funktion der Koeffizienten von $f(x)$ ist, heißt die Diskriminante der Funktion $f(x)$.

Der Algorithmus (6) kann uns noch eine weitere Aufgabe lösen: Aus der ersten dieser Gleichungen folgt:

$$A'' = A - Q'A', \tag{15}$$

und wenn man diesen Wert von A'' in die folgende Gleichung einsetzt:

$$A''' = (1 + Q'Q'')A' - Q''A,$$

also, wenn mit p, p' ganze rationale Funktionen bezeichnet werden,

$$A''' = pA + p'A'. \tag{16}$$

Setzt man die Ausdrücke (15), (16) in die dritte Gleichung (6) ein, so ergibt sich für A'''' wieder ein Ausdruck von der Form (16), und so kann man fortfahren und erhält schließlich:

$$A^{(\nu)} = PA + P'A', \tag{17}$$

worin P, P' ganze rationale Funktionen sind, deren Koeffizienten durch rationale Rechenoperationen aus den Koeffizienten von A und A' zusammengesetzt sind.

Die Gleichung (17) bleibt richtig, wenn man P durch $P - QA'$ und P' durch $P' + QA$ ersetzt, worin Q eine beliebige ganze

Funktion von x ist. Sind nun n und n' die Grade von A und A', so kann man Q so bestimmen, daß der Grad von $P - QA'$ nicht größer als $n' - 1$ wird, und dann folgt, da in (17) die höchsten Potenzen von x sich wegheben müssen, daß der Grad von $P' + QA$ nicht höher als $n - 1$ sein kann.

In der Formel (17) ist $A^{(\nu)}$ eine Konstante. Besonders wichtig ist dieser Satz in dem Falle, wo $A^{(\nu)}$ von Null verschieden, also A, A' relativ prim sind. Setzen wir in diesem Falle

$$P = A^{(\nu)}F(x), \qquad P' = A^{(\nu)}\Phi(x),$$

so können wir nach Weglassung des Faktors $A^{(\nu)}$ dem Satze folgenden Ausdruck geben:

II. Sind $f(x)$ und $\varphi(x)$ zwei ganze Funktionen ohne gemeinsamen Teiler von den Graden n und m, so kann man zwei andere ganze Funktionen $F(x)$ und $\Phi(x)$ bestimmen, deren Grade nicht höher als $m - 1$ und $n - 1$ sind, die der Gleichung

$$F(x)f(x) + \Phi(x)\varphi(x) = 1 \tag{18}$$

identisch genügen.

Der Satz läßt sich noch verallgemeinern. Multiplizieren wir die Gleichung (18) mit einer beliebigen ganzen Funktion $\chi(x)$, so folgt:

$$F(x)\chi(x)f(x) + \Phi(x)\chi(x)\varphi(x) = \chi(x), \tag{19}$$

und nun können wir

$$\Phi(x)\chi(x) = Q(x)f(x) + \psi(x)$$

setzen, so daß $Q(x)$, $\psi(x)$ ganze Funktionen von x sind, und der Grad von $\psi(x)$ kleiner ist als der von $f(x)$. Setzen wir dies in (19) ein und setzen an Stelle von $F(x)\chi(x) + Q(x)\varphi(x)$ wieder $F(x)$, so erhalten wir:

$$F(x)f(x) + \varphi(x)\psi(x) = \chi(x).$$

Wir können daher den vorigen Satz so verallgemeinern:

III. Sind $f(x)$, $\varphi(x)$, $\chi(x)$ gegebene ganze rationale Funktionen, und $f(x)$ und $\varphi(x)$ ohne gemeinsamen Teiler, so lassen sich die ganzen rationalen Funktionen $F(x)$, $\psi(x)$ und zwar $\psi(x)$ von niedrigerem Grade als $f(x)$ so bestimmen, daß die Gleichung

$$F(x)f(x) + \psi(x)\varphi(x) = \chi(x) \tag{20}$$

identisch befriedigt ist.

§ 16.

Zerlegung ganzer Funktionen in lineare Faktoren.

Wenn man eine ganze Funktion $n$ten Grades

(1) $$f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$$

nach § 15 durch eine lineare Funktion $x - \alpha$ dividiert, so wird der Rest eine Konstante. Setzen wir also

(2) $$f(x) = (x - \alpha)\, Q + C,$$

so enthält C die Variable x nicht mehr, und wenn wir

(3) $$Q = q_0 x^{n-1} + q_1 x^{n-2} + \cdots + q_{n-2} x + q_{n-1}$$

setzen, so folgt aus (2):

(4) $$\begin{aligned} f(x) = q_0 x^n &+ q_1 x^{n-1} + \cdots + q_{n-2} x^2 + q_{n-1} x + C \\ &- \alpha q_0 x^{n-1} - \cdots - \alpha q_{n-3} x^2 - \alpha q_{n-2} x - \alpha q_{n-1}, \end{aligned}$$

und aus der Vergleichung mit (1):

(5) $$\begin{aligned} q_0 &= a_0 \\ q_1 - \alpha q_0 &= a_1 \\ q_2 - \alpha q_1 &= a_2 \\ &\cdots\cdots\cdots \\ q_{n-1} - \alpha q_{n-2} &= a_{n-1} \\ C - \alpha q_{n-1} &= a_n. \end{aligned}$$

Daraus erhält man:

(6) $$\begin{aligned} q_0 &= a_0 \\ q_1 &= a_0 \alpha + a_1 \\ q_2 &= a_0 \alpha^2 + a_1 \alpha + a_2 \\ &\cdots\cdots\cdots \\ q_{n-1} &= a_0 \alpha^{n-1} + a_1 \alpha^{n-2} + \cdots a_{n-1} \\ C &= a_0 \alpha^n + a_1 \alpha^{n-1} + \cdots a_{n-1} \alpha + a_n = f(\alpha). \end{aligned}$$

C entsteht aus $f(x)$, wenn man $x = \alpha$ setzt, und kann also auch mit $f(\alpha)$ bezeichnet werden.

Demnach haben wir die Formel:

(7) $$\frac{f(x) - f(\alpha)}{x - \alpha} = Q(x),$$

worin $Q(x)$ eine ganze Funktion vom Grade $n - 1$ ist. Setzen wir darin $x = \alpha$, so ergibt sich:

(8) $$Q(\alpha) = f'(\alpha),$$

worin $f'(x)$ die Derivierte von $f(x)$ ist:

(9) $$f'(x) = n a_0 x^{n-1} + (n-1)\, a_1 x^{n-2} + \cdots + a_{n-1}.$$

Dies Resultat ließe sich natürlich auch durch rationale Rechnung bestätigen.

Wenn wir in den Ausdrücken (6) an Stelle der unbestimmten Größe α das Zeichen x setzen, so entsteht daraus eine Reihe von ganzen rationalen Funktionen von x, die wir, wenn wir der Einfachheit halber $a_0 = 1$ setzen, so schreiben:

$$\begin{aligned} f_0 &= 1, \\ f_1 &= x + a_1, \\ f_2 &= x^2 + a_1 x + a_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_{n-1} &= x^{n-1} + a_1 x^{n-2} + a_2 x^{n-3} \cdots + a_{n-1}. \end{aligned} \tag{10}$$

Diese Funktionen $f_0, f_1 \ldots f_{n-1}$ werden uns später noch gute Dienste leisten. Für jetzt fügen wir noch folgende Bemerkungen bei:

Man kann nach (10) die Potenzen $1, x, x^2 \ldots x^{n-1}$ von x linear ausdrücken durch die Funktionen $f_0, f_1 \ldots f_{n-1}$, und zwar so, daß in den Koeffizienten nur ganze rationale Verbindungen der a vorkommen, z. B.

$$\begin{aligned} 1 &= f_0, \\ x &= f_1 - a_1 f_0, \\ x^2 &= f_2 - a_1 f_1 + (a_1^2 - a_2) f_0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \end{aligned}$$

und daraus folgt, daß man jede ganze rationale Funktion von x, deren Grad nicht größer als $n - 1$ ist, gleichfalls linear durch $f_0, f_1 \ldots f_{n-1}$ ausdrücken kann in der Form

$$y_0 f_0 + y_1 f_1 + \cdots + y_{n-1} f_{n-1}, \tag{11}$$

worin die Koeffizienten $y_0, y_1 \ldots y_{n-1}$ von x unabhängig sind.

Ist also $F(x)$ eine beliebige ganze rationale Funktion von x, so kann man, indem man $f(x)$ als Divisor betrachtet,

$$F(x) = Pf(x) + y_0 f_0 + y_1 f_1 + \cdots + y_{n-1} f_{n-1} \tag{12}$$

setzen, worin auch P eine ganze rationale Funktion von x ist.

Zur rekurrenten Berechnung der Funktionen $f_\nu(x)$ ergibt sich aus (10) die Relation:

$$f_\nu(x) - x f_{\nu-1}(x) = a_\nu. \tag{13}$$

Ist $f(\alpha) = 0$, so heißt α eine Wurzel von $f(x)$, und die Gleichung (2) ergibt:

$$f(x) = (x - \alpha)\, Q(x). \tag{14}$$

Hat die Funktion $Q(x)$ vom $(n-1)$ten Grad eine Wurzel β, so können wir

$$Q(x) = (x - \beta)\, Q_1(x)$$

setzen.

Hat $Q_1(x)$ eine Wurzel γ usf., so ist, wenn $\alpha, \beta, \gamma \ldots \nu$ Wurzeln der Funktionen $f(x)$, $Q(x)$, $Q_1(x) \ldots Q_{n-2}$ sind, der Koeffizient der höchsten Potenz von x in allen Q gleich a_0 und es ergibt sich:

$$f(x) = a_0(x - \alpha)(x - \beta) \ldots (x - \nu). \tag{15}$$

Der Fundamentalsatz der Algebra, auf den wir weiterhin zurückkommen, besagt, daß jede ganze Funktion wenigstens eine Wurzel hat.

Setzen wir diesen voraus, so ergibt sich aus (15) das Theorem:

IV. Eine ganze Funktion $n$ten Grades läßt sich in n lineare Faktoren zerlegen.

Ohne den Fundamentalsatz vorauszusetzen, würde nur folgen, was in § 4 auf anderem Wege bewiesen ist, daß eine Funktion $n$ten Grades nicht mehr als n Wurzeln haben kann, außer wenn sie identisch verschwindet, d. h. alle ihre Koeffizienten $= 0$ sind.

Unter den Faktoren $x - \alpha$ von $f(x)$ kann derselbe auch zweimal oder öfter vorkommen, und man nennt dann α eine doppelte oder mehrfache Wurzel von $f(x)$.

Entwickelt man $f(x)$ nach dem Taylorschen Lehrsatze nach Potenzen von $(x - \alpha)$, so folgt, wenn mit $f'(x)$, $f''(x)$ die Derivierten von $f(x)$ bezeichnet werden:

$$f(x) = f(\alpha) + (x - \alpha) f'(\alpha) + \frac{(x-\alpha)^2}{1.2} f''(\alpha) + \cdots, \tag{16}$$

und wenn $f(\alpha) = 0$ ist, so läßt sich $f(x)$ nach (14) durch $(x - \alpha)$ dividieren, und man erhält:

$$Q(x) = f'(\alpha) + \frac{x-\alpha}{1.2} f''(\alpha) + \frac{(x-\alpha^2)}{1.2.3} f'''(\alpha) + \cdots,$$

also

$$Q(\alpha) = f'(\alpha)$$

oder

$$f'(\alpha) = a_0(\alpha - \beta)(\alpha - \gamma) \ldots (\alpha - \nu). \tag{17}$$

Daraus folgt:

V. Die notwendige und hinreichende Bedingung dafür, daß α eine mehrfache Wurzel ist, besteht darin, daß $f(\alpha)$ und $f'(\alpha)$ zugleich $= 0$ sind, und daß α eine $m$fache Wurzel ist darin, daß

$f(\alpha), f'(\alpha) \ldots f^{(m-1)}(\alpha)$ verschwinden, während $f^{(m)}(\alpha)$ nicht verschwindet.

Setzen wir jetzt $a_0 = 1$ und bezeichnen die n Wurzeln von $f(x)$ mit $\alpha_1, \alpha_2 \ldots \alpha_n$, so ergibt sich:

$$(18) \qquad \begin{aligned} f(x) &= (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n) \\ &= x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots a_n. \end{aligned}$$

Eine leichte Überlegung läßt folgendes Bildungsgesetz der Koeffizienten $a_1, a_2 \ldots a_n$ erkennen:

Es ist $-a_1$ gleich der Summe der α, a_2 gleich der Summe der Produkte von je zweien der α, $-a_3$ die Summe der Produkte von je dreien der α, allgemein $(-1)^\nu a_\nu$ die Summe der Produkte von je ν der Größen α, oder in Formeln ausgedrückt:

$$(19) \qquad \begin{aligned} -a_1 &= \Sigma \alpha_1 \\ +a_2 &= \Sigma \alpha_1 \alpha_2 \\ &\cdots\cdots\cdots\cdots \\ (-1)^\nu a_\nu &= \Sigma \alpha_1 \alpha_2 \ldots \alpha_\nu \\ &\cdots\cdots\cdots\cdots \\ (-1)^n a_n &= \alpha_1 \alpha_2 \ldots \alpha_n, \end{aligned}$$

ein Gesetz, das man aus den Formeln (5) durch vollständige Induktion bestätigt.

Die Anzahl der Glieder, die eine der Summen (18), etwa $(-1)^\nu a_\nu$, enthält, ist gleich der Anzahl der Kombinationen ohne Wiederholung von n Elementen zur νten Klasse, also gleich dem Binomialkoeffizienten

$$(20) \quad B_\nu^{(n)} = \frac{n(n-1)(n-2)\ldots(n-\nu+1)}{1.\,2 \ldots \nu} = \frac{\Pi(n)}{\Pi(\nu)\,\Pi(n-\nu)}.$$

Wenn

$$\Pi(m) = m. = 1.\,2.\,3 \ldots m$$

ist, und wenn man die α_i alle einander gleichsetzt, so erhält man aus (18) den binomischen Lehrsatz.

Die Binomialkoeffizienten $B_\nu^{(n)}$ sind nach (20) ganze Funktionen von n vom Grade ν und behalten ihren guten Sinn, wenn für n eine Variable x gesetzt wird:

$$(21) \qquad B_\nu^{(x)} = \frac{x.(x-1) \ldots (x-\nu+1)}{1.\,2 \ldots \nu}.$$

Man erhält daraus:

$$(22) \qquad B_\nu^{(x)} = B_\nu^{(x-1)} + B_{\nu-1}^{(x-1)},$$

woraus man schließt, was übrigens von vornherein selbstverständlich ist, daß die Binomialkoeffizienten $B_\nu^{(n)}$ für ganze positive n selbst ganze Zahlen sind.

Einen weiteren Satz über diese Koeffizienten wollen wir noch aus dem binomischen Lehrsatz ableiten. Danach ist, wenn $B_0^{(n)} = 1$ gesetzt wird:

$$\begin{aligned} 1 &= B_0^{(0)} \\ 1 + x &= B_0^{(1)} + B_1^{(1)} x \\ (1+x)^2 &= B_0^{(2)} + B_1^{(2)} x + B_2^{(2)} x^2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (1+x)^n &= B_0^{(n)} + B_1^{(n)} x + B_2^{(n)} x^2 + \cdots + B_n^{(n)} x^n. \end{aligned} \tag{23}$$

Wir machen von der Summenformel der geometrischen Reihe Gebrauch:

$$1 + (1+x) + (1+x)^2 + \cdots + (1+x)^n = \frac{(1+x)^{n+1} - 1}{x}.$$

Entwickelt man hier wieder $(1+x)^{n+1}$ nach der Binomialformel, so erhält man für die rechte Seite:

$$\frac{(1+x)^{n+1} - 1}{x} = B_1^{(n+1)} + B_2^{(n+1)} x + \cdots + B_{n+1}^{(n+1)} x^n.$$

Vergleicht man dies mit der Summe der rechten Seiten von (23) und setzt die Koeffizienten entsprechender Potenzen von x einander gleich, so folgt:

$$\begin{aligned} B_0^{(0)} + B_0^{(1)} + B_0^{(2)} + \cdots + B_0^{(n)} &= B_1^{(n+1)} \\ B_1^{(1)} + B_1^{(2)} + \cdots + B_1^{(n)} &= B_2^{(n+1)} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ B_n^{(n)} &= B_{n+1}^{(n+1)}, \end{aligned} \tag{24}$$

oder allgemein

$$B_\nu^{(\nu)} + B_\nu^{(\nu+1)} + \cdots + B_\nu^{(n)} = B_{\nu+1}^{(n+1)}. \tag{25}$$

Wenn man aber die Gleichungen (23) der Reihe nach mit

$$B_0^{(n)},\ -B_1^{(n)},\ +B_2^{(n)},\ \cdots \pm B_n^{(n)}$$

multipliziert (wo das obere Zeichen bei geradem, das untere bei ungeradem n gilt), so ergibt die Summe der linken Seiten nach der Binomialformel:

$$\begin{aligned} B_0^{(n)} - B_1^{(n)}(1+x) + B_2^{(n)}(1+x)^2 - \cdots \pm B_n^{(n)}(1+x)^n \\ = [1 - (1+x)]^n = (-x)^n, \end{aligned}$$

und die Gleichsetzung der Koeffizienten gleich hoher Potenzen auf der rechten und linken Seite liefert das Formelsystem

$$
\begin{aligned}
0 &= B_0^{(0)} B_0^{(n)} - B_0^{(1)} B_1^{(n)} + B_0^{(2)} B_2^{(n)} - \cdots \pm B_0^{(n)} B_n^{(n)} \\
0 &= \quad - B_1^{(1)} B_1^{(n)} + B_1^{(2)} B_2^{(n)} - \cdots \pm B_1^{(n)} B_n^{(n)} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\pm 1 &= \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \pm B_n^{(n)} B_n^{(n)}.
\end{aligned} \tag{26}
$$

Dieses Formelsystem läßt sich in umgekehrter Reihenfolge mit Benutzung der Relation $B_\nu^{(n)} = B_{n-\nu}^{(n)}$ auch so darstellen:

$$
\begin{aligned}
1 &= B_0^{(n)} B_0^{(n)} \\
0 &= B_1^{(n)} B_0^{(n)} - B_0^{(n-1)} B_1^{(n)} \\
0 &= B_2^{(n)} B_0^{(n)} - B_1^{(n-1)} B_1^{(n)} + B_0^{(n-2)} B_2^{(n)} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
0 &= B_n^{(n)} B_0^{(n)} - B_{n-1}^{(n-1)} B_1^{(n)} + B_{n-2}^{(n-2)} B_2^{(n)} - \cdots \pm B_0^{(0)} B_n^{(n)},
\end{aligned} \tag{27}
$$

oder auch in zusammenfassender Bezeichnung:

$$
\begin{aligned}
\sum_{0,\mu}^{\nu} (-1)^\nu B_{\mu-\nu}^{(n-\nu)} B_\nu^{(n)} &= 1 \qquad \mu = 0 \\
&= 0 \qquad \mu = 1, 2 \ldots n.
\end{aligned} \tag{28}
$$

§ 17.

Interpolation.

Wir wollen von den zuletzt gewonnenen Formeln eine Anwendung machen auf eine Aufgabe aus der Theorie der Interpolation.

Es soll eine ganze rationale Funktion $n$ten Grades bestimmt werden, die für $n+1$ gegebene Werte der Veränderlichen x vorgeschriebene Werte hat.

Es handelt sich also um die Bestimmung der $n+1$ Koeffizienten $a_0, a_1 \ldots a_n$ in

$$f(x) = a_0 x^n + a_1 x^{n-1} + \cdots a_{n-1} x + a_n \tag{1}$$

aus den $n+1$ Gleichungen

$$f(\alpha_0) = A_0, \quad f(\alpha_1) = A_1 \ldots f(\alpha_n) = A_n, \tag{2}$$

wenn $\alpha_0, \alpha_1 \ldots \alpha_n$ die gegebenen Werte von x, und $A_0, A_1 \ldots A_n$ die entsprechenden Werte von $f(x)$ sind. Die Gleichungen (2) sind ein System von Gleichungen ersten Grades für die Unbekannten $a_0, a_1 \ldots a_n$, die sich im allgemeinen durch Determinanten auflösen lassen. Sind die gegebenen Werte von x die $n+1$ ersten ganzen Zahlen $0, 1, 2 \ldots n$, so läßt sich die Aufgabe in folgender einfacherer Weise lösen. Nach § 16, (21) läßt sich

x^n und folglich jede ganze Funktion $n$ten Grades $f(x)$ von x linear ausdrücken durch $B_0^{(x)}$, $B_1^{(x)} \dots B_n^{(x)}$ in der Form:

$$f(x) = M_0 B_0^{(x)} + M_1 B_n^{(x)} + \cdots + M_n B_n^{(x)}, \tag{3}$$

worin M_0, $M_1 \dots M_n$ Konstanten sind, und die Funktion $f(x)$ ist bestimmt, wenn diese Konstanten bestimmt sind.

Es sei nun nach unserer Voraussetzung $f(0)$, $f(1)$, $f(2) \dots f(n)$ gegeben; da $B_\nu^{(x)}$ immer verschwindet, wenn x einen der Werte $0, 1, 2 \dots \nu - 1$ hat, so ergeben sich aus (3) die folgenden linearen Gleichungen für die Unbekannten M:

$$\begin{aligned} f(0) &= M_0 B_0^{(0)} \\ f(1) &= M_0 B_0^{(1)} + M_1 B_1^{(1)} \\ f(2) &= M_0 B_0^{(2)} + M_1 B_1^{(2)} + M_2 B_2^{(2)} \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ f(n) &= M_0 B_0^{(n)} + M_1 B_1^{(n)} + M_2 B_2^{(n)} + \cdots + M_n B_n^{(n)}. \end{aligned} \tag{4}$$

Diese Gleichungen sind nun in bezug auf M_0, $M_1 \dots M_n$ aufzulösen, was sehr leicht mit Hilfe der Gleichungen (27) des vorigen Paragraphen geschieht. Die erste Gleichung (4) ergibt nämlich direkt:

$$M_0 = f(0). \tag{5}$$

Multipliziert man die erste Gleichung (4) mit $B_0^{(1)}$, die zweite mit $-B_1^{(1)}$ und addiert, so erhält man nach dem erwähnten Formelsystem (auf $n = 1$ angewandt):

$$-M_1 = B_0^{(1)} f(0) - B_1^{(1)} f(1) \tag{6}$$

und so allgemein, indem man die ν ersten Gleichungen (4) der Reihe nach mit $B_0^{(\nu)}$, $-B_1^{(\nu)}$, $+B_2^{(\nu)}, \dots \pm B_\nu^{(\nu)}$ multipliziert und addiert:

$$\pm M_\nu = B_0^{(\nu)} f(0) - B_1^{(\nu)} f(1) + B_2^{(\nu)} f(2) - \cdots \pm B_\nu^{(\nu)} f(\nu), \tag{7}$$

wodurch nach (3) die Funktion $f(x)$ bestimmt und die Aufgabe gelöst ist. Es ist klar, daß, solange wir über die Werte $f(0)$, $f(1), \dots f(n)$ keine besondere Voraussetzung machen, in dieser Form jede beliebige ganze rationale Funktion von x dargestellt werden kann.

Die Formeln § 16, (22):

$$B_\nu^{(x+1)} = B_\nu^{(x)} + B_{\nu-1}^{(x)}, \qquad B_0^{(x+1)} = B_0^{(x)} = 1 \tag{8}$$

geben aber für die Koeffizienten M_0, $M_1 \dots M_n$ eine Bestimmungsweise, die für die praktische Rechnung viel bequemer ist.

Es ist nämlich nach (3)

$$f(x) = f(0) + M_1 B_1^{(x)} + M_2 B_2^{(x)} + \cdots + M_n B_n^{(x)}; \tag{9}$$

und wenn wir darin x durch $x + 1$ ersetzen und die Differenz

$$\Delta_x = f(x+1) - f(x) \tag{10}$$

bilden, mit Rücksicht auf (8)

$$\Delta_x = M_1 + M_2 B_1^{(x)} + M_3 B_2^{(x)} + \cdots + M_n B_{n-1}^{(x)}, \tag{11}$$

woraus sich ergibt [da (11) eine Gleichung von derselben Art wie (9) ist, nur daß $n - 1$ an Stelle von n getreten ist]:

$$M_1 = \Delta_0 = f(1) - f(0).$$

Setzen wir

$$\begin{aligned} \Delta_{x+1} - \Delta_x &= \Delta'_x \\ \Delta'_{x+1} - \Delta'_x &= \Delta''_x \\ &\cdots\cdots\cdots \\ \Delta_{x+1}^{(n-2)} - \Delta_x^{(n-2)} &= \Delta_x^{(n-1)}, \end{aligned} \tag{12}$$

so wird also hiernach

$$M_0 = f(0), \quad M_1 = \Delta_0, \quad M_2 = \Delta'_0 \ldots M_n = \Delta_0^{(n-1)},$$

und die Formel (9) ergibt:

$$f(x) = f(0) + \Delta_0 B_1^{(x)} + \Delta'_0 B_2^{(x)} + \cdots + \Delta_0^{(n-1)} B_n^{(x)}, \tag{13}$$

und ebenso kann man die Funktionen Δ_x, Δ'_x, Δ''_x ... ausdrücken:

$$\begin{aligned} \Delta_x &= \Delta_0 + \Delta'_0 B_1^{(x)} + \cdots + \Delta_0^{(n-1)} B_{n-1}^{(x)} \\ \Delta'_x &= \Delta'_0 + \Delta''_0 B_1^{(x)} + \cdots + \Delta_0^{(n-1)} B_{n-2}^{(x)} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{14}$$

Die Δ_x, Δ'_x, Δ''_x ... $\Delta_x^{(n-1)}$ sind ganze rationale Funktionen der Grade $n - 1$, $n - 2$ ... 0, die letzte von ihnen also konstant. Um sie alle darzustellen, braucht man nur die Werte $f(0)$, Δ_0, Δ'_0, Δ''_0 ... $\Delta_0^{(n-1)}$, die man am leichtesten berechnet, wenn man eine Tabelle anlegt, die für den Fall $n = 3$ z. B. folgende Form haben würde:

$$\begin{array}{c|c|c|c} f(0) & \Delta_0 & \Delta'_0 & \Delta''_0, \\ f(1) & \Delta_1 & \Delta'_1 & \\ f(2) & \Delta_2 & & \\ f(3) & & & \end{array}$$

und, wenn die $f(0)$, $f(1)$ ... gegeben sind, durch einfache Subtraktionen berechnet wird.

Noch anders wird das Interpolationsproblem nach Lagrange folgendermaßen gelöst:

Da eine Funktion $(n-1)$ten Grades n Koeffizienten enthält, so ist diese durch n vorgeschriebene Werte, die sie für die n willkürlichen Werte des Argumentes x

$$\alpha_1, \quad \alpha_2, \quad \alpha_3 \ldots \alpha_n$$

annehmen soll, bestimmt. Wir setzen

$$f(x) = (x - \alpha_1)\,(x - \alpha_2) \ldots (x - \alpha_n) \tag{15}$$

und erhalten dadurch eine Funktion $n$ten Grades, die für diese Werte α verschwindet, und in der x^n den Koeffizienten 1 hat.

Nehmen wir die n Werte α alle voneinander verschieden an, so sind die Derivierten $f'(\alpha_1)$, $f'(\alpha_2) \ldots f'(\alpha_n)$ alle von Null verschieden. Es ist z. B.

$$f'(\alpha_1) = (\alpha_1 - \alpha_2)\,(\alpha_1 - \alpha_3) \ldots (\alpha_1 - \alpha_n)$$

und in der Funktion

$$F(x) = \sum \frac{A_i f(x)}{f'(\alpha_i)(x - \alpha_i)} \tag{16}$$

$$= \frac{F(\alpha_1)\,f(x)}{f'(\alpha_1)(x-\alpha_1)} + \frac{F(\alpha_2)\,f(x)}{f'(\alpha_2)(x-\alpha_2)} + \cdots + \frac{F(\alpha_n)\,f(x)}{f'(\alpha_n)(x-\alpha_n)}$$

haben wir eine ganze Funktion $(n-1)$ten Grades, die den Bedingungen

$$F(\alpha_1) = A_1, \quad F(\alpha_2) = A_2 \ldots F(\alpha_n) = A_n$$

genügt.

Hierin sind die α_i ganz willkürlich und können daher auch als Variable angesehen werden.

Ist nun $\Phi(x)$ eine beliebige ganze Funktion von x, so setzen wir nach § 15

$$\Phi(x) = Q(x) f(x) + F(x), \tag{17}$$

worin $F(x)$ höchstens vom Grade $n-1$ ist, und es ist

$$\Phi(\alpha_i) = F(\alpha_i).$$

Demnach ergibt sich aus (16).

$$\Phi(x) - Q(x) f(x) = \sum \frac{\Phi(\alpha_i)\,f(x)}{f'(\alpha)\,(x - \alpha_i)} \tag{18}$$

und daraus ergibt sich durch Vergleichung der rechten und linken Seite nach § 15, I. das sehr bemerkenswerte Resultat:

VI. Ist $\Phi(x)$ eine beliebige ganze Funktion von x mit variablen Koeffizienten, so ist die Summe

$$\sum \frac{\Phi(\alpha_i)\,f(x)}{f'(\alpha_i)\,(x - \alpha_i)}$$

eine ganze Funktion, nicht nur von x, sondern auch von α oder den Koeffizienten von f und von den Koeffizienten von Φ.

Ist $\Phi(x) = x^{\nu+1}$, so ist $Q(x) = 0$, wenn $\nu < n-1$, und $Q(x) = 1$, wenn $\nu = n-1$ ist; setzt man dann $x = 0$, so erhält man aus (17):

$$(19)\qquad \begin{aligned} \sum \frac{\alpha_i^{\nu}}{f'(\alpha_i)} &= 0 \quad \nu = 0,\ 1,\ 2 \ldots n-2 \\ \sum \frac{\alpha_i^{n-1}}{f'(\alpha_i)} &= 1. \end{aligned}$$

§ 18.

Entwickelung einer gebrochenen Funktion nach fallenden Potenzen der Variablen.

Ist $f(x)$ irgend eine Funktion vom Grade n, $\Phi(x)$ vom Grade m, ν eine beliebige Zahl und Φ_ν vom Grade $n-1$, so ergibt sich durch Division:

$$(1)\qquad x^{\nu}\Phi(x) = f(x)\{c_{n-m-1}x^{m-n+\nu} + c_{n-m}x^{m-n+\nu-1} + \cdots + c_{\nu-1}\} + \Phi_\nu(x),$$

und die Koeffizienten von Φ_ν und die c sind rational aus den Koeffizienten von f und von Φ abgeleitet. Dividiert man diese Formel durch $x^{\nu}f(x)$, so folgt:

$$(2)\qquad \frac{\Phi(x)}{f(x)} = c_{n-m-1}x^{m-n} + c_{n-m}x^{m-n-1} + \cdots + c_{\nu-1}x^{-\nu} + \frac{\Phi_\nu(x)}{x^{\nu}f(x)}.$$

Der ins Unbestimmte fortgesetzte Teil dieses Ausdruckes

$$(3)\qquad c_{n-m-1}x^{m-n} + c_{n-m}x^{m-n-1} + \cdots$$

heißt die Entwickelung des Bruches $\Phi(x):f(x)$ nach fallenden Potenzen von x und die c_i die Entwickelungskoeffizienten. Das weggelassene Glied $x^{-\nu}\Phi_\nu(x):f(x)$ heißt der Rest. Mit welchem Rechte man unter Umständen die Entwickelung für die Funktion selbst setzen kann, ist eine Frage, die in die Analysis gehört, auf die wir hier nicht einzugehen brauchen.

Wenn $\Phi(x):f(x)$ eine echt gebrochene Funktion ist, so enthält die Entwickelung nach fallenden Potenzen nur negative Exponenten von x. Ist aber $\Phi:f$ unecht gebrochen, so scheidet

sich noch eine ganze Funktion oder eine Konstante ab. Der Teil mit negativen Potenzen

$$\frac{c_0}{x} + \frac{c_1}{x^2} + \frac{c_2}{x^3} + \cdots$$

ist für alle Funktionen Φ derselbe, die bei der Teilung durch f denselben Rest geben.

Die Entwickelung (2) ist in dem Sinne eindeutig bestimmt, daß, wenn wir einen Ausdruck von der Form annehmen:

$$\frac{\Phi(x)}{f(x)} = c'_{n-m-1}x^{m-n} + c'_{n-m}x^{m-n-1} + \cdots + c'_{\nu-1}x^{-\nu} + \frac{\Omega}{x^\nu}, \tag{4}$$

und nur voraussetzen, daß in der gebrochenen Funktion Ω der Grad des Nenners höher sei als der Grad des Zählers, dann notwendig

$$c'_i = c_i, \qquad \Omega = \frac{\Phi_\nu(x)}{f(x)} \tag{5}$$

sein muß. Es folgt nämlich zunächst aus der Annahme (4) durch Multiplikation mit $x^\nu f(x)$, daß $\Omega f(x)$ eine ganze Funktion von x ist, deren Grad also nach unserer Annahme höchstens $= n - 1$ ist, und dann werden die c'_i durch Vergleichung mit der Formel (1) den c_i gleich gefunden, was dann für Ω den Ausdruck (5) nach sich zieht.

Daraus folgt weiter, daß man die Entwickelungskoeffizienten für die Summe von zwei oder mehr gebrochenen Funktionen erhält, wenn man die entsprechenden Koeffizienten der einzelnen Summanden addiert, und daß man die Entwickelung eines Produktes zweier Funktionen dadurch bilden kann, daß man hinlänglich weit fortgesetzte Stücke der Entwickelungen der einzelnen Faktoren miteinander multipliziert, und das Ergebnis wieder nach absteigenden Potenzen der Variablen ordnet. Hierbei kann man einfach die Regel der Multiplikation auf ganze Funktionen von $1:x$ anwenden.

Für den einfachen Bruch

$$\frac{1}{x - \alpha}$$

ergibt sich die Entwickelung:

$$\frac{1}{x} + \frac{\alpha}{x^2} + \frac{\alpha^2}{x^3} + \frac{\alpha^3}{x^4} + \cdots,$$

und daraus, wenn man einen in Partialbrüche zerlegten Bruch hat:

$$\frac{\Phi(x)}{f(x)} = Q + \frac{\Phi(\alpha_1)}{f'(\alpha_1)(x-\alpha_1)} + \frac{\Phi(\alpha_2)}{f'(\alpha_2)(x-\alpha_2)} + \cdots + \frac{\Phi(\alpha_n)}{f'(\alpha_n)(x-\alpha_n)},$$

für die Koeffizienten c_0, c_1, c_2 ... die folgenden Ausdrücke:

$$(6)\quad c_0 = \sum \frac{\Phi(\alpha_i)}{f'(\alpha_i)},\quad c_1 = \sum \frac{\Phi(\alpha_i)\alpha_i}{f'(\alpha_i)},\quad c_2 = \sum \frac{\Phi(\alpha_i)\alpha_i^2}{f'(\alpha_i)} \cdots,$$

worin sich das Summenzeichen auf α_1, α_2 ... α_n erstreckt.

Nehmen wir insbesondere $\Phi(x) = f'(x)$ an, so werden die Größen c_0, c_1, c_2 ... identisch mit den Summen der Potenzen der α:

$$(7)\qquad c_0 = n,\quad c_1 = \Sigma\alpha_i,\quad c_2 = \Sigma\alpha_i^2 \ldots$$

§ 19.

Ganze Funktionen mehrerer Veränderlichen, Formen.

Unter einer ganzen rationalen Funktion $n$ten Grades mehrerer Veränderlichen $F(x, y, z \ldots)$ verstehen wir eine Summe von Gliedern:

$$\Sigma A_{\alpha,\beta,\gamma\ldots} x^\alpha y^\beta z^\gamma \ldots,$$

worin α, β, γ ... ganzzahlige, nicht negative Exponenten sind, deren Summe $\alpha + \beta + \gamma + \cdots$ den Wert n nicht übersteigt, und wenigstens in einem Gliede auch wirklich erreicht. Der Grad wird also bestimmt durch den größten Wert, den die Summe $\alpha + \beta + \gamma + \cdots$ annimmt. Die $A_{\alpha,\beta,\gamma\ldots}$ können beliebige Größen darstellen und heißen die Koeffizienten.

Wenn die Summe der Exponenten $\alpha + \beta + \gamma + \cdots$ in allen Gliedern denselben Wert hat, so heißt die Funktion homogen oder eine Form.

Eine fundamentale Eigenschaft der homogenen Funktionen $n$ten Grades ist die, daß, wenn alle Variablen mit demselben Faktor vervielfältigt werden, der Erfolg derselbe ist, wie wenn die Funktion mit der $n$ten Potenz vervielfältigt wird; in Zeichen, wenn t eine beliebige Veränderliche bedeutet:

$$(1)\qquad F(tx, ty, tz \ldots) = t^n F(x, y, z \ldots);$$

denn ersetzt man in dem Produkt $x^\alpha y^\beta z^\gamma \ldots$ die Variablen durch tx, ty, $tz \ldots$, so erhält es den Faktor

$$t^{\alpha+\beta+\gamma+\cdots};$$

hat nun $\alpha + \beta + \gamma + \cdots$ in allen Gliedern denselben Wert n, so kann der Faktor t^n vor die Summe F herausgenommen werden. Wenn man die Gleichung (1) rechts und links in bezug auf t differentiiert und dann $t = 1$ setzt, so ergibt sich die für homogene Funktionen charakteristische Gleichung von Euler:

$$(2) \qquad nF(x,y,z\ldots) = xF'(x) + yF'(y) + zF'(z) + \cdots$$

Durch Vermehrung der Veränderlichen kann man jede nicht homogene Funktion in eine homogene von gleichem Grade verwandeln. Ist nämlich $m - 1$ die Anzahl der Variablen in einer nicht homogenen Funktion $n$ten Grades, so setzen wir

$$x = \frac{x_1}{x_m}, \qquad y = \frac{x_2}{x_m}, \qquad z = \frac{x_3}{x_m} \cdots,$$

und erhalten in

$$x_m^n F\left(\frac{x_1}{x_m}, \frac{x_2}{x_m}, \frac{x_3}{x_m} \ldots\right)$$

eine ganze homogene Funktion $n$ten Grades der Variablen $x_1, x_2 \ldots x_m$, die wir mit

$$\Phi(x_1, x_2 \ldots x_m)$$

bezeichnen.

Es empfiehlt sich bisweilen, die homogenen Funktionen mehrerer Variablen mit den Polynomialkoeffizienten zu schreiben.

Wir setzen daher

$$(3) \qquad \Phi(x_1, x_2 \ldots x_m) = \sum \frac{\Pi(n)}{\Pi(\alpha_1)\,\Pi(\alpha_2) \ldots \Pi(\alpha_m)} A_{\alpha_1, \alpha_2 \ldots \alpha_m} x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m},$$

wo sich die Summe auf alle nicht negativen, der Bedingung

$$(4) \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_m = n$$

genügenden Zahlen erstreckt. Diese Bezeichnungsweise, ohne die Beschränkung (4), ist auch auf nicht homogene Funktionen anwendbar.

Man kann aber die homogene Funktion auch so darstellen:

$$(5) \qquad \Phi(x_1, x_2 \ldots x_m) = \Sigma A_{\nu_1, \nu_2 \ldots \nu_n} x_{\nu_1} x_{\nu_2} \ldots x_{\nu_n},$$

worin jeder der Indices $\nu_1, \nu_2 \ldots \nu_n$ von den übrigen unabhängig die Wertreihe 1, 2 ... m zu durchlaufen hat. Die Summe (4) besteht also aus m^n Gliedern, die aber nicht alle voneinander verschieden sind. Das Produkt $x_{\nu_1} x_{\nu_2} \ldots x_{\nu_n}$ bleibt nämlich ungeändert, wenn die Indices $\nu_1, \nu_2 \ldots \nu_n$ beliebig untereinander permutiert werden. Die Anzahl der Permutationen von n Elementen beträgt aber $n!$ oder $\Pi(n)$. Sind unter diesen Elementen je $\alpha_1, \alpha_2 \ldots$ einander gleich, so reduziert sich die Zahl der Permutationen auf

$$\frac{\Pi(n)}{\Pi(\alpha_1)\,\Pi(\alpha_2)\ldots},$$

woraus sich ergibt, daß in (5) irgend ein Produkt $x_1^{\alpha_1} x_2^{\alpha_2} \ldots$ genau

$$\frac{\Pi(n)}{\Pi(\alpha_1)\,\Pi(\alpha_2)\ldots}$$

mal vorkommt. Setzt man also noch fest, daß $A_{\nu_1, \nu_2 \ldots \nu_n}$ sich nicht ändern soll, wenn die Indices beliebig permutiert werden, so erweisen sich die Bezeichnungsweisen (3) und (5) als identisch, wenn durch Zusammenfassen gleicher Faktoren

$$x_{\nu_1} x_{\nu_2} \ldots x_{\nu_n} = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m}$$

und

$$A_{\nu_1, \nu_2 \ldots \nu_n} = A_{\alpha_1, \alpha_2 \ldots \alpha_m}$$

gesetzt wird.

Bezeichnen wir die Anzahl der Glieder, die in der Funktion Φ [nach (3)] auftreten, mit (m, n), so findet man, indem man zunächst die Glieder zählt, die den Faktor x_1 haben. und dann die übrigen, die eine homogene Funktion $n$ter Ordnung von den übrigen $m - 1$ Variablen bilden, die Rekursionsformel:

$$(6) \qquad (m, n) = (m, n-1) + (m-1, n),$$

mit deren Hilfe man durch vollständige Induktion den Ausdruck:

$$(7) \qquad (m, n) = \frac{m(m+1)\ldots(m+n-1)}{1.2\ldots n} = \frac{\Pi(m+n-1)}{\Pi(n)\,\Pi(m-1)},$$

der sowohl für $n = 1$ als für $m = 1$ richtig ist, als allgemein gültig erkennt.

Die ganzen homogenen Funktionen werden auch Formen genannt. Man unterscheidet nach der Anzahl der Variablen unäre (einfache Potenzen), binäre, ternäre, quaternäre Formen. Die Theorie der binären Formen ist im wesentlichen identisch mit der Theorie der ganzen rationalen Funktionen einer

Veränderlichen. Man gelangt von den binären Formen zu diesen Funktionen zurück, wenn man eine der homogenen Variablen durch eine Konstante, z. B. durch die Zahl 1 ersetzt.

§ 20.

Zerlegbare und unzerlegbare Funktionen. Primfunktionen.

Eine ganze Funktion der Variablen $x, y, z \ldots$ heißt geordnet, wenn jedes Glied $x^\alpha y^\beta z^\gamma \ldots$ nur einmal vorkommt. Die geordnete Funktion

$$\Sigma A_{\alpha,\beta,\gamma\ldots} x^\alpha y^\beta z^\gamma \ldots$$

wird nur dann gleich Null gesetzt, wenn die Koeffizienten $A_{\alpha,\beta,\gamma\ldots}$, die im allgemeinen beliebige Zahlen sind, alle den Wert Null haben. Man kann den Gliedern einer solchen Funktion in folgender Weise eine bestimmte Rangordnung geben, nachdem den Variablen $x, y, z \ldots$ eine feste Reihenfolge gegeben ist. Sind

$$X = x^\alpha y^\beta z^\gamma \ldots, \quad X' = x^{\alpha'} y^{\beta'} z^{\gamma'} \ldots$$

zwei Glieder der Funktion, so hat X einen höheren Rang als X', wenn $\alpha > \alpha'$ oder $\alpha = \alpha'$, $\beta > \beta'$, oder $\alpha = \alpha'$, $\beta = \beta'$, $\gamma > \gamma'$ ist, mit anderen Worten, wenn von den Differenzen

$$\alpha - \alpha', \quad \beta - \beta', \quad \gamma - \gamma' \ldots$$

die erste von Null verschiedene positiv ist. Dann kann man die Glieder der Funktion $F(x, y, z \ldots)$ in eindeutiger Weise nach der Höhe des Ranges ordnen.

Das Produkt von ganzen Funktionen ist wieder eine ganze Funktion. Multipliziert man zwei geordnete Funktionen, so folgt, daß das Produkt der Glieder höchsten Ranges in dem Produkt der Funktionen das Glied höchsten Ranges ergibt Daraus folgt, daß ein Produkt von mehreren ganzen Funktionen nur dann verschwinden kann, wenn wenigstens einer seiner Faktoren verschwindet.

Der Grad eines Produktes ist gleich der Summe der Grade seiner Faktoren.

Dies ist zunächst evident, wenn die Faktoren, und folglich auch das Produkt, homogene Funktionen sind, und es folgt dann allgemein aus der Bemerkung, daß sich jede nicht homogene ganze Funktion in eine Summe von homogenen Funktionen ver-

schiedener Grade zerlegen läßt, deren höchster den Grad der Funktion bestimmt.

Wir haben nun unter den ganzen Funktionen solche zu unterscheiden, die als Produkte von zwei oder mehr ganzen Funktionen, deren keine vom nullten Grade (also konstant) ist, dargestellt werden können, und solche, bei denen dies nicht möglich ist. Die ersten heißen zerlegbar, die anderen unzerlegbar.

Es ist dabei zu unterscheiden zwischen solchen Funktionen, die in keinem Rationalitätsbereich zerlegbar sind, wie z. B. $x^2 + y^2 + z^2$, und solchen, die zwar in einem, z. B. dem absoluten Rationalitätsbereich unzerlegbar sind, aber in einem erweiterten Rationalitätsbereich zerlegbar werden, wie z. B. $x^2 + y^2$, das im absoluten Rationalitätsbereich nicht zerlegt werden kann, obwohl es in die beiden komplexen Faktoren $(x + yi)$ $(x - yi)$ zerlegbar ist.

Hier wollen wir einen beliebigen, aber festen Rationalitätsbereich $\mathfrak{A}$ für die Koeffizienten voraussetzen und sprechen daher von Funktionen in $\mathfrak{A}$ und Zerlegbarkeit und Unzerlegbarkeit in $\mathfrak{A}$. Die absolute Unzerlegbarkeit ist darin als Spezialfall enthalten, weil man unter $\mathfrak{A}$ auch die Gesamtheit aller Zahlen verstehen kann.

Wenn eine ganze Funktion zerlegbar ist, so ist der Grad jedes der Faktoren niedriger als der Grad der Funktion selbst. Eine lineare Funktion ist also immer unzerlegbar, und jede ganze Funktion läßt sich in eine endliche Zahl unzerlegbarer Faktoren zerlegen.

Eine ganze Funktion W ist durch eine andere w teilbar, wenn eine dritte ganze Funktion (oder auch eine Konstante) w' existiert, so daß

$$W = w\,w'$$

ist. Daraus folgt dann, daß, wenn U eine durch w teilbare und V eine beliebige ganze Funktion ist, auch das Produkt $U\,V$ durch w teilbar ist, und daß, wenn U, U', U'' ... durch w teilbare, V, V', V'' ... beliebige ganze Funktionen sind, auch

$$U\,V + U'V' + U''V'' + \cdots$$

durch w teilbar ist.

Ist W durch w teilbar, so sagt man auch, w geht in W auf.

Zwei ganze Funktionen U, V, die nicht durch eine und dieselbe ganze Funktion teilbar sind, heißen relativ prim oder teilerfremd.

Wir beweisen den Satz:

1. Sind U, V, v ganze Funktionen irgend welcher Veränderlichen, sind U und v relativ prim und UV durch v teilbar, so ist V durch v teilbar.

Dieser Satz entspricht genau einem bekannten Fundamentalsatz aus der Lehre von den ganzen Zahlen, daß nämlich, wenn ein Produkt von zwei ganzen Zahlen durch eine dritte ganze Zahl teilbar ist, die zu dem einen Faktor teilerfremd ist, der andere Faktor durch diese Zahl teilbar sein muß.

Wir beweisen ihn durch vollständige Induktion. Sind U, V, v nur von einer Veränderlichen x abhängig, so ist der Satz richtig; denn nach § 15, II. kann man in diesem Falle, wenn U und v relativ prim sind, zwei andere ganze Funktionen P und p von x so bestimmen, daß

$$PU + pv = 1,$$

woraus durch Multiplikation mit V

$$PUV + pVv = V$$

folgt, und daraus ersieht man, daß, wenn UV durch v teilbar ist, auch V durch v teilbar sein muß.

Wir nehmen also an, der Satz, den wir beweisen wollen, sei für Funktionen von n und weniger Veränderlichen bewiesen, und wir leiten daraus seine Richtigkeit für Funktionen von $n + 1$ Veränderlichen ab.

Dazu ist erforderlich, daß wir aus dem als richtig vorausgesetzten Theorem 1. einige Folgerungen ziehen, als deren Schluß sich dann die Gültigkeit des Theorems für die nächst höhere Variablenzahl ergibt.

Ist v eine unzerlegbare Funktion, so ist eine andere Funktion U derselben Veränderlichen entweder relativ prim zu v oder durch v teilbar. Daraus folgt nach 1., daß ein Produkt von zwei Funktionen UV nur dann durch v teilbar sein kann, wenn einer der beiden Faktoren durch v teilbar ist. Dasselbe gilt für ein Produkt von mehreren Funktionen, und so folgt aus 1. das Theorem:

2. Ein Produkt aus mehreren ganzen Funktionen ist nur dann durch eine unzerlegbare Funktion v teilbar, wenn wenigstens einer der Faktoren des Produktes durch v teilbar ist.

Wenn eine ganze rationale Funktion U auf zwei Arten in unzerlegbare Faktoren zerlegt ist,

$$U = v\,v'\,v'' \ldots = w\,w'\,w'' \ldots,$$

so muß nach 2. wenigstens einer der Faktoren v, v', $v'' \ldots$ durch w teilbar sein, also etwa v. Dann aber kann, da auch v unzerlegbar ist, v von w nur durch einen konstanten Faktor verschieden sein.

Demnach ist, wenn c dieser konstante Faktor ist,

$$c\,v'\,v'' \ldots = w'\,w'' \ldots,$$

woraus folgt, daß eine der Funktionen v', $v'' \ldots$, etwa v', durch w' teilbar ist, und sich also von w' nur durch einen konstanten Faktor unterscheidet usf.

Wir erhalten also als zweite Folgerung aus dem Theorem 1.:

3. Eine ganze Funktion kann, von konstanten Faktoren abgesehen, nur auf eine Art in unzerlegbare Faktoren zerlegt werden.

Hieraus ergibt sich der Begriff des größten gemeinschaftlichen Teilers von zwei oder mehr ganzen rationalen Funktionen U, $V \ldots$ Man versteht darunter das Produkt aller unzerlegbaren Faktoren, die in den Zerlegungen jeder der Funktionen U, $V \ldots$ vorkommen, oder die Funktion möglichst hohen Grades, die in allen Funktionen U, $V \ldots$ aufgeht. Nach 3. ist diese Funktion, von einem konstanten Faktor abgesehen, für jedes Funktionensystem U, $V \ldots$ vollständig bestimmt.

Mehrere Funktionen U, V, $W \ldots$ heißen relativ prim, wenn es keine ganze Funktion gibt, die in allen aufgeht.

Alle diese Definitionen und Sätze sind genau analog mit sehr bekannten Sätzen der elementaren Zahlenlehre, die dort als Folgerungen des Algorithmus des größten gemeinschaftlichen Teilers auftreten, nur daß hier die unzerlegbaren Funktionen die Rolle der Primzahlen übernehmen.

Wir führen noch einen solchen Satz an:

4. Sind u, v relativ prim und Uu, Uv durch w teilbar, so ist U durch w teilbar.

Denn zerlegt man die ganzen rationalen Funktionen U, u, v, w in ihre unzerlegbaren Faktoren, so muß irgend ein Faktor von w, da er nicht in u und v zugleich vorkommen kann, in U aufgehen. Hebt man ihn aus U und w weg, so kann man ebenso für einen nächsten Faktor von w schließen usf.

Wir betrachten nun ganze rationale Funktionen einer Veränderlichen t

$$f(t) = u_0 t^m + u_1 t^{m-1} + \cdots + u_{m-1} t + u_m,$$

deren Koeffizienten $u_0, u_1 \ldots$ ganze rationale Funktionen von n Veränderlichen x sind, von denen t unabhängig ist.

Sind die Koeffizienten $u_0, u_1 \ldots u_m$ ohne gemeinsamen Teiler, so heißt $f(t)$ primitiv, im anderen Falle wird der größte gemeinschaftliche Teiler der Koeffizienten $u_0, u_1 \ldots u_m$ der Teiler der Funktion $f(t)$ genannt.

Wir schließen, immer unter Voraussetzung der Gültigkeit von 1.:

5. Das Produkt von zwei primitiven Funktionen $f(t)$ und $\varphi(t)$ ist wieder eine primitive Funktion und der Teiler eines Produktes zweier imprimitiver Funktionen ist gleich dem Produkt der Teiler beider Faktoren.

Es seien

$$\begin{aligned} f(t) &= u_0 t^m + u_1 t^{m-1} + \cdots + u_m \\ \varphi(t) &= v_0 t^\mu + v_1 t^{\mu-1} + \cdots + v_\mu \end{aligned} \tag{1}$$

zwei primitive Funktionen und

$$F(t) = U_0 t^{m+\mu} + U_1 t^{m+\mu-1} + \cdots + U_{m+\mu} \tag{2}$$

ihr Produkt. Es ist dann, wenn r und s irgend zwei Ziffern aus den Reihen 0, 1, 2 ... m und 0, 1, 2 ... μ bedeuten:

$$\begin{aligned} U_{r+s} = u_r v_s &+ u_{r-1} v_{s+1} + \cdots \\ &+ u_{r+1} v_{s-1} + \cdots \end{aligned} \tag{3}$$

Wenn nun w irgend eine unzerlegbare Funktion ist, die weder in allen u_r noch in allen v_s aufgeht, so wählen wir in (3) r und s so, daß u_r das erste nicht durch w teilbare u ist, also $u_{r-1}, u_{r-2} \ldots$ durch w teilbar sind, und daß ebenso v_s das erste, nicht durch w teilbare v wird. Dann kann auch U_{r+s} nicht durch w teilbar sein, weil alle Glieder, mit Ausnahme des ersten, durch w teilbar sind, d. h. $F(t)$ ist primitiv.

Daraus folgt unmittelbar, wenn p und q irgend welche ganze rationale Funktionen der x sind, daß pq der Teiler des Produktes der beiden imprimitiven Funktionen $pf(t)$, $q\varphi(t)$ ist, also der zweite Teil des Satzes 5. Daraus folgt weiter:

6. Wenn eine ganze rationale Funktion $F(t)$ der $n+1$ Variablen x und t in zwei Faktoren zerlegbar ist, die in bezug auf t ganz, in bezug auf die x wenigstens rational sind, so ist sie auch in zwei Funktionen zerlegbar, die in x und t ganz und rational sind.

Denn nach der Voraussetzung gibt es eine ganze rationale Funktion w der x allein und zwei ganze rationale Funktionen der x und t, $f_1(t)$, $\varphi_1(t)$, so daß:

$$wF(t) = f_1(t)\varphi_1(t),$$

und nach 5. muß w in dem Produkt der Teiler $f_1(t)$ und $\varphi_1(t)$ aufgehen, so daß wir auch:

$$F(t) = f(t)\varphi(t)$$

erhalten, worin $f(t)$, $\varphi(t)$ ebenso wie $f_1(t)$, $\varphi_1(t)$ ganze rationale Funktionen von x und t, und zwar in bezug auf t von demselben Grade wie $f_1(t)$ und $\varphi_1(t)$ sind, w. z. b. w.

Hieraus schließen wir weiter, daß zwei Funktionen $F(t)$, $f(t)$, die als ganze rationale Funktionen der $n+1$ Veränderlichen x und t betrachtet, in dem oben definierten Sinne relativ prim sind, sich auch, als Funktionen von t allein betrachtet und nach dem Algorithmus des größten gemeinschaftlichen Teilers behandelt, als relativ prim erweisen müssen. Denn wenn sie einen gemeinsamen Teiler hätten, der in bezug auf t ganz, in bezug auf die x gebrochen wäre, so ließe sich eine ganze Funktion T von x und t und eine ganze Funktion P der x allein so bestimmen, daß

$$PF(t) = TF_1(t), \qquad Pf(t) = Tf_1(t)$$

wäre, worin F_1, f_1 ganze Funktionen ohne gemeinsamen Teiler sind. Wenn also P_1, p_1, M die Teiler der Funktionen $F_1(t)$, $f_1(t)$, T sind, so sind P_1 und p_1 relativ prim, und P_1M und p_1M sind durch P teilbar, also ist nach 4. auch M und folglich T durch P teilbar; mithin sind $F(t)$ und $f(t)$ durch die ganze Funktion $T:P$ teilbar, also nicht relativ prim, wie doch vorausgesetzt war.

Es ergibt sich also nach § 15, daß sich zwei ganze Funktionen Q, q von x und t und eine ganze Funktion X von den x allein so bestimmen läßt, daß die Identität

$$QF(t) + qf(t) = X \tag{4}$$

besteht.

Ist nun $\Phi(t)$ eine weitere ganze Funktion von x und t, so folgt aus (4) durch Multiplikation mit $\Phi(t)$:

$$Q\Phi(t)\,F(t) + q\Phi(t)\,f(t) = X\Phi(t),$$

und wenn also $\Phi(t)\,F(t)$ durch $f(t)$ teilbar ist, so ist auch $X\Phi(t)$ durch $f(t)$ teilbar.

Demnach können wir, wenn $\varphi(t)$ und $\psi(t)$ wieder zwei ganze Funktionen von x und t bedeuten, setzen:

$$\begin{aligned} X\Phi(t) &= \varphi(t)\,f(t) \\ F(t)\Phi(t) &= \psi(t)\,f(t). \end{aligned} \tag{5}$$

Multipliziert man die zweite dieser Gleichungen mit X und setzt aus der ersten für $X\Phi(t)$ den Ausdruck $\varphi(t)\,f(t)$ ein, so läßt sich $f(t)$ wegheben und es folgt:

$$\varphi(t)\,F(t) = X\psi(t).$$

Es muß also X sowohl im Teiler von $\varphi(t)\,f(t)$ als in dem von $\varphi(t)\,F(t)$ aufgehen, und da $f(t)$ und $F(t)$ und mithin auch ihre Teiler relativ prim sind, so muß X im Teiler von $\varphi(t)$ aufgehen (nach 4.). Setzen wir demnach:

$$\varphi(t) = X\varphi_1(t),$$

so folgt aus (5):

$$\Phi(t) = \varphi_1(t)\,f(t),$$

d. h. $\Phi(t)$ ist durch $f(t)$ teilbar; also:

7. Sind $F(t)$ und $f(t)$ relativ prim und $\Phi(t)\,F(t)$ durch $f(t)$ teilbar, so ist $\Phi(t)$ durch $f(t)$ teilbar.

Dies aber ist nichts anderes als das Theorem 1. für Funktionen von $n + 1$ Variablen, und 1. somit allgemein bewiesen. Zugleich sind damit auch die aus 1. gezogenen Folgerungen bewiesen, insbesondere die, daß eine Funktion von einer beliebigen Anzahl von Variablen, abgesehen von konstanten Faktoren, nur auf eine Art in unzerlegbare Faktoren zerlegt werden kann. Die unzerlegbaren Funktionen heißen aus diesem Grunde auch Primfunktionen und die unzerlegbaren Faktoren einer Funktion ihre Primfaktoren.

Sind die Koeffizienten einer ganzen Funktion einer Variablen ganze Zahlen, so heißt der größte gemeinschaftliche Teiler dieser Koeffizienten der Teiler der Funktion, und wenn der Teiler $= 1$ ist, so heißt die Funktion primitiv. Wenn man also unter $u_0, u_1 \ldots v_0, v_1 \ldots$ ganze Zahlen versteht, so ergibt das Theorem V den Satz von Gauß:

8. Der Teiler eines Produktes von mehreren ganzzahligen Funktionen einer Variablen ist gleich dem Produkt der Teiler der Faktoren und das Produkt mehrerer primitiver Funktionen ist eine primitive Funktion.

Gauß gibt diesem Satz den folgenden Ausdruck:

Wenn die Koeffizienten $A, B, C \ldots N$; $a, b, c \ldots n$ zweier Funktionen der Form

$$(P) \quad x^m + Ax^{m-1} + Bx^{m-2} + Cx^{m-3} + \cdots N$$
$$(Q) \quad x^\mu + ax^{\mu-1} + bx^{\mu-2} + cx^{\mu-3} + \cdots n$$

alle rational sind, aber nicht alle ganzzahlig und das Produkt aus (P) und (Q)

$$= x^{m+\mu} + \mathfrak{A}x^{m+\mu-1} + \mathfrak{B}x^{m+\mu-2} \ldots + \mathfrak{Z},$$

so können nicht alle Koeffizienten $\mathfrak{A}, \mathfrak{B}, \ldots \mathfrak{Z}$ ganze Zahlen sein[1]).

§ 21.

Zerlegung ganzer Funktionen im absoluten Rationalitätsbereich.

Um zu entscheiden, ob eine ganze Funktion einer Variablen, deren Koeffizienten rationale Zahlen sind, zerlegbar oder unzerlegbar ist, und gegebenenfalls die Faktoren zu finden, hat Kronecker ein Verfahren angegeben, das durch eine endliche Zahl von Schritten zum Ziele führt.

Eine zerlegbare Funktion heißt auch reduzibel, eine unzerlegbare irreduzibel, wobei aber nochmals zu betonen ist, daß diese Eigenschaft nicht den Funktionen an sich anhaftet, sondern erst nach Festsetzung des Rationalitätsbereichs einen bestimmten Sinn erhält. Hier ist der Rationalitätsbereich der absolute.

---

[1]) Gauß, Disquisitiones arithmeticae, Art. 42.

Wir nehmen also eine ganze Funktion einer Variablen $F(x)$ vom Grade μ:

$$(1) \qquad F(x) = a_0 x^\mu + a_1 x^{\mu-1} + a_2 x^{\mu-2} + \cdots + a_\mu,$$

deren Koeffizienten ganze rationale Zahlen sind. Wollen wir $F(x)$ in seine irreduziblen Faktoren zerlegen, so genügt es, alle Faktoren $\varphi(x)$ von $F(x)$ zu ermitteln, deren Grad ν nicht größer als $^1/_2\,\mu$ ist, da, wenn $F(x)$ zerlegbar ist, wenigstens einer der Faktoren einen solchen Grad haben muß. Es sei also:

$$(2) \qquad F(x) = \varphi(x)\,\varphi_1(x).$$

Um das Verfahren auch auf Funktionen mit gebrochenen Koeffizienten anzuwenden, multiplizieren wir die Gleichung (2) mit einem so großen ganzzahligen Faktor (dem Hauptnenner), daß alle Koeffizienten ganze Zahlen werden. Dann ist der Teiler von F gleich dem Produkt der Teiler von φ und φ_1 und kann weggelassen werden. In (2) können wir also F, φ, φ_1 als primitive Funktionen annehmen. Ist dann r eine beliebige ganze rationale Zahl, so werden $F(r)$, $\varphi(r)$, $\varphi_1(r)$ auch ganze rationale Zahlen, und es muß $F(r)$ wegen (2) durch $\varphi(r)$ teilbar sein. Nehmen wir nun $\nu + 1$ willkürliche, voneinander verschiedene ganze Zahlen $r_0, r_1, r_2 \ldots r_\nu$, so müssen sich die Zahlen

$$(3) \qquad \varphi(r_0),\ \varphi(r_1),\ \ldots\ \varphi(r_\nu)$$

unter den Teilern der Zahlen

$$(4) \qquad F(r_0),\ F(r_1) \ldots F(r_\nu)$$

finden, und da die Zahlen (4) durch die Funktion F selbst gegeben sind, so gibt es nur eine endliche Anzahl von zulässigen Annahmen für die Zahlen (3). Durch die Werte (3) ist aber die Funktion $\varphi(x)$ selbst vollkommen bestimmt, etwa wenn

$$f(x) = (x - r_0)\,(x - r_1) \ldots (x - r_\nu)$$

gesetzt wird, durch die Interpolationsformel von Lagrange (§ 17):

$$(5) \qquad \varphi(x) = \sum_{0,\nu}^{i} \frac{\varphi(r_i)\,f(x)}{(x - r_i)\,f'(r_i)}.$$

Man erhält so eine endliche Anzahl von möglichen Bestimmungen der Funktion $\varphi(x)$, und muß mit jeder dieser Funktionen den Versuch machen, ob sie in $F(x)$ enthalten ist.

Das hier geschilderte Verfahren ist unter Umständen auch auf Funktionen in anderen Körpern anwendbar, dann nämlich,

wenn in dem Körper die Zerlegung der ganzen Größen in ihre Primfaktoren ebenso wie bei den ganzen Zahlen als bestimmt vorausgesetzt werden kann.

Durch vollständige Induktion kann man dieses Verfahren auch auf Funktionen mehrerer Variablen anwenden.

Wir nehmen an, daß man bereits ein Verfahren kennt, um Funktionen von $n-1$ Variablen $y, z \ldots$ in Primfaktoren zu zerlegen. Dann sind die Koeffizienten der Funktionen F, φ, φ_1 nicht mehr ganze Zahlen, sondern ganze Funktionen von $y, z \ldots$. Nach § 20 sind diese Funktionen in Primfaktoren zu zerlegen und daher gilt auch der Satz, daß der Teiler von $F(x)$ gleich dem Produkt der Teiler von $\varphi(x)$ und $\varphi_1(x)$ ist. Für $r_0, r_1, r_2 \ldots r_\nu$ können dann ebenfalls ganze Zahlen oder auch ganze Funktionen von $y, z \ldots$ gesetzt werden[1]).

Wir schließen diese Betrachtungen mit einem speziellen, aber sehr einfachen und nützlichen Kennzeichen für irreduzible Funktionen, das von Schönemann herrührt[2]).

Sind die ganzen rationalen Koeffizienten $a_1, a_2 \ldots a_n$ einer Funktion $n$ten Grades:

$$(6) \quad f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n$$

durch eine Primzahl p teilbar, a_n aber nicht durch p^2, so ist $f(x)$ irreduzibel.

Um dies zu beweisen, sei $f(x)$ zerlegbar in die beiden Faktoren:

$$(7) \quad \begin{aligned} \varphi(x) &= x^\mu + b_1 x^{\mu-1} + b_2 x^{\mu-2} + \cdots + b_{\mu-1} x + b_\mu, \\ \psi(x) &= x^\nu + c_1 x^{\nu-1} + c_2 x^{\nu-2} + \cdots + c_{\nu-1} x + c_\nu, \end{aligned}$$

worin die Koeffizienten b, c rationale und folglich nach § 20, 8. ganze Zahlen sind.

Vergleichen wir das Produkt der beiden Funktionen (7) mit (6), so ergibt sich:

$$(8) \quad \begin{aligned} a_n &= b_\mu c_\nu \\ a_{n-1} &= b_\mu c_{\nu-1} + b_{\mu-1} c_\nu \\ a_{n-2} &= b_\mu c_{\nu-2} + b_{\mu-1} c_{\nu-1} + b_{\mu-2} c_\nu \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_\nu &= b_\mu c_{\nu-\mu} + b_{\mu-1} c_{\nu-\mu-1} + \cdots + b_1 c_{\nu-1} + c_\nu. \end{aligned}$$

---

[1]) Kronecker, Crelles Journal, Bd. 94; Runge, ebenda, Bd. 99.

[2]) Der Satz wird auch nach Eisenstein benannt.

Diese Formeln sind auch dann noch richtig, wenn $\mu \gtrless \nu$ ist, falls wir $c_0 = 1$ und jedes c mit negativem Index $= 0$ setzen.

Wenn nun a_n durch p, aber nicht durch p^2 teilbar ist, so ist von den beiden Faktoren b_μ, c_ν nur der eine durch p teilbar, nicht der andere; sei also c_ν durch p nicht teilbar, b_μ durch p teilbar. Dann folgt, wenn die übrigen $a_{n-1}, a_{n-2} \ldots a_\nu$ alle durch p teilbar sind, daß $b_{\mu-1}, b_{\mu-2} \ldots b_1$ durch p teilbar sein müssen, was nach der letzten Gleichung (8) unmöglich ist. Damit ist die Zerlegung $f = \varphi\,\psi$ als unmöglich dargetan.

Dritter Abschnitt.

Symmetrische Funktionen.

§ 22.

Symmetrische Grundfunktionen.

Wir betrachten jetzt ganze Funktionen beliebigen Grades von einer beliebigen Anzahl n von Veränderlichen

$$\alpha_1, \alpha_2 \ldots \alpha_n.$$

Eine solche Funktion heißt symmetrisch, wenn sie ungeändert bleibt, wenn die Variablen $\alpha_1, \alpha_2 \ldots \alpha_n$ einer beliebigen Permutation unterworfen werden, und solche symmetrische Funktionen sind es, deren Eigenschaften und Bildungsgesetze wir jetzt genauer kennen lernen müssen.

Damit eine Funktion $\Phi(\alpha_1, \alpha_2, \ldots \alpha_n)$ symmetrisch sei, ist es genügend, daß sie sich bei der Vertauschung von je zweien der Argumente $\alpha_1, \alpha_2, \ldots \alpha_n$ nicht ändere. Denn um aus der Anordnung von n Ziffern 1, 2, 3 ... n eine beliebige andere Anordnung derselben Ziffern $a_1, a_2, a_3 \ldots a_n$ zu erhalten, vertauscht man zunächst die Ziffern 1 und a_1, und hat dann die Aufgabe auf die Bildung einer Permutation von nur $n - 1$ Ziffern zurückgeführt. Die Gesamtzahl aller möglichen Permutationen ist:

$$1.\ 2.\ 3 \ldots n = \Pi(n).$$

Die Funktion Φ wird im allgemeinen nicht homogen sein, sondern Glieder verschiedener Dimension enthalten; wenn man aber alle Glieder gleicher Dimension zusammenfaßt, so läßt sich jedes Φ durch eine Summe homogener Funktionen verschiedener Grade darstellen, und wenn Φ symmetrisch sein soll, so muß jeder homogene Bestandteil eines bestimmten Grades für sich symmetrisch sein, da durch die Permutationen der Variablen die Dimensionen der Glieder nicht geändert werden.

Wir können uns hiernach auf die Betrachtung homogener, symmetrischer Funktionen beschränken, aus denen sich alle anderen zusammensetzen lassen.

Die allgemeine Form einer symmetrischen Funktion erhalten wir, wenn wir in einem Gliede einer solchen Funktion

$$\alpha_1^{\nu_1}\,\alpha_2^{\nu_2}\ldots\,\alpha_n^{\nu_n}$$

die unteren Indices auf alle mögliche Art permutieren und die Summe aller so gebildeten Glieder nehmen. Eine solche Funktion können wir einen Elementarbestandteil einer symmetrischen Funktion nennen. Nehmen wir mehrere solcher Elementarbestandteile, multiplizieren sie mit beliebigen, von den α unabhängigen Faktoren und addieren sie, so erhalten wir die allgemeinste symmetrische Funktion. Die Anzahl der Glieder eines dieser Elementarbestandteile ist, wenn die Exponenten $\nu_1, \nu_2 \ldots \nu_n$ alle voneinander verschieden sind, $\Pi(n)$; wenn aber ein und derselbe Exponent ν mehrmals vorkommt, so hat man die Permutationen, die keine verschiedenen Glieder geben, wegzulassen. Es ist z. B. bei drei Veränderlichen, wenn ν_1, ν_2, ν_3 verschieden sind, ein Elementarbestandteil:

$$\begin{aligned}\alpha_1^{\nu_1}\alpha_2^{\nu_2}\alpha_3^{\nu_3} + \alpha_1^{\nu_1}\alpha_3^{\nu_2}\alpha_2^{\nu_3} + \alpha_2^{\nu_1}\alpha_1^{\nu_2}\alpha_3^{\nu_3} + \alpha_2^{\nu_1}\alpha_3^{\nu_2}\alpha_1^{\nu_3}\\ + \alpha_3^{\nu_1}\alpha_1^{\nu_2}\alpha_2^{\nu_3} + \alpha_3^{\nu_1}\alpha_2^{\nu_2}\alpha_1^{\nu_3},\end{aligned}$$

wenn aber $\nu_3 = \nu_2$ ist:

$$\alpha_1^{\nu_1}\alpha_2^{\nu_3}\alpha_3^{\nu_2} + \alpha_2^{\nu_1}\alpha_1^{\nu_2}\alpha_3^{\nu_2} + \alpha_3^{\nu_1}\alpha_1^{\nu_2}\alpha_2^{\nu_2}.$$

Das einfachste Beispiel einer symmetrischen Funktion ist die Summe der Variablen

$$\alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_n.$$

Ebenso gehört das Produkt $\alpha_1 . \alpha_2 \ldots \alpha_n$ dazu.

Diese beiden sind die extremen Fälle einer Reihe von symmetrischen Funktionen, die wir die symmetrischen Grundfunktionen nennen und folgendermaßen erhalten:

Das Produkt

$$(1) \qquad f(x) = (x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n)$$

ist, wenn x eine von den α unabhängige Variable ist, eine symmetrische Funktion von $\alpha_1, \alpha_2 \ldots \alpha_n$. Wenn wir also die Multiplikation der einzelnen Faktoren ausführen und nach Potenzen von x ordnen, so sind die Koeffizienten der einzelnen Potenzen von x gleichfalls symmetrische Funktionen. Denn sie ändern

sich bei der Vertauschung der α ebensowenig, wie die Funktion $f(x)$. Wir setzen:

$$(2) \quad f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n$$

und erhalten:

$$(3) \quad \begin{aligned} a_1 &= -\Sigma \alpha_1 \\ a_2 &= +\Sigma \alpha_1 \alpha_2 \\ a_3 &= -\Sigma \alpha_1 \alpha_2 \alpha_3 \\ &\cdots\cdots\cdots\cdots \\ a_n &= \pm \Sigma \alpha_1 \alpha_2 \ldots a_n. \end{aligned}$$

Diese Summen sind die symmetrischen Grundfunktionen.

Das Ziel unserer Betrachtungen ist der Beweis des Hauptsatzes:

I. Alle symmetrischen Funktionen der α lassen sich rational durch die Grundfunktionen ausdrücken.

§ 23.

Die Potenzsummen.

Wir beschäftigen uns zunächst mit einer anderen speziellen Art symmetrischer Funktionen, den Potenzsummen. Bedeutet nämlich ν irgend einen ganzzahligen positiven Exponenten, so gehört

$$(1) \quad s_\nu = \alpha_1^\nu + \alpha_2^\nu + \cdots + \alpha_n^\nu$$

offenbar zu den symmetrischen Funktionen, und s_ν wird die νte Potenzsumme genannt. Wir wollen Formeln ableiten, nach denen die Potenzsummen durch die Grundfunktionen ausdrückbar sind.

Wir bezeichnen in der Folge immer, wenn $\varphi(x)$ irgend eine Funktion von x ist, mit

$$S[\varphi(\alpha)]$$

die Summe, die wir erhalten, wenn x in $\varphi(x)$ durch jede der Variablen $\alpha_1, \alpha_2 \ldots \alpha_n$ ersetzt und die so gebildeten Funktionen addiert werden, also

$$S[\varphi(\alpha)] = \varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_n).$$

Hiernach ist z. B.

$$S[\alpha^\nu] = s_\nu$$

die νte Potenzsumme.

Wir dividieren nun $f(x)$ durch eine beliebige lineare Funktion $x - \alpha$ und erhalten:

$$(2)\quad \frac{f(x)}{x-\alpha} = x^{n-1} + f_1(\alpha)x^{n-2} + f_2(\alpha)x^{n-3} + \cdots + f_{n-1}(\alpha) + \frac{f(\alpha)}{x-\alpha},$$

worin nach § 16, (10)

$$(3)\quad \begin{aligned} f_1(\alpha) &= \alpha + a_1 \\ f_2(\alpha) &= \alpha^2 + a_1\alpha + a_2 \\ f_3(\alpha) &= \alpha^3 + a_1\alpha^2 + a_2\alpha + a_3 \\ &\cdots\cdots\cdots\cdots \\ f_{n-1}(\alpha) &= \alpha^{n-1} + a_1\alpha^{n-2} + a_2\alpha^{n-3} + \cdots + a_{n-1}. \end{aligned}$$

Wir setzen in (2) für α jede der Größen $\alpha_1, \alpha_2 \ldots \alpha_n$, wodurch $f(\alpha) = 0$ wird, bilden die Summe S und erhalten:

$$(4)\quad S\left[\frac{f(x)}{x-\alpha}\right] = f'(x)$$

$$= n x^{n-1} + (n-1)a_1 x^{n-2} + (n-2)a_2 x^{n-3} + \cdots + a_{n-1},$$

während die rechte Seite

$$(5)\quad n x^{n-1} + x^{n-2}S[f_1(\alpha)] + x^{n-3}S[f_2(\alpha)] + \cdots + S[f_{n-1}(\alpha)]$$

wird, und die Vergleichung der Koeffizienten gleicher Potenzen von x in (4) und (5) ergibt:

$$(6)\quad \begin{aligned} S[f_1(\alpha)] &= (n-1)a_1 \\ S[f_2(\alpha)] &= (n-2)a_2 \\ &\cdots\cdots\cdots \\ S[f_{n-1}(\alpha)] &= a_{n-1}. \end{aligned}$$

Es ist aber nach (1) und (3)

$$\begin{aligned} S[f_1(\alpha)] &= s_1 + n a_1 \\ S[f_2(\alpha)] &= s_2 + a_1 s_1 + n a_2 \\ &\cdots\cdots\cdots\cdots \\ S[f_{n-1}(\alpha)] &= s_{n-1} + a_1 s_{n-2} + a_2 s_{n-3} + \cdots + n a_{n-1}, \end{aligned}$$

und demnach erhält man aus (6) das folgende von Newton herrührende Formelsystem[1]):

$$(7)\quad \begin{aligned} 0 &= s_1 + a_1 \\ 0 &= s_2 + a_1 s_1 + 2a_2 \\ 0 &= s_3 + a_1 s_2 + a_2 s_1 + 3a_3 \\ &\cdots\cdots\cdots\cdots \\ 0 &= s_{n-1} + a_1 s_{n-2} + a_2 s_{n-3} + \cdots + (n-1)a_{n-1}. \end{aligned}$$

[1]) Newton, Arithmetica universalis, edit. s'Gravesande, p. 592.

Dieses Formelsystem läßt sich aber noch weiter fortsetzen; denn da

$$S[f(\alpha)] = 0, \quad S[\alpha f(\alpha)] = 0, \quad S[\alpha^2 f(\alpha)] = 0 \ldots$$

ist, so folgt:

$$\begin{aligned} 0 &= s_n + a_1 s_{n-1} + a_2 s_{n-2} + \cdots + n a_n \\ 0 &= s_{n+1} + a_1 s_n + a_2 s_{n-1} + \cdots + a_n s_1 \\ 0 &= s_{n+2} + a_1 s_{n+1} + a_2 s_n + \cdots + a_n s_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{8}$$

Durch die Formeln (7), (8) ist nun die Aufgabe gelöst, der Reihe nach die Funktionen s_1, s_2, s_3 ... bis zu beliebiger Höhe als ganze rationale Funktionen der symmetrischen Grundfunktionen darzustellen, z. B.:

$$\begin{aligned} s_1 &= -a_1 \\ s_2 &= +a_1^2 - 2a_2 \\ s_3 &= -a_1^3 + 3a_1a_2 - 3a_3 \\ s_4 &= +a_1^4 - 4a_1^2a_2 + 4a_1a_3 + 2a_2^2 - 4a_4 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{9}$$

und die Bildungsweise dieser Ausdrücke zeigt, daß die Koeffizienten in diesen Darstellungen ganze Zahlen sind.

Man kann auch umgekehrt mittels der Formeln (7) und (8) die symmetrischen Grundfunktionen a_1, a_2, a_3 ... rational durch die Potenzsummen s_1, s_2, s_3 ... ausdrücken. Diese Darstellung ist aber insofern weit weniger einfach, als die Koeffizienten nicht ganze, sondern gebrochene Zahlen sind, z. B.:

$$\begin{aligned} a_1 &= -s_1 \\ 2a_2 &= +s_1^2 - s_2 \\ 6a_3 &= -s_1^3 + 3s_1s_2 - 2s_3 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{10}$$

Man kann das Formelsystem (8) auch nach der entgegengesetzten Richtung fortsetzen, wenn man die Gleichungen

$$S[\alpha^{-1} f(\alpha)] = 0, \qquad S[\alpha^{-2} f(\alpha)] = 0 \ldots$$

bildet. Man erhält dadurch ein Mittel, um die Potenzsummen s_ν auch für negative Exponenten ν durch die Grundfunktionen auszudrücken. Diese Summen der negativen Potenzen gehören gleichfalls zu den symmetrischen Funktionen, wenn auch nicht mehr zu den ganzen, sondern zu den gebrochenen. Sie gehen erst durch Multiplikation mit Potenzen des Produktes $\alpha_1 \alpha_2 \ldots \alpha_n$ in

ganze Funktionen über. Der Vollständigkeit wegen setzen wir die zwei ersten dieser Formeln hierher:

$$(11)\quad \begin{aligned} 0 &= s_{n-1} + a_1 s_{n-2} + a_2 s_{n-3} + \cdots + n a_{n-1} + a_n s_{-1} \\ 0 &= s_{n-2} + a_1 s_{n-3} + a_2 s_{n-4} + \cdots + a_{n-1} s_{-1} + a_n s_{-2}, \end{aligned}$$

die sich nach (7) auch so darstellen lassen:

$$a_{n-1} + a_n s_{-1} = 0, \quad 2 a_{n-2} + a_{n-1} s_{-1} + a_n s_{-2} = 0.$$

Aus § 17, (19) erhält man noch:

$$(12)\quad \begin{aligned} S\left[\frac{f_\nu(\alpha)}{f'(\alpha)}\right] &= 0, \qquad \nu = 0, 1, 2 \ldots n-2, \\ S\left[\frac{f_{n-1}(\alpha)}{f'(\alpha)}\right] &= 1. \end{aligned}$$

§ 24.

Beweis des Hauptsatzes.

Wir gehen nunmehr zum Beweis des Fundamentalsatzes der Theorie der symmetrischen Funktionen über, daß sie alle rational durch die symmetrischen Grundfunktionen ausdrückbar sind. Da wir die vollständige Induktion als Beweismittel anwenden, so leiten wir den Satz zunächst unter der Voraussetzung ab, daß nur zwei unabhängige Veränderliche α, β gegeben seien, aber auf einem Wege, der zugleich für den allgemeinen Beweis den leitenden Gedanken hervortreten lassen wird.

Wir bezeichnen die symmetrischen Grundfunktionen mit

$$(1)\quad a = -(\alpha + \beta), \quad b = \alpha\beta,$$

und setzen demgemäß

$$(2)\quad f(x) = (x - \alpha)(x - \beta) = x^2 + ax + b.$$

Es sei nun $S(\alpha, \beta)$ irgend eine ganze rationale und symmetrische Funktion von α und β. Wir können für β aus (1) den Wert $-(a + \alpha)$ einsetzen und erhalten, wenn wir nach Potenzen von α ordnen,

$$(3)\quad \begin{aligned} S(\alpha, \beta) &= S(\alpha, -a - \alpha) \\ &= A_0 \alpha^m + A_1 \alpha^{m-1} + \cdots + A_{m-1}\alpha + A_m, \end{aligned}$$

worin die Koeffizienten A_0, $A_1 \ldots A_m$ nur von a und von den in S etwa noch vorkommenden Koeffizienten abhängen. Wir bemerken aber ausdrücklich, daß, wenn in $S(\alpha, \beta)$ keine gebrochenen Zahlenkoeffizienten vorkommen, auch in den Koeffizienten $A_0, A_1 \ldots A_m$ keine Brüche auftreten.

Wir setzen nun

$$\Phi(x) = A_0 x^m + A_1 x^{m-1} + \cdots + A_{m-1} x + A_m, \tag{4}$$

dividieren $\Phi(x)$ durch $f(x)$ und erhalten einen Quotienten Q und einen Rest, der in bezug auf x höchstens vom ersten Grade ist, also:

$$\Phi(x) = Qf(x) + A + Bx; \tag{5}$$

hierin sind nun A und B ganze Funktionen von a und b, und auch sie enthalten keinerlei gebrochene Zahlenkoeffizienten, wenn in S keine solche vorkommen. Wenn wir nun $x = \alpha$ setzen, so ergibt sich aus (5) und (3), da $f(\alpha)$ verschwindet,

$$S(\alpha, \beta) = A + B\alpha. \tag{6}$$

Da aber $S(\alpha, \beta)$ und ebenso A, B symmetrisch sind, so folgt durch Vertauschung von α und β

$$S(\alpha, \beta) = A + B\beta. \tag{7}$$

Hieraus schließt man, da α und β voneinander unabhängige Variable sind, daß $B = 0$ und folglich

$$S(\alpha, \beta) = A \tag{8}$$

sein muß, womit der Fundamentalsatz für diesen Fall bewiesen ist.

Wir setzen nun voraus, der Fundamentalsatz sei bewiesen für symmetrische Funktionen von $n - 1$ Veränderlichen und leiten ihn durch ein Verfahren, das dem eben für zwei Variable angewandten ganz analog ist, für n Variable her.

Es sei wieder

$$S = S(\alpha_1, \alpha_2 \ldots \alpha_n) \tag{9}$$

eine ganze symmetrische Funktion der n Veränderlichen $\alpha_1, \alpha_2 \ldots \alpha_n$.

Wenn wir sie nach Potenzen von α_1 ordnen und demgemäß setzen

$$S = S_0 \alpha_1^{\mu} + S_1 \alpha_1^{\mu-1} + \cdots + S_{\mu-1}\alpha_1 + S_\mu, \tag{10}$$

so sind die Koeffizienten S_0, $S_1 \ldots S_\mu$ ganze symmetrische Funktionen der $n - 1$ Veränderlichen $\alpha_2, \ldots \alpha_n$.

Bezeichnen wir die symmetrischen Grundfunktionen dieser letzteren Variablen mit $a_1', a_2' \ldots a_{n-1}'$, so können wir die Koeffizienten S_0, $S_1 \ldots S_\mu$ nach unserer Voraussetzung rational durch diese ausdrücken. Es ist aber nach § 23, (3):

$$\begin{aligned} a_1' &= f_1(\alpha_1) = \alpha_1 + a_1 \\ a_2' &= f_2(\alpha_1) = \alpha_1^2 + a_1\alpha_1 + a_2 \\ a_3' &= f_3(\alpha_1) = \alpha_1^3 + a_1\alpha_1^2 + a_2\alpha_1 + a_3 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \end{aligned}$$

d. h. die Koeffizienten S_0, $S_1 \dots S_\mu$ können ganz und rational durch α_1, a_1, $a_2 \dots a_n$ ausgedrückt werden. Wenn wir also, nachdem diese Ausdrücke eingeführt sind, in (10) aufs neue nach Potenzen von α_1 ordnen, so erhalten wir:

$$S = A_0 \alpha_1^m + A_1 \alpha_1^{m-1} + \cdots + A_{m-1}\alpha_1 + A_m, \tag{11}$$

worin im allgemeinen m ein von μ verschiedener Exponent sein wird, und die Koeffizienten A_0, $A_1 \dots A_m$ ganz und rational von a_1, $a_2 \dots a_n$ abhängen. Wir setzen wieder

$$\Phi(x) = A_0 x^m + A_1 x^{m-1} + \cdots + A_{m-1} x + A_m \tag{12}$$

und dividieren $\Phi(x)$ durch

$$\begin{aligned} f(x) &= x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n \\ &= (x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n). \end{aligned}$$

Es ergibt sich ein Quotient und ein Rest, der in bezug auf x höchstens vom Grade $n - 1$ ist. Wir setzen also:

$$\Phi(x) = Q f(x) + \psi(x) \tag{13}$$

$$\psi(x) = C_0 x^{n-1} + C_1 x^{n-2} + \cdots + C_{n-2} x + C_{n-1}, \tag{14}$$

und hierin sind C_0, $C_1 \dots C_{n-1}$ ganze rationale Funktionen von a_1, $a_2 \dots a_n$, in denen, wenn S in seiner ursprünglichen Gestalt keine gebrochenen Koeffizienten enthält, auch keine Brüche vorkommen.

Nun ist aber, da $f(\alpha_1)$ verschwindet und $S = \Phi(\alpha_1)$ ist, nach (13)

$$S = \psi(\alpha_1), \tag{15}$$

und hierin kann, da S symmetrisch ist, α_1 durch α_2, $\alpha_3 \dots \alpha_n$ ersetzt werden. Hieraus ergeben sich die folgenden n Gleichungen:

$$\begin{aligned} &C_0 \alpha_1^{n-1} + C_1 \alpha_1^{n-2} + \cdots + C_{n-2}\alpha_1 + (C_{n-1} - S) = 0 \\ &C_0 \alpha_2^{n-1} + C_1 \alpha_2^{n-2} + \cdots + C_{n-2}\alpha_2 + (C_{n-1} - S) = 0 \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &C_0 \alpha_n^{n-1} + C_1 \alpha_n^{n-2} + \cdots + C_{n-2}\alpha_n + (C_{n-1} - S) = 0. \end{aligned} \tag{16}$$

Betrachten wir dies System als ein System homogener linearer Gleichungen mit den n Unbekannten

$$C_0, C_1 \dots C_{n-2}, C_{n-1} - S,$$

so ist seine Determinante

$$\begin{vmatrix} \alpha_1^{n-1}, \alpha_1^{n-2} \dots \alpha_1, 1 \\ \alpha_2^{n-1}, \alpha_2^{n-2} \dots \alpha_2, 1 \\ \dots\dots\dots\dots\dots \\ \alpha_n^{n-1}, \alpha_n^{n-2} \dots \alpha_n, 1 \end{vmatrix},$$

die nach § 3, (9) gleich dem Produkt aller Differenzen $\alpha_1 - \alpha_2$, $\alpha_2 - \alpha_3 \ldots$ ist, von Null verschieden, und folglich ist

$$C_0 = 0, \quad C_1 = 0 \ldots \quad C_{n-2} = 0$$

und

$$S = C_{n-1}, \tag{17}$$

worin der zu beweisende Fundamentalsatz enthalten ist.

Zu demselben Resultat gelangt man auch daraus, daß die Funktion $(n-1)$ten Grades von x

$$C_0 x^{n-1} + C_1 x^{n-2} + \cdots + C_{n-2} x + C_{n-1} - S$$

für $x = \alpha_1, \alpha_2 \ldots \alpha_n$, also für mehr als $n - 1$ Werte verschwindet, und folglich alle ihre Koeffizienten gleich Null haben muß (§ 4).

Der Beweis, den wir hier für das Fundamentaltheorem im Anschluß an Cauchy[1]) gegeben haben, bietet zugleich ein Mittel, in besonderen Fällen den Ausdruck einer symmetrischen Funktion durch die Grundfunktionen wirklich zu berechnen. Dieselbe Möglichkeit bieten auch die anderen Beweise, die für das Theorem bekannt sind. Wir wollen noch einen zweiten Beweis hier mitteilen, der zu einer oft einfacheren Berechnungsart führt[2]).

§ 25.

Zweiter Beweis des Satzes von den symmetrischen Funktionen.

Es sei S eine ganze symmetrische Funktion der Variablen $\alpha_1, \alpha_2 \ldots \alpha_n$. Die einzelnen Glieder dieser Funktion sind alle von der Form

$$M \alpha_1^{\nu_1} \alpha_2^{\nu_2} \ldots \alpha_n^{\nu_n},$$

worin M ein von den α unabhängiger Koeffizient ist. Diese Glieder sollen nun nach der im § 20 gegebenen Vorschrift geordnet werden, d. h. es soll, nachdem die Reihenfolge der Variablen $\alpha_1, \alpha_2 \ldots \alpha_n$ festgesetzt ist, von zwei Gliedern

$$A = M \alpha_1^{\nu_1} \alpha_2^{\nu_2} \ldots \alpha_n^{\nu_n}, \qquad A' = M' \alpha_1^{\nu'_1} \alpha_2^{\nu'_2} \ldots \alpha_n^{\nu'_n}$$

---

[1]) Cauchy, Exercices de mathématiques, 4eme année.

[2]) Waring, Meditationes algebraicae. Gauß, Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. Werke, Bd. III, S. 36. Deutsch von Netto in Ostwalds Klassikern, Nr. 14.

A das höhere genannt werden, wenn die erste der Differenzen

$$\nu_1 - \nu_1', \quad \nu_2 - \nu_2' \ldots \nu_n - \nu_n',$$

die von Null verschieden ist, einen positiven Wert hat, wenn also entweder $\nu_1 > \nu_1'$ oder $\nu_1 = \nu_1', \nu_2 > \nu_2'$ oder $\nu_1 = \nu_1', \nu_2 = \nu_2', \nu_3 > \nu_3'$ usw.

Da wir alle Glieder, in denen sämtliche Exponenten $\nu_1, \nu_2 \ldots \nu_n$ übereinstimmen, in ein Glied vereinigt voraussetzen, so ist hiernach von je zwei Gliedern entschieden, welches das höhere ist, und wenn A höher als A', A' höher als A'' ist, so ist auch A höher als A''.

Ist nun nach dieser Anordnung

$$A = M\alpha_1^{\nu_1}\alpha_2^{\nu_2} \ldots \alpha_n^{\nu_n}$$

das höchste Glied unserer Funktion S, so folgt, daß

$$\nu_1 \geqq \nu_2$$

sein muß; denn wäre $\nu_1 < \nu_2$, so würde das Glied

$$M\alpha_1^{\nu_2}\alpha_2^{\nu_1} \ldots \alpha_n^{\nu_n},$$

das wegen der Symmetrie gleichfalls in S vorkommen muß, in der Ordnung höher stehen als A, gegen die Voraussetzung. Ebenso folgt, daß $\nu_2 \geqq \nu_3$ sein muß, denn wäre $\nu_2 < \nu_3$, so würde

$$M\alpha_1^{\nu_1}\alpha_2^{\nu_3}\alpha_3^{\nu_2} \ldots \alpha_n^{\nu_n}$$

höher stehen als A usf. Wir schließen also, daß die Exponenten in A

$$\nu_1, \nu_2, \nu_3 \ldots \nu_n$$

eine abnehmende oder wenigstens niemals wachsende Zahlenreihe bilden, oder daß die Differenzen

$$\nu_1 - \nu_2, \quad \nu_2 - \nu_3, \quad \ldots \nu_{n-1} - \nu_n$$

alle positiv oder wenigstens nicht negativ sind.

Wir schließen zweitens, daß in keinem Gliede der Funktion S ein höherer Exponent als ν_1 vorkommen kann; denn sonst würde auch ein Glied vorkommen, in dem α_1 diesen höheren Exponenten hätte, und dies wäre gegen die Voraussetzung von höherer Ordnung als A. Es sind also die Glieder, die in einer symmetrischen Funktion überhaupt vorkommen können, von den Koeffizienten abgesehen, durch das höchste Glied vollkommen bestimmt, und die Anzahl der möglichen Glieder ist bei gegebenem höchsten Gliede nur eine endliche.

Wir ordnen die symmetrischen Funktionen nach der Exponentenreihe ihres höchsten Gliedes

$$(\nu_1, \nu_2 \dots \nu_n)$$

in der Weise an, daß von zwei solchen Funktionen S, S' mit den Exponentenreihen $(\nu_1, \nu_2 \dots \nu_n)$ und $(\nu'_1, \nu'_2 \dots \nu'_n)$ S als die höhere gilt, wenn die erste nicht verschwindende unter den Differenzen $\nu_1 - \nu'_1, \nu_2 - \nu'_2 \dots \nu_n - \nu'_n$ positiv ist.

Die niedrigste Ordnung ist hiernach die, in der alle diese Exponenten $= 0$ sind, und die symmetrische Funktion also eine Konstante wird. Die nächst niedrige Ordnung ist

$$(1, 0, 0 \dots 0),$$

der die einzige symmetrische Funktion

$$M(\alpha_1 + \alpha_2 + \cdots + \alpha_n) = - Ma_1$$

entspricht. Die nächstfolgende Ordnung ist

$$(1, 1, 0 \dots 0),$$

der die symmetrische Funktion

$$Ma_1 + M'a_2$$

entspricht, und so erhalten wir in den n ersten Ordnungen die Grundfunktionen selbst und ihre linearen Verbindungen, und keine anderen.

Die nächstfolgende Ordnung ist charakterisiert durch

$$(2, 0, 0 \dots 0)$$

und enthält die linearen Verbindungen der Grundfunktionen mit der Summe der Quadrate.

In allen diesen Fällen ist die Darstellbarkeit der symmetrischen Funktionen bereits bewiesen, und wir werden also jetzt den allgemeinen Beweis dadurch führen, daß wir die Darstellung der Funktion S von der Ordnung $(\nu_1, \nu_2 \dots \nu_n)$ durch die Grundfunktionen unter der Voraussetzung ableiten, daß sie für die Funktionen niedrigerer Ordnung schon bekannt sei, also durch das Schlußverfahren der vollständigen Induktion.

Wir bemerken hierzu, daß durch die Multiplikation zweier symmetrischer Funktionen S und S', deren höchste Glieder

$$A = M\alpha_1^{\nu_1}\alpha_2^{\nu_2} \dots \alpha_n^{\nu_n}, \qquad A' = M'\alpha_1^{\nu'_1}\alpha_2^{\nu'_2} \dots \alpha_n^{\nu'_n}$$

sind, eine symmetrische Funktion entsteht, deren höchstes Glied das Produkt

$$AA' = MM'\alpha_1^{\nu_1+\nu'_1}\alpha_2^{\nu_2+\nu'_2} \dots \alpha_n^{\nu_n+\nu'_n}$$

ist. Denn nehmen wir an, es gebe in SS' ein höheres Glied als AA', das aus der Multiplikation der beiden Glieder

$$B = N\alpha_1^{\mu_1}\alpha_2^{\mu_2} \dots \alpha_n^{\mu_n}, \qquad B' = N'\alpha_1^{\mu'_1}\alpha_2^{\mu'_2} \dots \alpha_n^{\mu'_n}$$

entsteht, also

$$BB' = NN'\alpha_1^{\mu_1+\mu'_1}\alpha_2^{\mu_2+\mu'_2} \dots \alpha_n^{\mu_n+\mu'_n},$$

und es sei nicht gleichzeitig $B = A$ und $B' = A'$, so wäre die erste der Differenzen

$$\mu_1 - \nu_1 + \mu'_1 - \nu'_1,\ \mu_2 - \nu_2 + \mu'_2 - \nu'_2 \dots \mu_n - \nu_n + \mu'_n - \nu'_n,$$

die von Null verschieden ist, positiv, was unmöglich ist, da die erste nicht verschwindende unter den Differenzen

$$\mu_1 - \nu_1,\ \mu_2 - \nu_2 \dots \mu_n - \nu_n$$

und unter den Differenzen

$$\mu'_1 - \nu'_1,\ \mu'_2 - \nu'_2 \dots \mu'_n - \nu'_n$$

nach der Voraussetzung negativ ist. Durch Wiederholung dieses Schlusses ergibt sich der allgemeinere Satz, daß man das höchste Glied eines Produktes aus mehreren Faktoren dadurch erhält, daß man die höchsten Glieder der einzelnen Faktoren multipliziert.

Die höchsten Glieder der symmetrischen Grundfunktionen

$$a_1,\ a_2,\ a_3 \dots a_n$$

sind nun, abgesehen vom Vorzeichen,

$$\alpha_1,\ \alpha_1\alpha_2,\ \alpha_1\alpha_2\alpha_3, \dots\ \alpha_1\alpha_2\alpha_3 \dots \alpha_n,$$

und wenn wir also das Produkt bilden

$$(1) \qquad P = \pm M a_1^{\nu_1-\nu_2} a_2^{\nu_2-\nu_3} \dots a_{n-1}^{\nu_{n-1}-\nu_n} a_n^{\nu_n},$$

so erhalten wir eine symmetrische Funktion, deren höchstes Glied gleichfalls A ist, wie in S.

Die Differenz

$$S - P = S'$$

ist also wieder eine symmetrische Funktion, deren höchstes Glied niedriger ist als das von S, und die wir nach der Voraussetzung rational durch die Grundfunktionen darstellen können. Da P ebenso dargestellt ist, ist das Ziel erreicht.

Nennen wir die α_i die Wurzeln, die a_i die Koeffizienten der Funktion

$$(2) \qquad \begin{aligned} f(x) &= (x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n) \\ &= x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n, \end{aligned}$$

so können wir unser Resultat auch so aussprechen:

II. Die symmetrischen Funktionen der Wurzeln lassen sich rational durch die Koeffizienten ausdrücken Hat die symmetrische Funktion der α_i ganzzahlige Koeffizienten, so hat auch die ganze Funktion der a_i ganzzahlige Koeffizienten.

Betrachten wir jetzt eine ganze Funktion $F(a)$ der Variablen $a_1, a_2 \dots a_n$ und bezeichnen eines ihrer Glieder mit

$$P = M a_1^{h_1} a_2^{h_2} \dots a_n^{h_n}. \tag{3}$$

Ersetzen wir darin die Variablen a durch die symmetrischen Grundfunktionen der α, so erhalten wir daraus eine symmetrische Funktion $\Phi(a)$ der α, deren höchstes Glied nach (1)

$$A = M \alpha_1^{\nu_1} \alpha_2^{\nu_2} \dots \alpha_n^{\nu_n} \tag{4}$$

wird, wenn

$$\begin{aligned} \nu_1 &= h_1 + h_2 + \dots + h_n, \\ \nu_2 &= h_2 + \dots + h_n, \\ &\dots\dots\dots\dots \\ \nu_n &= h_n \end{aligned} \tag{5}$$

gesetzt wird; denn daraus folgt:

$$h_1 = \nu_1 - \nu_2, \quad h_2 = \nu_2 - \nu_3, \quad \dots \; h_n = \nu_n.$$

Wir ordnen nun die Glieder P der Funktion $F(a)$ in der Weise, daß wir einem Glied P einen höheren Rang geben als einem Glied P', wenn das durch die Substitution der α entstandene Glied A einen höheren Rang hat, als das entsprechende aus P' entspringende Glied A'. Dadurch werden die Glieder P der Funktion $F(a)$ in eindeutiger Weise in eine Rangordnung gebracht und das höchste Glied P von $F(a)$ hat denselben Koeffizienten M, wie das höchste Glied A von $\Phi(\alpha)$.

Daraus folgt:

III. Eine nicht identisch verschwindende Funktion der a kann nicht dadurch identisch verschwinden, daß man für die a die symmetrischen Grundfunktionen der α setzt.

Diesem Satz können wir die beiden gleichwertigen Ausdrücke geben:

IV. Eine identische Gleichung zwischen den symmetrischen Grundfunktionen der α bleibt richtig, wenn man für die Grundfunktionen unabhängige Variable setzt.

V. Eine symmetrische Funktion der Variablen α läßt sich nur auf eine Weise durch die Grundfunktionen ausdrücken.

Denn wären $F(a)$, $F_1(a)$ zwei Darstellungen von einer Funktion $\Phi(\alpha)$, so würde die nicht identisch verschwindende Funktion $F(a) - F_1(a)$ durch Substitution der symmetrischen Grundfunktionen der α identisch verschwinden.

Ersetzt man $a_1, a_2 \dots a_n$ durch

$$\frac{a_1}{a_0}, \frac{a_2}{a_0} \dots \frac{a_n}{a_0},$$

so ist nach (1) $a_0^{\nu_1} F$ eine ganze homogene Funktion $\nu_1$ten Grades der $n + 1$ Variablen $a_0, a_1 \dots a_n$.

Die Funktion $a_0^{\nu_1} F$ ist nicht durch a_0 teilbar, denn aus (1) erhält man das erste Glied von $a_0^{\nu_1} F$ in der Form

$$P = \pm M a_1^{\nu_1 - \nu_2} a_2^{\nu_2 - \nu_3} \dots a_n^{\nu_n},$$

das nicht durch a_0 teilbar ist und auch nicht durch eines der folgenden Glieder zerstört werden kann, da die Exponenten $\nu_1 - \nu_2$, $\nu_2 - \nu_3 \dots \nu_n$ nicht in zwei Gliedern gleich sein können.

Demnach ist $a_0^{\nu_1}$ die niedrigste Potenz von a_0, mit der die symmetrische Funktion F mit dem höchsten Glied $(\nu_1, \nu_2 \dots \nu_n)$ multipliziert werden muß, um sie in eine ganze Funktion von $a_0, a_1 \dots a_n$ zu verwandeln.

§ 26.

Diskriminanten.

Wir wenden die bewiesenen Sätze an auf das Differenzenprodukt

$$P = (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \dots (\alpha_1 - \alpha_n) \quad (1)$$
$$(\alpha_2 - \alpha_3) \dots (\alpha_2 - \alpha_n)$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$(\alpha_{n-1} - \alpha_n).$$

Dieses ändert durch Umstellung zweier Elemente nur das Vorzeichen. Das Quadrat wird also bei allen Vertauschungen zweier der Variablen α und demnach auch bei allen Permutationen ungeändert bleiben, d. h. eine symmetrische Funktion der Variablen sein. Die Ordnung dieser symmetrischen Funktion ist, wie man aus (1) sieht, gleich

$$(2n - 2,\ 2n - 4,\ \dots\ 2,\ 0).$$

Wenn wir also die Funktion $f(x)$, deren Wurzeln die $\alpha_1, \alpha_2 \ldots \alpha_n$ sind, in der Form annehmen:

$$f(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n,$$

so daß die höchste Potenz von x nicht gerade den Koeffizienten 1 hat, so ist P^2 eine ganze rationale Funktion von

$$\frac{a_1}{a_0}, \frac{a_2}{a_0} \cdots \frac{a_n}{a_0}, \tag{2}$$

die durch Multiplikation mit a_0^{2n-2} in eine homogene Funktion von $a_0, a_1 \ldots a_n$ übergeht. Wir nennen diese homogene Funktion, die also durch

$$a_0^{2n-2} P^2 = D \tag{3}$$

definiert ist, die Diskriminante der Funktion $f(x)$. Setzen wir $a_0 = 1$, so erhalten wir die unhomogene Form der Diskriminante.

VI. Das Verschwinden der Diskriminante ist die notwendige und hinreichende Bedingung dafür, daß zwei der Werte $\alpha_1, \alpha_2 \ldots \alpha_n$ einander gleich werden.

Wir können für die Diskriminante noch einen zweiten Ausdruck aufstellen.

Es ist nach § 16, (17)

$$\begin{aligned} f'(\alpha_1) &= a_0 (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3) \ldots (\alpha_1 - \alpha_n) \\ f'(\alpha_2) &= a_0 (\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3) \ldots (\alpha_2 - \alpha_n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f'(\alpha_n) &= a_0 (\alpha_n - \alpha_1)(\alpha_n - \alpha_2) \ldots (\alpha_n - \alpha_{n-1}), \end{aligned} \tag{4}$$

folglich, da hier jeder Faktor $(\alpha_h - \alpha_k)$ zweimal mit entgegengesetzten Zeichen vorkommt,

$$D = (-1)^{\frac{n(n-1)}{2}} a_0^{n-2} f'(\alpha_1) f'(\alpha_2) \ldots f'(\alpha_n). \tag{5}$$

Um die Diskriminante wirklich darzustellen, haben wir zunächst ein allgemeines Mittel. Es war nach § 3, (9)

$$(-1)^{\frac{n(n-1)}{2}} P = \begin{vmatrix} 1, & \alpha_1, & \alpha_1^2 & \ldots & \alpha_1^{n-1} \\ 1, & \alpha_2, & \alpha_2^2 & \ldots & \alpha_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1, & \alpha_n, & \alpha_n^2 & \ldots & \alpha_n^{n-1} \end{vmatrix}, \tag{6}$$

und wenn wir das Quadrat dieser Determinante durch Multiplikation nach Vertikalreihen als eine neue Determinante darstellen und

$$s_m = \alpha_1^m + \alpha_2^m + \cdots \alpha_n^m$$

setzen:

$$(7) \qquad D = a_0^{2n-2} \begin{vmatrix} s_0, & s_1, & s_2, & \ldots & s_{n-1} \\ s_1, & s_2, & s_3, & \ldots & s_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ s_{n-1}, & s_n, & s_{n+1} & \ldots & s_{2n-2} \end{vmatrix}.$$

Die Größen s_m sind aber die Potenzsummen, die wir im § 23 durch die Koeffizienten ausgedrückt haben; nur sind an Stelle der $a_1, a_2 \ldots a_n$ die Quotienten (2) zu setzen.

So finden wir z. B. für $n = 2$

$$(8) \qquad D = a_0^2(s_0 s_2 - s_1^2) = a_1^2 - 4 a_0 a_2$$

und für $n = 3$

$$(9) \qquad D = a_0^4(s_0 s_2 s_4 - s_0 s_3^2 + 2 s_1 s_2 s_3 - s_1^2 s_4 - s_2^3).$$

Hierin ist nach § 23 zu setzen:

$$\begin{aligned} s_0 &= 3, \quad a_0 s_1 = -a_1 \\ a_0^2 s_2 &= a_1^2 - 2 a_0 a_2 \\ a_0^3 s_3 &= -a_1^3 + 3 a_0 a_1 a_2 - 3 a_0^2 a_3 \\ a_0^4 s_4 &= a_1^4 - 4 a_0 a_1^2 a_2 + 4 a_0^2 a_1 a_3 + 2 a_0^2 a_2^2, \end{aligned}$$

wodurch man erhält:

$$(10) \quad D = a_1^2 a_2^2 + 18 a_0 a_1 a_2 a_3 - 4 a_0 a_2^3 - 4 a_1^3 a_3 - 27 a_0^2 a_3^2.$$

VII. Die Diskriminante D ist eine unzerlegbare Funktion der Variablen $a_0, a_1 \ldots a_n$.

Denn angenommen, D habe einen rationalen Teiler D_1, so muß D_1 homogen sein, und läßt sich daher, von einer Potenz von a_0 als Faktor abgesehen, rational durch die Variablen α_1, $\alpha_2 \ldots \alpha_n$ ausdrücken. Da nun D durch keine Potenz von a_0 teilbar ist, so muß D_1, wenn es nicht konstant ist, die Variablen $\alpha_1, \alpha_2 \ldots \alpha_n$ enthalten und muß daher wegen (3) durch eine der Differenzen $\alpha_i - \alpha_k$ teilbar sein. Da aber andererseits D_1 in bezug auf die $\alpha_1, \alpha_2 \ldots \alpha_n$ symmetrisch ist, so ist es auch durch das Quadrat von $\alpha_i - \alpha_k$ und durch das ganze Produkt P^2 teilbar, also, von einem konstanten Faktor abgesehen, mit D identisch.

Das Verschwinden der Diskriminante D ist die Bedingung dafür, daß zwei unter den Wurzeln $\alpha_1, \alpha_2 \ldots \alpha_n$ der Gleichung $f(x) = 0$ einander gleich werden. Man kann aber diesen Satz

noch verschärfen und so ein genaues Kennzeichen für die Anzahl der verschiedenen Wurzeln von $f(x) = 0$ herleiten.

Wenn, wie bisher, $s_0, s_1, s_2 \ldots$ die Potenzsummen der α bezeichnen, so setzen wir nun für ein beliebiges $\nu \gtreqless n$

$$D_\nu = \begin{vmatrix} s_0, & s_1 \ldots s_{\nu-1} \\ s_1, & s_2 \ldots s_\nu \\ \ldots\ldots\ldots\ldots\ldots \\ s_{\nu-1}, & s_\nu \ldots s_{2\nu-2} \end{vmatrix}, \tag{11}$$

so daß, wenn $a_0 = 1$ angenommen wird, D_n mit der Diskriminante D von $f(x)$ übereinstimmt. Um diese Funktionen D_ν durch die Wurzeln α_i auszudrücken, beweisen wir folgenden Satz:

VIII. Man wähle beliebig ν unter den Größen α_i aus, etwa $\alpha_1, \alpha_2 \ldots \alpha_\nu$, und bezeichne das Differenzenprodukt dieser ν Größen mit

$$P_\nu = \begin{vmatrix} 1, & \alpha_1 \ldots \alpha_1^{\nu-1} \\ 1, & \alpha_2 \ldots \alpha_2^{\nu-1} \\ \ldots\ldots\ldots\ldots\ldots \\ 1, & \alpha_\nu \ldots \alpha_\nu^{\nu-1} \end{vmatrix}. \tag{12}$$

Dann ist

$$D_\nu = \Sigma P_\nu^2, \tag{13}$$

worin sich die Summe Σ auf alle möglichen Arten, die ν Größen $\alpha_1, \alpha_2 \ldots \alpha_\nu$ auszuwählen, bezieht.

Man kann diesen Satz auch so ausdrücken, daß D_ν die Summe der Diskriminanten aller der Gleichungen ist, die irgend ν von den Größen α_i zu Wurzeln haben.

Um ihn zu beweisen, sei ν irgend eine unter n gelegene Zahl. Schreiben wir dann die Determinante D_ν in der folgenden Form:

$$D_\nu = \begin{vmatrix} \Sigma\alpha_i^0\alpha_i^0, & \Sigma\alpha_i^0\alpha_i^1 & \ldots \Sigma\alpha_i^0\alpha_i^{\nu-1} \\ \Sigma\alpha_i^1\alpha_i^0, & \Sigma\alpha_i^1\alpha_i^1 & \ldots \Sigma\alpha_i^1\alpha_i^{\nu-1} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \Sigma\alpha_i^{\nu-1}\alpha_i^0, & \Sigma\alpha_i^{\nu-1}\alpha_i^1 & \ldots \Sigma\alpha_i^{\nu-1}\alpha_i^{\nu-1} \end{vmatrix},$$

so können wir den Satz § 6, XXIII. anwenden, wenn wir dort

$$a_h^{(s)} = b_h^{(s)} = \alpha_s^{h-1}$$

setzen, und dann geht jene Formel unmittelbar in die Formel (13) über.

Daraus erkennen wir zunächst, daß $a_0^{2\nu-2} D_\nu$ eine ganze Funktion der $a_0, a_1 \ldots a_n$ ist, die den Faktor a_0 nicht mehr

hat. Es ergibt sich aber auch daraus die folgende Bedeutung der D_ν für die Erkennung der Wurzelgleichheit:

IX. Die notwendige und hinreichende Bedingung dafür, daß unter den $\alpha_1, \alpha_2 \ldots \alpha_n$ nicht mehr und nicht weniger als ϱ verschiedene vorkommen, ist

$$D_{\varrho+1} = 0,\ D_{\varrho+2} = 0 \ \ldots D_n = 0,$$

D_ϱ von Null verschieden[1]).

Denn ist ν größer als ϱ, so kommen in jedem der Differenzenprodukte P_ν der Formel (13) zwei gleiche α vor, und sie verschwinden also alle. Ist aber $\nu = \varrho$, so enthält die Summe (13) ein nicht verschwindendes Glied, nämlich das Differenzenprodukt der ϱ verschiedenen Größen $\alpha_1, \alpha_2 \ldots \alpha_\varrho$, während alle anderen Glieder der Summe verschwinden.

§ 27.

Resultanten.

Es seien

$$\alpha_1,\ \alpha_2 \ldots \alpha_n$$
$$\beta_1,\ \beta_2 \ldots \beta_m$$

zwei Reihen von n und m unabhängigen Variablen und es werde gesetzt:

$$(1) \qquad \begin{aligned} f(x) &= a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n) \\ &= a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n, \end{aligned}$$

$$(2) \qquad \begin{aligned} \varphi(x) &= b_0(x-\beta_1)(x-\beta_2)\ldots(x-\beta_m) \\ &= b_0 x^m + b_1 x^{m-1} + \cdots + b_m. \end{aligned}$$

Es sind dann die Quotienten

$$(3) \qquad \frac{a_1}{a_0},\ \frac{a_2}{a_0} \ldots \frac{a_n}{a_0}$$

die symmetrischen Grundfunktionen der α_i und ebenso

$$(4) \qquad \frac{b_1}{b_0},\ \frac{b_2}{b_0} \ldots \frac{b_m}{b_0}$$

die symmetrischen Grundfunktionen der β_i.

Das Produkt $\varphi(\alpha_1)\,\varphi(\alpha_2)\ldots\varphi(\alpha_n)$ ist eine symmetrische Funktion der α_i und kann daher rational durch die Verhältnisse (3) ausgedrückt werden, und ist außerdem nach (2) eine ganze

[1]) L. Baur, Mathematische Annalen, Bd. 50.

Funktion der b_k. Die Ordnung dieser symmetrischen Funktion der α ist (§ 25) $(m, m, \ldots m)$, und sie wird daher durch Multiplikation mit a_0^m in eine ganze Funktion der a_k übergeführt.

Wir setzen

$$R = a_0^m \varphi(\alpha_1)\ \varphi(\alpha_2) \ldots \varphi(\alpha_n). \tag{5}$$

Wegen (2) kann R auch so ausgedrückt werden:

$$R = a_0^m b_0^n \overset{i,k}{\Pi}(\alpha_i - \beta_k), \tag{6}$$

wenn das Produkt Π sich auf alle Indices $i = 1, 2 \ldots n$, $k = 1, 2 \ldots m$ bezieht.

Die Funktion R kann als ganze homogene Funktion $m$ten Grades der $a_0, a_1 \ldots a_n$ und als ganze homogene Funktion $n$ten Grades der $b_0, b_1 \ldots b_n$ dargestellt werden und wird die Resultante der beiden Funktionen f und φ genannt und auch mit $R(f, \varphi)$ bezeichnet.

Vertauscht man f mit φ, so ergibt sich aus (6):

$$\begin{aligned} R(f, \varphi) &= (-1)^{mn}\ R(\varphi, f) \\ R(f, \varphi) &= (-1)^{mn}\ b_0^n f(\beta_1) \ldots f(\beta_2) \ldots f(\beta_m). \end{aligned} \tag{7}$$

Aus der Definition der Resultante gehen die folgenden Sätze hervor.

Aus (6) ergibt sich, wenn h, k unabhängige Variable sind:

$$R(hf, k\varphi) = h^m k^n R(f, \varphi). \tag{8}$$

Ferner ist, wenn λ eine Variable bedeutet:

$$f(\beta_i) + \lambda\varphi(\beta_i) = f(\beta_i),$$

woraus man nach (7) erhält:

$$R(f + \lambda\varphi, \varphi) = (-1)^{m\nu} b_0^\nu f(\beta_1)\ f(\beta_2) \ldots f(\beta_m),$$

worin ν der Grad von $f + \lambda\varphi$ ist, der mit n übereinstimmt, wenn $m \gtreqless n$ ist, und $= m$ ist, wenn $m > n$ ist. Nach (7) folgt hieraus:

$$R(f + \lambda\varphi, \varphi) = (-1)^{m(\nu-n)}\ b_0^{\nu-n} R(f, \varphi). \tag{9}$$

Ist im besonderen $n = m$, so folgt:

$$R(f + \lambda\varphi, \varphi) = R(f, \varphi). \tag{10}$$

Die Formel (9) läßt sich mit (8) verbinden und gibt dann ein allgemeineres und, besonders wenn f und φ von gleichem Grade sind, also $m = n$ ist, einfacheres Resultat. Man kann dann (10) anwenden und erhält, wenn $\alpha, \beta, \gamma, \delta$ wieder neue Variable sind:

$$R(\alpha f + \beta\varphi, \gamma f + \delta\varphi) = (\alpha\delta - \beta\gamma)^n R(f, \varphi). \tag{11}$$

Die Resultante ist eine unzerlegbare Funktion der a_i, b_k. Denn hätte sie einen Teiler R_1, so müßte dieser, als Funktion der α, β nach (6) wenigstens durch eine der Differenzen $\alpha_i - \beta_k$ teilbar sein. Da aber R_1 eine symmetrische Funktion ist, so wäre es durch alle diese Differenzen und mithin durch R selbst teilbar.

Für die folgende Betrachtung ist es besser, die Funktionen $f(x)$, $\varphi(x)$ in nicht homogener Form zu betrachten, also a_0 und $b_0 = 1$ zu setzen.

Es seien nun in den beiden Funktionen $n$ten und $m$ten Grades:

$$(12)\qquad \begin{aligned} f(x) &= x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n, \\ \varphi(x) &= x^m + b_1 x^{m-1} + \cdots + b_{m-1} x + b_m \end{aligned}$$

die Koeffizienten a, b voneinander unabhängige Variable. Der Euklidische Algorithmus (§ 15) hat uns gezeigt, wie wir durch rationale Rechnung zwei Funktionen $F(x)$, $\Phi(x)$ von den Graden $m-1$, $n-1$ bestimmen können, die der Bedingung

$$F(x) f(x) + \Phi(x)\ \varphi(x) = 1$$

genügen. Die Koeffizienten von $F(x)$ und $\Phi(x)$ sind rationale gebrochene Funktionen der a_i und b_i, und wenn wir mit dem Hauptnenner A heraufmultiplizieren, so können wir für diese Gleichung auch setzen:

$$(13)\qquad F(x) f(x) + \Phi(x)\ \varphi(x) = A,$$

worin A, $F(x)$ und $\Phi(x)$ ganze Funktionen der a_i, b_i sind.

Wir nehmen außerdem an, daß die drei Funktionen $A, F(x), \Phi(x)$ als Funktionen aller Variablen a_i, b_i, x keinen gemeinschaftlichen Teiler haben; denn ein solcher könnte in (13) weggehoben werden. Wir beweisen nun den Satz:

X. Wenn eine identische Gleichung (13) besteht, in der die drei ganzen Funktionen $F(x)$, $\Phi(x)$, A ohne gemeinsamen Teiler sind, und A von x frei ist, F und Φ den Grad $m-1$ und $n-1$ nicht übersteigen, so ist A bis auf einen numerischen Faktor die Resultante von f und φ.

Zum Beweis führen wir noch zwei weitere Reihen von n und m Variablen ein:

$$(14)\qquad \begin{matrix} \alpha_1, & \alpha_2 \ldots \alpha_n \\ \beta_1, & \beta_2 \ldots \beta_m. \end{matrix}$$

Wir haben dann nach § 25 den Satz, daß die Gleichung (13) richtig bleibt, wenn die Variablen

$$a_1, \quad a_2 \ldots a_n$$

durch die symmetrischen Grundfunktionen der α_i und die Variablen

$$b_1, \quad b_2 \ldots b_m$$

durch die symmetrischen Grundfunktionen der β_i ersetzt werden.

Durch dieselbe Substitution gehen $f(x)$ und $\varphi(x)$ über in

$$\begin{aligned} f(x) &= (x - \alpha_1)\,(x - \alpha_2) \ldots (x - \alpha_n), \\ \varphi(x) &= (x - \beta_1)\,(x - \beta_2) \ldots (x - \beta_m), \end{aligned} \tag{15}$$

und die Resultante R in

$$R(f, \varphi) = \overset{i,k}{\Pi}(\alpha_i - \beta_k) = \overset{i}{\Pi}\varphi(\alpha i) = (-1)^{mn}\overset{k}{\Pi} f(\beta_k). \tag{16}$$

Wenn man nach dieser Substitution $x = \beta_k$ setzt, so verschwindet $\varphi(x)$ und es ergibt sich aus (13)

$$F(\beta_k)\, f(\beta_k) = A. \tag{17}$$

Daraus folgt, daß A als Funktion der α, β wenigstens durch eine der Differenzen $\alpha_i - \beta_k$ teilbar ist, und da A eine symmetrische Funktion der α und der β ist, so muß es durch alle diese Differenzen, also nach (16) durch R teilbar sein.

Setzen wir demnach

$$A = HR \tag{18}$$

und fassen A und R als Funktionen der α, β auf, so folgt, daß auch H eine ganze symmetrische Funktion der α und der β ist.

Setzen wir

$$R = f(\beta_k) R_k, \tag{19}$$

so ist R_k nach (16) gleichfalls eine ganze, aber nicht symmetrische Funktion der α, β, und aus (17) und (18) folgt:

$$F(\beta_k) = HR_k. \tag{20}$$

Hiernach können wir durch die Lagrangesche Formel (§ 17) die Funktion $F(x)$ [die den $(m-1)$ten Grad nicht übersteigt] bestimmen. Es ergibt sich

$$F(x) = H \sum^{k} \frac{R_k f(x)}{f'(\beta_k)\,(x - \beta_k)}. \tag{21}$$

Da nun nach § 17, VI. die Summe nach k eine ganze Funktion der Variablen x, α, β ist, so ist $F(x)$ durch H teilbar, und H ist ein gemeinschaftlicher Teiler von $F(x)$ und A. Dann

aber müßte, da H von x unabhängig ist, also mit $\varphi(x)$ keinen Teiler $x - \beta_k$ gemein haben kann, H nach (13) auch in $\Phi(x)$ aufgehen. Das ist aber nach der Voraussetzung nur möglich, wenn H von den Variablen unabhängig, also eine Zahl ist. Hiermit ist unser Satz bewiesen.

Um diesen Satz für die Diskriminante anzuwenden, hat man nur $\varphi(x) = f'(x)$ zu setzen. Dann bleiben zwar die a_i, b_i nicht voneinander unabhängige Variable, sondern es ist $b_i = (n-i)a_i$. Es ändert dies aber in der Schlußweise nichts.

Hat man (durch den Euklidischen Algorithmus) eine Gleichung

$$F(x)f(x) + \Phi(x)f'(x) = A \tag{22}$$

gefunden, worin A von x frei ist und $F(x)$, $\Phi(x)$, A ganze Funktionen der a_i ohne gemeinsamem Teiler sind, so ist

$$\Phi(\alpha_i)f'(\alpha_i) = A$$

und folglich A teilbar durch

$$\Delta = \Pi(\alpha_i - \alpha_k).$$

Setzt man $A = H\Delta$ und $\Delta = f'(\alpha_i)\Delta_i$, so folgt

$$\Phi(\alpha_i) = H\Delta_i$$

$$\Phi(x) = H\sum^{i} \frac{\Delta_i f(x)}{f'(\alpha_i)(x - \alpha_i)}$$

und folglich ist $\Phi(x)$ durch H teilbar, woraus wie oben geschlossen wird, daß H eine Zahl sein muß.

Wir sprechen den in X. enthaltenen Satz aus:

XI. Wenn eine identische Gleichung (22) besteht, in der $F(x)$, $\Phi(x)$, A drei ganze Funktionen ohne gemeinsamem Teiler sind, und F den Grad $n-2$, Φ den Grad $n-1$ nicht übersteigt, so ist A bis auf einen numerischen Faktor die Diskriminante von $f(x)$.

§ 28.

Grad und Gewicht der Resultanten.

Die einzelnen Glieder der Resultante der zwei Funktionen $f(x)$ und $\varphi(x)$ vom $n$ten und $m$ten Grade haben, von einem numerischen (ganzzahligen) Faktor abgesehen, die Gestalt

$$a_0^{\nu_0} a_1^{\nu_1} \dots a_n^{\nu_n} b_0^{\mu_0} b_1^{\mu_1} \dots b_m^{\mu_m}, \tag{1}$$

worin, wie aus den oben bestimmten Graden hervorgeht,

$$(2)\qquad \begin{aligned} \nu_0 + \nu_1 + \cdots + \nu_n &= m \\ \mu_0 + \mu_1 + \cdots + \mu_n &= n. \end{aligned}$$

Wir können aber noch eine andere Relation zwischen diesen Exponenten angeben.

Da nämlich R eine homogene Funktion nmten Grades von den $n + m$ Variablen α, β ist, da ferner a_0 vom nullten, a_1 vom ersten, a_2 vom zweiten usw. Grade in den α, allgemein a_k und b_k vom $k$ten Grade in diesen Variablen sind, so folgt

$$(3)\qquad \begin{aligned} &\nu_1 + 2\nu_2 + 3\nu_3 + \cdots + n\nu_n \\ &+ \mu_1 + 2\mu_2 + 3\mu_3 + \cdots + m\mu_m = nm. \end{aligned}$$

Wenn wir den Index k von a_k oder b_k das Gewicht der betreffenden Größe nennen und unter dem Gewicht eines Produktes die Summe der Gewichte der einzelnen Faktoren verstehen, so können wir der Formel (3) auch den wörtlichen Ausdruck geben:

XII. Alle Glieder der entwickelten Resultante haben ein und dasselbe Gewicht mn.

Summen, deren Glieder alle dasselbe Gewicht haben, heißen isobarische Funktionen, und das Gewicht eines jeden Gliedes heißt das Gewicht der Funktion. Durch Multiplikation mehrerer isobarischer Funktionen entstehen wieder isobarische Funktionen, und das Gewicht des Produktes ist die Summe der Gewichte der Faktoren.

Die Resultanten sind isobarische Funktionen der Variablen a_k, b_k vom Gewichte mn.

Die Funktionen f und φ selbst kann man dadurch zu isobarischen machen, daß man der Variablen x das Gewicht 1 beilegt.

XIII. Ersetzt man die Variablen in einer isobarischen Funktion durch homogene Funktionen irgendwelcher anderer Variablen vom Grade des Gewichtes der betreffenden Variablen, so entsteht aus der isobarischen Funktion eine homogene Funktion der neuen Variablen, deren Grad dem Gewichte der Funktion gleich ist.

Wenn die Funktionen $f(x)$, $\varphi(x)$ durch Ordnen zweier homogener Funktionen der Variablen x, y, z ... vom Grade n und m

nach einer der Variablen x entstanden sind, so sind die Koeffizienten a_k, b_k homogene Funktionen vom Grade k der Variablen y, z ... und aus XII. und XIII. ergibt sich der Satz:

XIV. Durch Elimination einer Variablen x aus zwei homogenen Funktionen $n$ten und $m$ten Grades ergibt sich eine Endgleichung vom Grade mn in den übrigen Variablen.

Hierin ist das Theorem von Bézout enthalten, das in geometrischer Einkleidung so ausgesprochen wird, daß sich eine Kurve $n$ter und eine Kurve $m$ter Ordnung in mn Punkten schneiden.

Gibt man den Variablen a_k, b_k nicht die Gewichte k, sondern $(N + k)$ und $(M + k)$, worin N, M irgend zwei Zahlen sind, so wird das Gewicht der Resultante von f und φ, wie aus (2) und (3) hervorgeht,

$$mN + nM + mn. \tag{4}$$

Dies ist auch der Grad der Resultante, wenn die a_k und b_k durch Funktionen von neuen Variablen der Grade $N + k$ und $M + k$ ersetzt werden.

§ 29.

Größter gemeinschaftlicher Teiler.

Wir kehren noch einmal zur Resultante der beiden Funktionen $f(x)$ und $\varphi(x)$ vom $n$ten und $m$ten Grade zurück, um ihr noch eine bessere Form zu geben.

Wir bedienen uns dabei der in § 16 eingeführten Funktion $f_s(x)$, die wir in bezug auf die a_i homogen schreiben:

$$\begin{aligned} f_0(x) &= a_0 \\ f_1(x) &= a_0 x + a_1 \\ f_2(x) &= a_0 x^2 + a_1 x + a_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n(x) &= a_0 x^n + a_1 x^{n-1} + \cdots + a_n, \end{aligned} \tag{1}$$

so daß $f_n(x)$ mit der gegebenen Funktion $f(x)$ übereinstimmt. Die entsprechende Bedeutung sollen die Funktionen $\varphi_s(x)$ haben:

$$\varphi_s(x) = b_0 x^s + b_1 x^{s-1} + \cdots + b_s. \tag{2}$$

Wir können dann setzen:

$$\begin{aligned} f(x) &= x^s f_{n-s}(x) + F_{n-s}(x) \\ \varphi(x) &= x^s \varphi_{m-s}(x) + \Phi_{m-s}(x), \end{aligned} \tag{3}$$

worin

$$(4)\quad \begin{aligned} F_{n-s}(x) &= a_{n-s+1}x^{s-1} + a_{n-s+2}x^{s-2} + \cdots + a_n, \\ \Phi_{m-s}(x) &= b_{m-s+1}x^{s-1} + b_{m-s+2}x^{s-2} + \cdots + b_m, \end{aligned}$$

und aus (3) ergibt sich:

$$(5)\quad \varphi(x) f_{n-s}(x) - f(x)\varphi_{m-s}(x) = f_{n-s}\Phi_{m-s} - \varphi_{m-s}F_{n-s}.$$

Nehmen wir jetzt

$$(6)\qquad m \gtrless n$$

an, so ist

$$f_{n-s}\Phi_{m-s} - \varphi_{m-s}F_{n-s}$$

eine Funktion $n-1$ ten Grades von x, und wenn wir also s in (5) durch $n-s+1$ ersetzen, so folgt

$$(7)\quad \varphi f_{s-1} - f\varphi_{m-n+s-1} = A_{1,s}x^{n-1} + A_{2,s}x^{n-2} + \cdots A_{n,s}.$$

Da φ_s für negative s nicht definiert ist, so hat die Formel (7) nur so lange eine Bedeutung, als

$$(8)\qquad s > n - m$$

ist, also für

$$s = n - m + 1, \quad n - m + 2, \quad \ldots n$$

[für $s = n + 1$ würde die linke Seite von (7) identisch Null].

Wir definieren daher für

$$(9)\qquad s \leqq n - m:$$

$$(10)\qquad \varphi x^{s-1} = A_{1,s}x^{n-1} + A_{2,s}x^{n-2} + \cdots + A_{n,s}.$$

Die $A_{r,s}$ sind dann unter der Voraussetzung (8) bilineare Funktionen der a_i, b_i, und unter der Voraussetzung (9) lineare Funktionen der b_i allein, nämlich

$$A_{r,s} = b_{m-n+r+s-1}$$

oder $= 0$, wenn $m - n + r + s < 1$.

Die Gleichungen (7), (10) lassen sich in die eine zusammenfassen:

$$(11)\quad \begin{aligned} \varphi p_s - f q_s &= A_{1,s}x^{n-1} + A_{2,s}x^{n-2} + \cdots + A_{n,s} \\ &= \sum_{1,n}^{r} A_{r,s}x^{n-r}, \end{aligned}$$

wenn wir setzen:

$$(12)\quad \begin{aligned} p_s &= x^{s-1} \\ q_s &= 0 \end{aligned} \qquad \text{für } s = 1, 2, 3, \ldots n - m$$

$$(13)\quad \begin{aligned} p_s &= f_{s-1} \\ q_s &= \varphi_{m-n+s-1} \end{aligned} \qquad \text{für } s = n - m + 1, n - m + 2, \ldots n.$$

Multiplizieren wir die Gleichungen (11) mit willkürlichen Faktoren c_s, lassen s von 1 bis n gehen und setzen:

$$\begin{aligned} \sum_{1,n}^{s} c_s p_s &= P(x), \\ \sum_{1,n}^{s} c_s q_s &= Q(x), \\ \sum_{1,n}^{s} c_s A_{r,s} &= C_r, \end{aligned} \tag{14}$$

so ergibt sich:

$$\varphi(x) P(x) - f(x) Q(x) = \sum_{1,n}^{r} C_r x^{n-r}, \tag{15}$$

worin $P(x)$ und $Q(x)$ ganze rationale Funktionen von x von den Graden $n-1$, $m-1$ sind, die nach (12) nur dann identisch verschwinden können, wenn die sämtlichen c_s verschwinden.

Wir bezeichnen jetzt mit

$$R = \begin{vmatrix} A_{1,1}, & A_{2,1} \dots & A_{n,1} \\ A_{1,2}, & A_{2,2} \dots & A_{n,2} \\ \dots\dots & \dots\dots & \dots\dots \\ A_{1,n}, & A_{2,n} \dots & A_{n,n} \end{vmatrix} \tag{16}$$

die Determinante der $A_{i,k}$ und die Unterdeterminanten mit

$$R_{r,s} = \frac{\partial R}{\partial A_{r,s}}. \tag{17}$$

R ist nach (8) und (9) eine homogene Funktion $n$ten Grades der b_i und $m$ten Grades der a_i

Setzen wir dann

$$c_1 = R_{n,1}, \quad c_2 = R_{n,2}, \quad \dots c_n = R_{n,n}, \tag{18}$$

so ergibt sich aus (14):

$$C_1 = 0, \quad C_2 = 0, \quad \dots C_n = R,$$

und aus (15) folgt:

$$\varphi(x) P(x) - (x) Q(x) = R. \tag{19}$$

Es ist unmöglich, daß infolge der Annahme (18) die Funktionen P, Q, R identisch verschwinden; denn nehmen wir den speziellen Fall

$$f(x) = a_0 x^n, \quad \varphi(x) = b_0 x^m, \tag{20}$$

worin n beliebig, $m = 0, 1, 2, \dots n$ sein kann, so ergibt sich aus (12) und (13):

$$(21)\qquad \begin{aligned} p_s &= x^{s-1} \\ q_s &= 0 \end{aligned} \qquad s = 1, 2, 3, \ldots n - m$$

$$\begin{aligned} p_s &= c_0 x^{s-1} \\ q_s &= b_0 x^{m-n+s-1} \end{aligned} \qquad s = n - m + 1, n - m + 2, \ldots n.$$

$$(22)\qquad \begin{aligned} \varphi p_s - f q_s &= b_0 x^{m+s-1}, && s = 1, 2, 3, \ldots n - m \\ &= 0, && s = n - m + 1, \ldots n, \end{aligned}$$

also wird nach (11)

$$(23)\qquad A_{n-m-s+1,s} = b_0 \qquad s \lesseqgtr n - m$$

und alle übrigen $A_{r,s}$ sind $= 0$.

Nehmen wir $m = 0$, so wird die Determinante R:

$$\begin{vmatrix} 0, & 0 \ldots 0, & b_0 \\ 0, & 0 \ldots b_0, & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ b_0, & 0 \ldots 0, & 0 \end{vmatrix} = \pm b_0^n$$

von Null verschieden; nehmen wir für m einen anderen Wert, so geht s in (22) nur bis $n - m$ und die Hauptunterdeterminante

$$R^{(n-m)} = \Sigma \pm A_{1,1} A_{2,2} A_{n-m,n-m}$$

erhält dieselbe Form wie R und den Wert $\pm b_0^{n-m}$, während die anderen Hauptunterdeterminanten verschwinden.

Nun sind P und Q von den Graden $n - 1$, $m - 1$ (höchstens) in x, und R ist von x unabhängig und in bezug auf die a_i, b_i von den Graden m und n, und wir haben nach § 27, X, den Satz:

XV. Die Determinante R ist die Resultante $R(f'\varphi)$ der beiden Funktionen $f(x)$ und $\varphi(x)$.

Um die Frage nach den gemeinschaftlichen Teilern der zwei Funktionen f und φ allgemeiner zu behandeln, führen wir die Hauptunterdeterminanten von R ein:

$$R^{(n)},\quad R^{(n-1)},\quad R^{(n-2)},\quad \ldots R^{(1)},$$

indem wir setzen:

$$(24)\qquad \begin{aligned} R^{(\nu)} &= \Sigma \pm A_{1,1} A_{2,2} \ldots A_{\nu,\nu} && \nu = 1, 2 \ldots n \\ R^{(\nu)}_{h,k} &= \frac{\partial R^{(\nu)}}{\partial A_{h,k}} && h = 1, 2 \ldots \nu \\ & && k = 1, 2 \ldots \nu. \end{aligned}$$

Im besonderen ist $R^{(n)} = R$ und $R^{(1)} = A_{1,1}$.

Nun setzen wir in den Formeln (14) für irgend einen Index μ zwischen 1 und n:

$$(25)\qquad \begin{gathered} c_1 = R^{(\mu)}_{\mu,1},\quad c_2 = R^{(\mu)}_{\mu,2} \ldots c_\mu = R^{(\mu)}_{\mu,\mu}, \\ c_{\mu+1} = 0,\quad c_{\mu+2} = 0 \ldots c_n = 0 \end{gathered}$$

und erhalten:

$$(26)\qquad C_r = \sum_{1,\mu}^{s} R^{(\mu)}_{\mu,s} A_{r,s} = \begin{vmatrix} A_{1,1}, & A_{2,1}, \ldots & A_{\mu-1,1}, & A_{r,1} \\ A_{1,2}, & A_{2,2}, \ldots & A_{\mu-1,2}, & A_{r,2} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ A_{1,\mu}, & A_{2,\mu}, \ldots & A_{\mu-1,\mu}, & A_{r,\mu} \end{vmatrix},$$

und daraus folgt:

$$C_1 = 0,\quad C_2 = 0 \ldots C_{\mu-1} = 0,\quad C_\mu = R^{(\mu)},$$

während für $\nu > \mu$ die C_ν ganze Funktionen der a_i, b_i werden, und weil die p_s Funktionen vom Grade $s - 1$ sind, so ergibt sich bei dieser Annahme, daß $P(x)$ vom Grade $\mu - 1$ ist.

Demnach ergibt sich aus (15)

$$(27)\qquad \varphi(x) P(x) - f(x) Q(x) = T_\mu(x),$$

worin

$$(28)\qquad T_\mu(x) = R^{(\mu)} x^{n-\mu} + C_{\mu+1} x^{n-\mu-1} + \cdots + C_n$$

eine ganze Funktion $(n - \mu)$ten Grades von x ist, die man nach (26) auch als Determinante schreiben kann:

$$T_\mu = \begin{vmatrix} A_{1,1}, & A_{2,1} \ldots A_{\mu-1,1}, & A_{\mu,1} x^{n-\mu} + A_{\mu+1,1} x^{n-\mu-1} + \cdots + A_{n,1} \\ A_{1,2}, & A_{2,2} \ldots A_{\mu-1,2}, & A_{\mu,2} x^{n-\mu} + A_{\mu+1,2} x^{n-\mu-1} + \cdots + A_{n,2} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ A_{1,\mu}, & A_{2,\mu} \ldots A_{\mu-1,\mu}, & A_{\mu,\mu} x^{n-\mu} + A_{\mu+1,\mu} x^{n-\mu-1} + \cdots + A_{n,\mu} \end{vmatrix}.$$

Daß auch hier $P(x)$, $Q(x)$, $P_\mu(x)$ nicht identisch verschwinden, ergibt sich wie oben, wenn man in der speziellen Annahme (20) $m = n - \mu$ setzt, wobei $R^{(\mu)} = \pm b_0^{n-m}$, also nicht Null wird.

Hieraus schließen wir folgendes:

Jeder gemeinschaftliche Teiler von f und φ muß ein Teiler von T_μ sein, und wenn $R^{(\mu)}$ nicht verschwindet, so können f und φ keinen Teiler von höherem als dem $(n - \mu)$ten Grade haben.

Haben f und φ einen gemeinschaftlichen Teiler vom Grade $n - \mu$, und ist $R^{(\mu)}$ von Null verschieden, so ist dieser Teiler T_μ.

Haben f und φ einen gemeinschaftlichen Teiler vom Grade $n - \mu$, ist aber $R^{(\mu)} = 0$, so muß T_μ identisch verschwinden und es folgt aus (27):

$$\varphi(x) P(x) = f(x) Q(x).$$

Es muß also $\varphi(x) P(x)$ durch $f(x)$ teilbar sein. Wenn nun selbst $f(x)$ durch $P(x)$ teilbar wäre, so wäre der Quotient immer

noch vom Grade $n - \mu + 1$, weil $P(x)$ vom Grade $\mu - 1$ ist, und durch diesen Quotienten wäre $f(x)$ und $\varphi(x)$ teilbar. Bei dieser Annahme ist also der größte gemeinschaftliche Teiler von f und φ mindestens vom Grade $n - \mu + 1$.

Gehen wir also von der Tatsache aus, daß $R = 0$ die Bedingung dafür ist, daß $f(x)$ und $\varphi(x)$ einen gemeinschaftlichen Teiler mindestens vom ersten Grade haben, so kommen wir durch vollständige Induktion zu dem Satze:

XVI. Die notwendige und hinreichende Bedingung dafür, daß die Funktionen $f(x)$ und $\varphi(x)$ einen gemeinschaftlichen Teiler vom Grade $n - \mu$, aber nicht von höherem Grade haben, ist die, daß die Hauptunterdeterminanten

$$R^{(n)},\ R^{(n-1)},\ \ldots\ R^{(\mu+1)}$$

verschwinden, daß aber $R^{(\mu)}$ nicht verschwindet. Der größte gemeinschaftliche Teiler ist durch die Determinante T_μ ausgedrückt.

Vierter Abschnitt.

Wurzeln.

§ 30.

Über die Gaußschen Beweise des Fundamentalsatzes.

Der Fundamentalsatz der Algebra besteht in der Behauptung, daß jede algebraische Gleichung $f(x) = 0$, deren Koeffizienten numerisch gegeben sind, eine Wurzel hat, d. h. daß man einen reellen oder imaginären Zahlenwert α angeben könne, der, für x gesetzt, dem Ausdruck $f(x)$ einen Wert verleiht, der dem absoluten Werte nach kleiner ist als eine beliebig klein anzunehmende Größe. Durch eine scharfe Formulierung des Zahlbegriffes kann man diesen Satz auch so ausdrücken:

> Es gibt eine reelle oder komplexe Zahl α, die der Bedingung $f(\alpha) = 0$ genügt.

Die Koeffizienten der Funktion f sind beliebig gegebene reelle oder komplexe Zahlen.

Es liegt keine Beschränkung darin, wenn wir die Koeffizienten von f reell annehmen. Denn wenn $f(x)$ imaginäre Koeffizienten hat, so nehme man für diese Koeffizienten überall die konjugiert imaginären Werte und bezeichne die so entstandene konjugierte Funktion mit $\varphi(x)$. Dann ist

$$f(x)\,\varphi(x) = F(x)$$

eine Funktion mit reellen Koeffizienten. Hat diese eine Wurzel $x = \alpha$, so ist entweder

$$f(\alpha) = 0 \quad \text{oder} \quad \varphi(\alpha) = 0.$$

Ist etwa $f(\alpha) = 0$ und β die zu α konjugiert imaginäre Zahl, so ist auch $\varphi(\beta) = 0$, d. h. sowohl $f(x)$ als $\varphi(x)$ hat eine Wurzel.

Wenn $f(\alpha) = 0$ ist, so ist $f(x)$ durch $x - \alpha$ teilbar. Man setze also

$$f(\alpha) = (x - \alpha) f_1(x),$$

worin der Grad der ganzen Funktion $f_1(x)$ um eins niedriger ist als der von $f(x)$.

Wenn nun $f_1(x)$ wieder eine Wurzel α_1 hat, so kann man

$$f_1(x) = (x - \alpha_1) f_2(x)$$

setzen usf., wo die $f(x)$, $f_1(x)$, $f_2(x)$ ... ganze Funktionen von abnehmenden Graden sind. In diesen Funktionen hat die höchste Potenz von x, nämlich x^n, x^{n-1}, x^{n-2} ... immer denselben Koeffizienten a_0, und so kann man den Fundamentalsatz auch so aussprechen:

I. Eine ganze Funktion $n$ten Grades

(1) $$f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

kann in n lineare Faktoren zerlegt werden:

(2) $$f(x) = a_0 (x - \alpha)(x - \alpha_1) \ldots (x - \alpha_{n-1}),$$

worin die $\alpha, \alpha_1, \ldots \alpha_{n-1}$ reelle oder komplexe Zahlen sind, die auch zum Teil einander gleich sein können.

Den ersten Beweis dieses Satzes hat Gauß in seiner Doktor-Dissertation vom Jahre 1799 gegeben.

Er stützt sich dabei auf eine geometrische Anschauung. Er setzt nämlich

(3) $$z = x + iy,$$

(4) $$f(z) = X + iY,$$

worin X und Y reelle Funktionen von x und y sind, selbst in dem Falle, wo die Koeffizienten von $f(z)$ imaginär sind; $f(x)$ wird dann und nur dann $= 0$ sein, wenn gleichzeitig

(5) $$X = 0, \quad Y = 0$$

ist Deutet man x, y als Cartesische Koordinaten, so stellen die Gleichungen (5) zwei Kurven dar, und es kommt also darauf an, sich zu überzeugen, daß diese Kurven wenigstens einen Schnittpunkt haben.

50 Jahre später, 1849, ist Gauß auf diesen Beweis zurückgekommen und hat ihm eine noch bessere Gestalt gegeben.

Der zweite Beweis des Fundamentalsatzes von Gauß vom Jahre 1816 beruht auf anderer Grundlage. Er stützt sich zwar

auch auf die Stetigkeit, auf der der Satz beruht, daß eine stetige reelle Funktion $f(x)$, die für zwei reelle Werte $x = a$ und $x = b$ entgegengesetzte Zeichen hat, zwischen diesen Werten wenigstens einmal gleich Null sein muß, woraus dann folgt, daß eine Gleichung ungeraden Grades mit reellen Koeffizienten mindestens eine reelle Wurzel haben muß, weil eine Funktion $f(x)$ von ungeradem Grade für hinlänglich große positive oder negative Werte von x sowohl positiv als negativ gemacht werden kann. Es wird sodann durch eine elegante Transformation die Auflösung irgend einer Gleichung auf eine Gleichung ungeraden Grades zurückgeführt.

Der dritte Beweis endlich vom selben Jahre, 1816, beruht auf funktionentheoretischen Sätzen über die Integration einer Funktion komplexen Arguments, die Gauß schon früh gefunden hatte, aber erst viel später durch Cauchy, Riemann u. a. weiteren Kreisen bekannt wurden.

Es schließen sich daran andere Beweise von Cauchy, Weierstrass, Lipschitz, die den bei allen benutzten Begriff der Stetigkeit tiefer zu begründen suchen, ohne den es einmal nicht abgeht. Zu diesen gehört auch der in der großen Ausgabe der Algebra mitgeteilte.

Einen schönen Beweis hat Gordan gegeben, der sich wie der zweite Beweis von Gauß auf die einleuchtende Tatsache stützt, daß eine reelle Gleichung ungeraden Grades immer eine Wurzel hat, die Zurückführung einer Gleichung geraden Grades auf eine von ungeradem Grade aber sehr vereinfacht. Diesem Gordanschen Beweis wollen wir hier folgen[1]).

§ 31.

Beweis des Fundamentalsatzes nach Gordan.

Es sei $f(x)$ eine Funktion mit reellen Koeffizienten, von der die Existenz einer Wurzel nachgewiesen werden soll. Der Grad von $f(x)$ sei

$$n = 2^{\mu} m, \tag{1}$$

worin μ eine positive, m eine ungerade Zahl sei. Um für die vollständige Induktion eine Grundlage zu gewinnen, nennt Gordan

[1]) Gordan, Mathematische Annalen **10** (1876) und Vorlesungen über Invariantentheorie **1**, § 12 (1885).

von zwei Gleichungen der Größe:

$$n = 2^{\mu} m, \quad n' = 2^{\mu'} m',$$

die zweite „leichter auflösbar" als die erste, wenn entweder

$$\mu' < \mu$$

oder

$$\mu' = \mu, \quad m' < m, \tag{2}$$

und dadurch sind die Gleichungen nach der Schwierigkeit der Auflösung in Klassen geteilt, deren erste für $\mu = 0$ die Gleichungen ungeraden Grades sind, die nach Voraussetzung „auflösbar" sind.

Das Ziel wird erreicht sein, wenn wir die Auflösung der Gleichung $n$ten Grades unter der Voraussetzung nachweisen können, daß alle leichter auflösbaren Gleichungen eine Wurzel haben.

Nun sei u eine Variable und die Koeffizienten $a_1, a_2, \ldots a_n$ von $f(x)$ seien vorläufig auch unabhängige Variable. Wir setzen:

$$\begin{aligned} \frac{f(x+u) + f(x-u)}{2} &= F(x) \\ \frac{f(x+u) - f(x-u)}{2u} &= \Phi(x). \end{aligned} \tag{3}$$

Beides sind ganze Funktionen von x, sie sind aber auch ganze Funktionen von u, weil $f(x+u) - f(x-u)$ durch u teilbar ist, und ganze Funktionen der $a_1, a_2, \ldots a_n$. Sie ändern sich nicht, wenn u in $-u$ verwandelt wird und enthalten folglich nur die geraden Potenzen von u.

Aus (3) ergibt sich:

$$\begin{aligned} f(x+u) &= F(x) + u\Phi(x) \\ f(x-u) &= F(x) - u\Phi(x). \end{aligned} \tag{4}$$

Jetzt untersuchen wir die Resultante $R(F, \Phi)$ und diese ist nach dem Satze § 27, (9) gleich der Resultanten

$$R[F(x) - u\Phi(x), \Phi(x)] = R[f(x-u), \Phi(x)] = P(u). \tag{5}$$

Sie ist eine ganze Funktion von u, deren Grad wir jetzt bestimmen müssen.

Um diese Resultante für variable a_i zu bilden, führen wir n neue unabhängige Variable $\alpha_1, \alpha_2, \ldots \alpha_n$ ein, und dann bleibt jede identische Gleichung zwischen den Variablen $a_1, a_2, \ldots a_n$ richtig, wenn wir für die a_i die symmetrischen Grund-

funktionen der α_i setzen. Dadurch geht die Funktion

$$f(x) = x^n + a_1 x^{n-1} + \cdots a_{n-1} x + a_n$$

in das Produkt

$$(x - \alpha_1)(x - a_2) \ldots (x - \alpha_n)$$

über und die Variablen α_i können die Wurzeln von $f(x)$ genannt werden, ohne daß damit über die Existenz der Wurzeln einer bestimmten numerischen Gleichung etwas vorausgesetzt wird.

Es ist dann:

$$f(x + u) = \overset{i}{\Pi}(x + u - \alpha_i),$$

$$f(x - u) = \overset{i}{\Pi}(x - u - \alpha_i),$$

und die Wurzeln von $f(x - u)$ sind $u + \alpha_i$.

Demnach ist die Resultante (5) nach § 27, (5)

$$P(u) = \overset{i}{\Pi}\Phi(u + \alpha_i).$$

Es ist aber nach (3)

$$\Phi(u + \alpha_i) = \frac{1}{2u} f(2u + \alpha_i) \tag{6}$$

und folglich

$$P(u) = \frac{\overset{i}{\Pi} f(2u + \alpha_i)}{(2u)^n}. \tag{7}$$

Ordnet man den Zähler $f(2u + \alpha_i)$ nach absteigenden Potenzen von u, so läßt er sich wegen $f(\alpha_i) = 0$ durch u dividieren und man kann $P(u)$ durch die α_i und u ausdrücken. Das höchste Glied von $P(u)$ ist $= (2u)^{n(n-1)}$ in bezug auf u, und der Grad $n(n-1)$ kann sich nicht erniedrigen, wenn auch für die α_i irgendwelche spezielle Zahlenwerte gesetzt werden.

Nun enthalten die Funktionen $F(x)$ und $\Phi(x)$ nur die geraden Potenzen von u und dasselbe gilt daher auch von der Resultanten $R(F, \Phi)$, d. h. von $P(u)$, deren Grad in u^2 also gleich

$$\tfrac{1}{2} n(n-1)$$

ist. Dieser Grad ist also nach (1) nur durch $2^{\mu-1}$, nicht durch 2^μ teilbar.

Die Resultante $P(u) = R(F, \Phi)$ läßt sich aber aus F und Φ durch rationale Rechnung bilden und hat daher wie F und Φ reelle Koeffizienten.

Werden jetzt für die α_i irgendwelche reelle Zahlen gesetzt, so folgt, daß $P(u)$ in bezug auf u^2 leichter auflösbar ist als

$f(x)$, und weil eine Quadratwurzel aus einer reellen und aus einer komplexen Zahl immer existiert, so folgt aus der Voraussetzung, daß es einen reellen oder komplexen Wert v gibt, der $P(v) = 0$ macht.

Setzen wir $u = v$, so erhalten $F(x)$ und $\Phi(x)$ und folglich nach (3) auch $f(x+v)$, $f(x-v)$ und $\Phi(x)$ als Funktionen von x einen gemeinschaftlichen Teiler, der, weil $\Phi(x)$ nur vom Grade $n-1$ ist, von niedrigerem Grade als n sein muß. Es hat also auch $f(x)$ selbst einen echten Teiler $\varphi(x)$, der aber auch komplex sein kann.

Ist dann $f(x) = \varphi(x,v)\,\varphi_1(x,v)$, so ist jedenfalls der Grad von einem der beiden Faktoren $\gtreqless \frac{1}{2}n$, und wenn wir diesen für $\varphi(x,v)$ nehmen, so können wir also als bewiesen annehmen, daß $\varphi(x,v)$ von einem Grade $\gtreqless \frac{1}{2}n$ ist. Ist v' die zu v konjugierte Zahl, so ist die reelle Funktion $f(x)$ auch durch $\varphi(x,v')$ teilbar und folglich auch durch das kleinste gemeinschaftliche Vielfache $f_1(x)$ von $\varphi(x,v)$ und $\varphi(x,v')$, was jedenfalls reell ist, und dessen Grad kleiner ist als n, außer wenn $\varphi(x,v)$ vom Grade $\frac{1}{2}n$ und teilerfremd zu $\varphi(x,v')$ ist; abgesehen von diesem Ausnahmefall folgt also aus unserer Annahme ein reeller echter Teiler von $f(x)$.

Es bleibt noch der Fall zu betrachten, daß in der Zerlegung von $f(x)$ in $\varphi(x,v)\,\varphi_1(x,v)$ beide Faktoren vom Grade $\frac{1}{2}n$ und nicht weiter in Faktoren zerlegbar sind. Wir setzen in diesem Falle

$$f(x) = \varphi(x)\,\psi(x), \tag{8}$$

wo dann $\varphi(x)$ und $\psi(x)$ unzerlegbare konjugiert imaginäre Funktionen vom Grade $\frac{1}{2}n$ sind. Wenn wir den Koeffizienten der höchsten Potenz von x in $f(x)$ gleich 1 annehmen, so können wir auch die Koeffizienten von $x^{\frac{1}{2}n}$ in $\varphi(x)$ und $\psi(x)$ gleich 1 annehmen.

Nach (8) können wir bilden:

$$\begin{aligned} f(x+v) &= \varphi(x+v)\,\psi(x+v) \\ f(x-v) &= \varphi(x-v)\,\psi(x-v). \end{aligned} \tag{9}$$

Diese beiden Funktionen müssen einen gemeinschaftlichen Teiler haben, und da die Funktionen φ und ψ unzerlegbar sind und es gleichgültig ist, welchen der beiden Faktoren wir mit φ und welchen mit ψ bezeichnen, so muß entweder

$$\varphi(x+v) = \varphi(x-v) \tag{10}$$

oder

$$\varphi(x+v) = \psi(x-v) \tag{11}$$

sein. Setzen wir

$$\varphi(x) = x^{n/2} + c_1 x^{n/2-1} + \cdots \tag{12}$$

und bezeichnen mit $c_1' \ldots$ die zu $c_1 \ldots$ konjugierten Werte, so ist

$$\begin{aligned} \psi(x) &= x^{n/2} + c_1' x^{n/2-1} + \cdots \\ \varphi(x+v) - \varphi(x-v) &= n v x^{n/2-1} + \cdots \end{aligned} \tag{13}$$

und dies müßte im Falle (10) identisch verschwinden.

Da n nicht $= 0$ ist, so müßte $v = 0$ sein und es müßte nach (3) $f(x)$ und $f'(x)$ einen gemeinsamen Teiler, also $f(x)$ einen echten Teiler haben.

Ferner folgt

$$\varphi(x+v) - \psi(x-v) = (nv + c_1 - c_1')x^{n/2-1}$$

und folglich müßte im Falle (11)

$$nv = c_1' - c_1$$

sein. Es wäre also v rein imaginär und folglich, da φ und ψ konjugiert imaginär sind, $\varphi(x+v)$ wegen (11) reell. Es hat also nach (9) $f(x+v)$ einen reellen Faktor $\varphi(x+v) = \chi(x)$, der als vom Grade $n/2$ leichter löslich ist als $f(x)$ und folglich als gelöst gilt. Es hat also $f(x+v)$ einen linearen Faktor $x - \alpha$ und folglich $f(x)$ den linearen Faktor $x - v - \alpha$; folglich auch, wenn $v + \alpha$ imaginär ist, den konjugierten Faktor, und $f(x)$ hat einen reellen Faktor ersten oder zweiten Grades. Damit ist bewiesen:

II. Die reelle Funktion $f(x)$ hat einen reellen Faktor $f_1(x)$, dessen Grad niedriger als n ist und nur im Falle $n = 2$ gleich n sein kann.

Ist

$$f(x) = f_1(x) f_2(x)$$

und sind

$$n = 2^\mu m, \quad n_1 = 2^{\mu_1} m_1, \quad n_2 = 2^{\mu_2} m_2$$

die Grade der Funktion f, f_1, f_2, so ist $n = n_1 + n_2$ und folglich

$$2^\mu m = 2^{\mu_1} m_1 + 2^{\mu_2} m_2,$$

und es können nicht beide Zahlen μ_1 und μ_2 größer als μ sein. Sei $\mu_1 \gtreqless \mu$. Ist $\mu_1 = \mu$, so ist $m_1 < m$ und folglich $f_1(x)$ leichter lösbar als $f(x)$; und ist $\mu_1 < \mu$, so ist gleichfalls $f_1(x)$ leichter lösbar als $f(x)$. Damit ist durch vollständige Induktion bewiesen:

III. Jede Funktion $f(x)$ hat eine reelle oder imaginäre Wurzel

und also zugleich das Theorem I.

§ 32.

Stetigkeit. Anderer Beweis des Fundamentalsatzes.

Ein anderer Weg, zu dem Grundsatze über algebraische Gleichungen zu gelangen, geht aus von dem Begriff der Stetigkeit, über den wir einige allgemeine Betrachtungen vorausschicken müssen.

Eine reelle Funktion $f(x)$ heißt in dem Intervall

$$a \leqq x \leqq b \tag{1}$$

stetig, wenn sie in dem ganzen Intervall endlich bleibt und die Eigenschaft hat, daß

$$f(x \pm h) - f(x) \tag{2}$$

für jeden Wert x des Intervalls unter einer beliebig gegebenen Zahl η liegt, sobald das positive h kleiner als eine hinlänglich kleine Zahl ε bleibt. (Für $x = a$ nehmen wir nur das obere, für $x = b$ nur das untere Zeichen, um nicht aus dem Intervall herauszukommen [1]).

Für die reellen Zahlen gilt nun der Dedekindsche Begriff der Stetigkeit, den wir so formulieren:

IV. Werden die reellen Zahlen so in zwei Teile $\mathfrak{A}$ und $\mathfrak{B}$ geteilt, daß jede Zahl in $\mathfrak{A}$ kleiner ist als jede Zahl in $\mathfrak{B}$, so gibt es eine und nur eine Zahl ξ, die die Zahlen $\mathfrak{A}$ von den Zahlen $\mathfrak{B}$ trennt, d. h. ξ ist großer als jede Zahl in $\mathfrak{A}$ und kleiner, als jede Zahl in $\mathfrak{B}$, wobei ξ selbst entweder zu $\mathfrak{A}$ oder zu $\mathfrak{B}$ gehören kann.

Eine solche Einteilung heißt ein Schnitt und wird mit $\mathfrak{A}|\mathfrak{B}$ bezeichnet [2]).

---

[1]) Die Funktion $f(x)$ heißt gleichmäßig stetig, wenn sich für jedes gegebene η ein für das ganze Intervall gültiges ε bestimmen läßt, das dieser Forderung genügt. Hier kommt diese Unterscheidung nicht in Betracht. Übrigens hat Lüroth (Mathem. Ann. 6) bewiesen, daß eine stetige Funktion einer Variablen immer gleichmäßig stetig ist.

[2]) Dedekind vollzieht den Schnitt nur im Bereich der rationalen Zahlen und definiert dadurch die irrationalen Zahlen.

Aus diesem Satze ergeben sich sehr leicht die Folgerungen:

V. Wenn S ein beliebiges System reeller Zahlen bedeutet, die alle größer sind als eine bestimmte positive oder negative Zahl C, so existiert eine untere Grenze g für die Zahlen S. Ist S' ein Teil von S, so haben auch die Zahlen S' eine untere Grenze g', die entweder gleich g oder größer als g ist.

Unter einer unteren Grenze ist g hier eine Zahl zu verstehen, die von keiner der Zahlen unterschritten wird, aber so, daß, wenn δ eine beliebig gegebene positive Größe ist, zwischen g und $g + \delta$ (mit Einschluß der Grenzen) immer wenigstens eine Zahl aus S liegt.

Gehört g selbst zu den Zahlen S, ist also g die kleinste Zahl in S, so heißt g das Minimum von S. Dies tritt z. B. dann ein, wenn S nur eine endliche Anzahl von Zahlen enthält.

VI. Ist $f(x)$ eine reelle stetige Funktion und ist $f(a)$ negativ, $f(b)$ positiv, so gibt es im Intervall $(a\,b)$ einen Wert ξ, für den $f(\xi) = 0$ ist.

Darauf beruht der schon benutzte Satz, daß eine reelle Gleichung ungeraden Grades immer eine reelle Wurzel hat.

VII. Eine reelle stetige Funktion $f(x)$ hat in dem Intervall $(a\,b)$ ein Minimum.

Ist g die untere Grenze der Funktionswerte $f(x)$ in dem Intervall, so ist nach V. die untere Grenze der Funktionswerte in einem Teil des Intervalls entweder gleich g oder größer als g.

Wenn nun $f(a) = g$ ist, so ist das zu Beweisende richtig; ist aber $f(a) > g$, so kann man wegen der Stetigkeit von $f(x)$ ein Intervall $a \leqq x \leqq a + h$ angeben, in dem $f(x) > g$ bleibt und also die untere Grenze von $f(x)$ größer als g ist.

Wir konstruieren nun in dem Intervall einen Schnitt $\mathfrak{A}|\mathfrak{B}$ derart, daß wir einen Wert α des Intervalls zu $\mathfrak{A}$ rechnen, wenn die untere Grenze von $f(x)$ in dem Intervall $a \leqq x \leqq \alpha$ größer als g ist, und einen Wert β zu $\mathfrak{B}$, wenn die untere Grenze im Intervall $a \leqq x \leqq \beta$ gleich g ist.

$\mathfrak{B}$ kann möglicherweise aus dem einzigen Werte b bestehen; dann aber muß $f(b) = g$ sein, und unsere Behauptung ist für $x = b$ erfüllt; denn wäre $f(b) > g$, so könnte man eine Größe

g' zwischen $f(b)$ und g und wegen der Stetigkeit ein Intervall $b - h \leqq x \leqq b$ so annehmen, daß in diesem Intervall alle Funktionswerte $f(x)$ größer als g', ihre untere Grenze also gleich oder größer als g' und daher sicher größer als g wäre. Da aber $b - h$ zu $\mathfrak{A}$ gehört, so ist auch in dem Intervall $a \leqq x \leqq b - h$ und folglich in dem ganzen Intervall (1) die untere Grenze größer als g, gegen die Annahme.

Ist also $f(b) > g$, so definiert der Schnitt $\mathfrak{A} \mid \mathfrak{B}$ eine Zahl ξ im Intervall, von der wir zeigen, daß

$$f(\xi) = g \tag{3}$$

ist.

Ist $f(\xi) > g' > g$, so können wir wegen der Stetigkeit von $f(x)$ ein Intervall

$$\xi - h \lessgtr x \lessgtr \xi + h$$

bestimmen, in dem $f(x)$ größer als g' und daher die untere Grenze größer als g ist. Da aber $\xi - h$ zu $\mathfrak{A}$ gehört, so ist auch in dem ganzen Intervall $a \leqq x \leqq \xi + h$ die untere Grenze größer als g, während doch $\xi + h$ zu $\mathfrak{B}$ gehört. Also muß $f(\xi) = g$ sein.

VIII. Die ganze Funktion $n$ten Grades $f(x)$ wird mit x zugleich unendlich.

Das will sagen: Wenn $f(x)$ eine ganze Funktion mit komplexen oder reellen Koeffizienten ist und die Variable x ebenfalls komplex ist, wenn ferner C ein beliebiger positiver Wert ist, so kann eine positive Größe R so angenommen werden, daß

$$|f(x)| > C \tag{4}$$

ist, sobald

$$|x| \geqq R \tag{5}$$

wird. Um den Satz zu beweisen, setze man $f(x)$ in die Form

$$f(x) = x^n \left(A + \frac{f_1(x)}{x^n} \right),$$

worin A von Null verschieden und $f_1(x)$ höchstens vom $(n-1)$ten Grade ist.

Es läßt sich also R so annehmen, daß $|f_1(x)/x^n|$, das nur negative Potenzen von x enthält, unter der Bedingung (5) beliebig klein wird, und da nach einem bekannten Satz der abso-

lute Wert einer Summe niemals kleiner ist als die Differenz der absoluten Werte, so folgt

$$|f(x)| \gtrless R^n \left(|A| - \left| \frac{f_1(x)}{x^n} \right| \right),$$

und da $|A|$ einen positiven Wert hat, so kann man R so groß nehmen, daß (4) erfüllt ist.

Da x^n mit von Null an wachsendem x fortwährend und über alle Grenzen wächst, so ist $x^n - a$ für genügend große x positiv und für $x = 0$ negativ; also gibt es nach VI. ein positives x, für das $x^n - a = 0$ ist, d. h. $\sqrt[n]{a}$ hat für jedes positive a einen und nur einen reellen positiven Wert und bei ungeradem n hat $\sqrt[n]{a}$ auch für negative a einen negativen Wert. Diesen Satz erweitern wir so:

IX. Die reine Gleichung

(6) $$x^n - a = 0$$

hat für jedes reelle oder komplexe a mindestens eine Wurzel.

Da wir, wie aus den Elementen bekannt, eine Quadratwurzel aus einer reellen oder imaginären Größe immer durch Quadratwurzeln aus reellen positiven Größen ausdrücken können, so genügt es, den Satz für ein ungerades n zu beweisen. Denn es ist

$$\sqrt[2n]{a} = \sqrt[2]{\sqrt[n]{a}}, \quad \sqrt[4n]{a} = \sqrt[2]{\sqrt[2n]{a}} \ \ldots$$

Es sei also die Gleichung gegeben:

(7) $$(y + iz)^n = b + ic.$$

Der Fall $c = 0$ ist durch das Vorstehende schon erledigt, und wenn c nicht $= 0$ ist, kann auch z nicht $= 0$ sein.

Aus der Gleichung (7) folgt aber auch die Gleichung

(8) $$(y - iz)^n = b - ic$$

und daraus

$$(y^2 + z^2)^n = b^2 + c^2$$

(9) $$\frac{(y+iz)^n(b-ic) - (y-iz)^n(b+ic)}{2i} = 0.$$

Aus der ersten dieser Gleichungen ist der absolute Wert von $x = y + iz$ bestimmt. Setzen wir dann

$$y = \lambda z,$$

so folgt aus der zweiten Gleichung (9):

$$\frac{(\lambda + i)^n (b - ic) - (\lambda - i)^n (b + ic)}{2i} = 0,$$

was eine reelle Gleichung $n$ten, also ungeraden Grades für λ ist, die jedenfalls eine reelle Wurzel hat. Hat man diese, so ergibt sich aus der ersten Gleichung (9):

$$z = \sqrt[2]{\frac{\sqrt[n]{b^2 + c^2}}{1 + \lambda^2}}, \qquad y = \lambda z,$$

wodurch IX. bewiesen ist.

Wir kehren zurück zu einer beliebigen ganzen reellen oder imaginären Funktion $f(x)$ und stellen zur Veranschaulichung das durch die Ungleichung

$$| x | < R$$

begrenzte Gebiet (R) für die Variable x durch eine Kreisfläche vom Radius R in der Ebene der Variablen $x = y + iz$ dar.

Wenn wir in VIII. C größer annehmen, als irgend einen Wert von $| f(x) |$ im Inneren des Gebietes (R), zum Beispiel größer als $| f(0) |$, so wird $| f(x) |$ gewiß im Inneren des Gebietes (R) kleiner werden, als an der Begrenzung. Da nun im ganzen Inneren von (R) die Funktion $| f(x) |$, die wir zur Abkürzung jetzt mit X bezeichnen wollen, nicht negativ wird, so gibt es für die Werte von X eine untere Grenze g und wir haben den Satz zu beweisen, der die Übertragung des Satzes VII auf eine Funktion von zwei Variablen ist:

X. Es existiert ein Wert ξ von x im Inneren des Gebietes (R), so daß

$$| f(\xi) | = g$$

wird, daß also die untere Grenze auch hier ein Minimum ist.

Die untere Grenze von X in irgend einem Teile des Bereiches ist entweder gleich g oder größer als g.

Ein Größengebiet, das durch die Ungleichheitsbedingungen

$$| x | \leqq R, \qquad -R \leqq y \leqq \alpha \tag{10}$$

bestimmt ist, wird in unserer Fig. 2 durch das Segment $(P Q \alpha Q')$, das wir das Segment (P, α) nennen wollen, dargestellt. Wir bestimmen nun zunächst einen Wert η von y durch einen Schnitt $\mathfrak{A} \mid \mathfrak{B}$ folgendermaßen:

Eine Zahl α zwischen $-R$ und $+R$ wird in $\mathfrak{A}$ aufgenommen, wenn die untere Grenze von X in dem Segment (P, α) größer als g ist, und ein Wert β wird in $\mathfrak{B}$ aufgenommen, wenn die untere Grenze von X in dem Segment (P, β) gleich g ist. Dieser Schnitt $\mathfrak{A} \mid \mathfrak{B}$ definiert eine Zahl η von der Eigenschaft, daß die untere Grenze von X in dem Bereich (P, y), wenn $y < \eta$ ist, größer als g, und wenn $y > \eta$ ist, gleich g ist.

Nun ist $| f(\eta + i\zeta) |$ als ganze Funktion von ζ eine stetige Funktion von ζ in dem Intervall

$$(11) \qquad -\sqrt{R^2 - \eta^2} \leqq \zeta \leqq \sqrt{R^2 - \eta^2},$$

das in der Fig. 2 durch M, M' bezeichnet ist.

Diese Funktion erhält also nach VIII. für einen Wert ξ von ζ in diesem Intervall einen Minimumwert γ, so daß, wenn

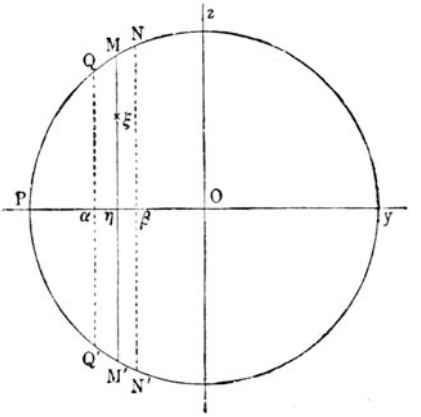
$$\xi = \eta + i\xi$$

gesetzt wird,

$$(12) \qquad | f(\xi) | = \gamma$$

wird.

Fig. 2.



Es ist nun ferner leicht einzusehen, daß $\gamma = g$ sein muß; denn da g die untere Grenze aller Werte von X innerhalb (R) ist, so kann γ zunächst nicht kleiner als g sein.

Es kann aber γ auch nicht größer als g sein; denn alle Werte, die X auf der Strecke MM' annimmt, sind $\geqq \gamma$; ist aber $\gamma > g$, so kann man eine Größe g' so annehmen, daß $\gamma > g' > g$ ist, und wegen der Stetigkeit von X lassen sich die Zahlen α, β so bestimmen, daß

$$\alpha < \eta < \beta$$

und daß in dem ganzen Bereich $(P, \beta) - (P, \alpha) = (\alpha, \beta)$ ($QNN'Q'$ in der Fig. 2) X größer als g' bleibt. Folglich ist die untere Grenze von X in (α, β) größer als g. Da nun die untere Grenze von X in (P, α) größer als g ist, so ist sie auch in (P, β), was aus (P, α) und (α, β) zusammengesetzt ist, größer als g; β gehört aber zu $\mathfrak{B}$, woraus ein Widerspruch mit der Definition von $\mathfrak{B}$ folgen würde.

Es bleibt also nur übrig, daß $\gamma = g$ ist, und der Satz X. ist damit nachgewiesen, nämlich, daß es einen Wert ξ im Inneren von (R) gibt, für den

$$|f(\xi)| = g$$

ist, daß also $|f(x)|$ einen Minimumwert erreicht. Wir beweisen nun:

XI. Wenn α irgend ein Wert von x ist, für den $f(\alpha)$ von Null verschieden ist, so läßt sich h so annehmen, daß

$$|f(\alpha + h)| < |f(\alpha)| \tag{13}$$

ausfällt.

Daraus folgt dann, daß, wenn $f(\xi)$ von Null verschieden wäre, g nicht das Minimum der Funktion $|f(x)|$ sein könnte; da dies aber bewiesen ist, so muß

$$f(\xi) = 0 \tag{14}$$

sein, und ξ ist eine Wurzel der Gleichung $f(x) = 0$. Der Fundamentalsatz wird also dann bewiesen sein.

Von den derivierten Funktionen $f'(\alpha), f''(\alpha)\ldots$ können einige verschwinden; die letzte $f^{(n)}(\alpha)$, die gleich $\Pi(n)a_0$ ist, ist aber von Null verschieden; es mögen also $f'(\alpha), f''(\alpha)\ldots f^{(m-1)}(\alpha)$ verschwinden, $f^{(m)}(\alpha)$ nicht verschwinden. Wir haben dann nach der Taylorschen Entwickelung:

$$f(\alpha+h) = f(\alpha) + \frac{h^m}{\Pi(m)} f^{(m)}(\alpha) + \frac{h^{m+1}}{\Pi(m+1)} f^{(m+1)}(\alpha) + \cdots \tag{15}$$

Wir wählen, was stets möglich ist, eine positive Zahl ε so, daß

$$\left| \frac{h}{m+1} \frac{f^{(m+1)}(\alpha)}{f^{(m)}(\alpha)} + \frac{h^2}{(m+1)(m+2)} \frac{f^{(m+2)}(\alpha)}{f^{(m)}(\alpha)} + \cdots \right| < 1 \tag{16}$$

ist, sobald

$$|h| < \varepsilon, \tag{17}$$

und einen positiven echten Bruch δ so, daß

$$\delta |f(\alpha)| < \left| \frac{\varepsilon^m f^{(m)}(\alpha)}{\Pi(m)} \right|, \tag{18}$$

was, wenn $f(\alpha)$ nicht verschwindet, gleichfalls möglich ist. Dann bestimmen wir (auf Grund von IX.) h aus der Gleichung:

$$\frac{h^m f^{(m)}(\alpha)}{\Pi(m)} = -\delta f(\alpha), \tag{19}$$

woraus nach (17)

$$|h| < \varepsilon$$

folgt. Aus (14) ergibt sich jetzt, wenn man für h^m den Wert aus (18) setzt,

$$f(\alpha + h) = (1 - \delta) f(\alpha)$$
$$- \delta f(\alpha) \left(\frac{h}{m+1} \frac{f^{(m+1)}(\alpha)}{f^{(m)}(\alpha)} + \frac{h^2}{(m+1)(m+2)} \frac{f^{(m+2)}(\alpha)}{f^{(m)}(\alpha)} \cdots \right),$$

also

$$|f(\alpha + h)| \gtreqless |f(\alpha)| (1 - \delta) \tag{20}$$
$$+ \delta \left| f(\alpha) \right| \left| \frac{h f^{(m+1)}(\alpha)}{(m+1) f^{(m)}(\alpha)} + \frac{h^2}{(m+1)(m+2)} \frac{f^{(m+2)}(\alpha)}{f^{(m)}(\alpha)} \cdots \right|,$$

und wegen (15)

$$|f(\alpha + h)| < |f(\alpha)| \tag{21}$$

w. z. b. w.

Damit ist also der Beweis des Fundamentalsatzes beendigt.

§ 33.

Stetigkeit der Wurzeln.

XII. Die Wurzeln einer algebraischen Gleichung sind stetige Funktionen der Koeffizienten.

Es ist zunächst die Bedeutung dieser Behauptung zu erklären. Es sei

$$f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots \tag{1}$$

eine ganze Funktion von x vom $n$ten Grade, und in lineare Faktoren zerlegt, sei

$$f(x) = (x - \alpha)^a (x - \beta)^b (x - \gamma)^c \ldots, \tag{2}$$

worin die $\alpha, \beta, \gamma \ldots$ voneinander verschiedene reelle oder komplexe Zahlen und $a, b, c \ldots$ positive ganze Zahlen sind.

Die Wurzeln $\alpha, \beta, \gamma \ldots$ werden sich mit den Koeffizienten $a_1, a_2, a_3 \ldots$ ändern; auch der Grad ihrer Vielfachheit kann ein anderer werden.

Wir bezeichnen die Änderungen von $a_1, a_2 \ldots$ mit $\varepsilon_1, \varepsilon_2 \ldots$ und setzen

$$\varphi(x) = \varepsilon_1 x^{n-1} + \varepsilon_2 x^{n-2} + \cdots \tag{3}$$
$$f(x) + \varphi(x) = f_1(x). \tag{4}$$

Wir umgeben die Punkte $\alpha, \beta, \gamma \ldots$ mit Gebieten von solcher Kleinheit, daß diese Gebiete sich gegenseitig ausschließen, etwa dadurch, daß wir die Punkte $\alpha, \beta, \gamma \ldots$ durch Kreisperipherien

mit den Radien ϱ, ϱ', ϱ''... umgeben, und bezeichnen diese Gebiete durch (ϱ), (ϱ'), (ϱ'') ...

Wenn die absoluten Werte von ε_1, ε_2 ... unter hinlänglich kleinen Werten liegen, so können wir von der Funktion $f_1(x)$ zunächst beweisen,

daß sie keine Wurzeln außerhalb der Gebiete (ϱ), (ϱ'), (ϱ'') ... hat,

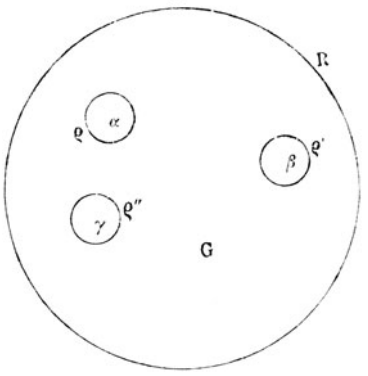
und zweitens,

daß die Anzahl der Wurzeln von $f_1(x)$ innerhalb (ϱ) genau a, innerhalb (ϱ') genau b, innerhalb (ϱ'') genau c usf. beträgt.

Bei dem letzten Teil des Satzes ist aber zu beachten, daß, wenn $f_1(x)$ mehrfache Wurzeln hat, diese nach ihrer Vielfachheit gezählt werden müssen.

Wir konstruieren nach VIII. in der Ebene x einen Kreis mit dem Radius R und dem Nullpunkt als Mittelpunkt, der die Gebiete (ϱ), (ϱ'), (ϱ'')... einschließt, so daß außerhalb dieses Kreises keine Wurzeln von $f_1(x)$ mehr liegen, und bezeichnen das innerhalb dieses Kreises, aber außerhalb (ϱ), (ϱ'), (ϱ'') ... liegende Gebiet mit G. Es ist dazu noch zu bemerken, daß R von den ε_1, ε_2 ... unabhängig angenommen werden kann, solange für die absoluten Werte dieser Größen eine bestimmte obere Grenze festgesetzt wird.

Fig. 3.



Ist nun x ein Punkt des Gebietes G, so ist der absolute Wert von $x - \alpha$ größer als ϱ, der von $x - \beta$ größer als ϱ' usf., und mithin nach (2)

$$|f(x)| > \varrho^a \varrho'^b \varrho''^c \ldots \tag{5}$$

Ist α_1 eine Wurzel von $f_1(x)$, so folgt aus (4):

$$f(\alpha_1) = -\varphi(\alpha_1), \tag{6}$$

und α_1 kann daher nicht in dem Gebiete G liegen, wenn man dafür sorgt, daß innerhalb G überall

$$|\varphi(x)| < \varrho^a \varrho'^b \varrho''^c \ldots, \tag{7}$$

was durch genügende Verkleinerung der oberen Grenze von $\varepsilon_1, \varepsilon_2 \ldots$ immer möglich ist. Hiermit ist der erste Teil unserer Behauptung erwiesen, daß, wenn die Ungleichung (7) befriedigt ist, keine Wurzeln von $f_1(x)$ außerhalb der Gebiete (ϱ), (ϱ'), $(\varrho'')\ldots$ liegen.

Um den zweiten Teil zu beweisen, setzen wir

$$f(x) = \psi(x)\,(x-\alpha)^a, \tag{8}$$

$$\psi(x) = (x-\beta)^b\,(x-\gamma)^c\ldots \tag{9}$$

Wir wählen nun eine positive Zahl A, die aber von ϱ unabhängig sein soll, so daß, solange x im Inneren oder an der Peripherie des Gebietes (ϱ) liegt,

$$|\psi(x)| > A, \qquad (\text{für } |x-\alpha| \leqq \varrho) \tag{10}$$

bleibt. Eine solche Zahl erhalten wir z. B., wenn wir l gleich der Hälfte des kleinsten unter den Abständen (α, β), $(\alpha, \gamma)\ldots$ annehmen und

$$A = l^{b+c+\cdots}$$

setzen; dann ist die in (10) ausgedrückte Forderung wenigstens so lange erfüllt, als ϱ kleiner als l ist. Ist nur ein einziger Punkt α vorhanden, so ist $\psi(x) = 1$ zu setzen, und für A kann jeder beliebige echte Bruch gesetzt werden.

Es seien nun $\alpha_1, \alpha_2 \ldots$ die Wurzeln von $f_1(x)$ innerhalb (ϱ), und $a_1, a_2 \ldots$ die Grade ihrer Vielfachheit, und $\beta_1, \beta_2 \ldots$, $b_1, b_2 \ldots$, $\gamma_1, \gamma_2 \ldots$, $c_1, c_2 \ldots$ sollen dieselbe Bedeutung für die Gebiete (ϱ'), $(\varrho'')\ldots$ haben; es ist dann

$$n = a_1 + a_2 + \cdots + b_1 + b_2 + \cdots + c_1 + c_2 + \cdots, \tag{11}$$

da die Gesamtzahl aller Wurzeln gleich n sein muß. Die Funktion $f_1(x)$ läßt sich dann so darstellen:

$$f_1(x) = \psi_1(x)\,(x-\alpha_1)^{a_1}\,(x-\alpha_2)^{a_2}\ldots, \tag{12}$$

worin

$$\begin{aligned}\psi_1(x) = {} & (x-\beta_1)^{b_1}\,(x-\beta_2)^{b_2}\ldots \\ & (x-\gamma_1)^{c_1}\,(x-\gamma_2)^{c_2}\ldots\end{aligned} \tag{13}$$

Nun können wir eine von ϱ, ϱ', ϱ'' unabhängige positive Zahl B bestimmen, so daß, solange x innerhalb oder an der Grenze von (ϱ) bleibt,

$$|\psi_1(x)| < B \qquad (\text{für } |x-\alpha| \leqq \varrho). \tag{14}$$

Wir können z. B. eine Größe L wählen, die größer ist als die doppelte Entfernung des Punktes α von einem der Punkte $\beta, \gamma \ldots$ und jedenfalls größer als 1 und dann

$$B > L^n$$

annehmen.

Nun folgt aus (4):

$$|f(x)| \leqq |f_1(x)| + |\varphi(x)|, \tag{15}$$

also nach (8) und (12):

$$|\psi(x)| \cdot |x-\alpha|^a \leqq |\psi_1(x)| \cdot |x-\alpha_1|^{a_1} \cdot |x-\alpha_2|^{a_2} \cdots + |\varphi(x)|, \tag{16}$$

und wenn wir nun x auf der Peripherie von (ϱ) annehmen, also

$$|x-\alpha| = \varrho$$

setzen, so ist

$$|x-\alpha_1| < 2\varrho, \qquad |x-\alpha_2| < 2\varrho \ldots,$$

folglich nach (10) und (16):

$$\varrho^a A < (2\varrho)^{a_1+a_2+\cdots} B + |\varphi(x)|. \tag{17}$$

Wir wollen nun die oberen Grenzen für die Koeffizienten $\varepsilon_1, \varepsilon_2, \ldots$ von φ so klein annehmen, daß

$$|\varphi(x)| < \varrho^a A'$$

wird, worin A' eine Zahl bedeutet, die kleiner als A ist; dann folgt aus (17):

$$A - A' < 2^{a_1+a_2+\cdots} \varrho^{-a+a_1+a_2+\cdots} B. \tag{18}$$

Dies aber würde für ein hinreichend kleines ϱ nicht mehr möglich sein, wenn a kleiner als die Summe $a_1 + a_2 + \cdots$ wäre. Es folgt also:

$$a \overline{\overline{\geqq}} a_1 + a_2 + \cdots \tag{19}$$

Ebenso läßt sich beweisen, daß

$$\begin{aligned} b &\overline{\overline{\geqq}} b_1 + b_2 + \cdots \\ c &\overline{\overline{\geqq}} c_1 + c_2 + \cdots \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{20}$$

sein muß. Da aber die Summen der linken Seiten sowohl als der rechten in den Ungleichungen (19), (20) gleich n sein müssen, so können nur die Gleichheitszeichen bestehen, also:

$$\begin{aligned} a &= a_1 + a_2 + \cdots \\ b &= b_1 + b_2 + \cdots \\ c &= c_1 + c_2 + \cdots \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

wodurch auch der zweite Teil unseres Satzes bewiesen ist.

Wir wollen dem bewiesenen Satze noch folgende, auf den Fall mehrfacher Wurzeln bezügliche Bemerkung beifügen:

Die notwendige und hinreichende Bedingung dafür, daß $f(x) = 0$ überhaupt mehrfache Wurzeln habe, ist die, daß $f(x)$ und $f'(x)$ einen gemeinsamen Faktor haben, daß also die Diskriminante von f verschwindet. Diese Diskriminante ist eine ganze Funktion $F(a_1, a_2 \dots a_n)$ der Koeffizienten $a_1, a_2 \dots a_n$, die, da es jedenfalls Gleichungen ohne mehrfache Wurzeln gibt, nicht identisch verschwindet. Man kann daher über die $\varepsilon_1, \varepsilon_2 \dots$, auch wenn für diese beliebig enge Grenzen festgesetzt sind, so verfügen, daß $F(a_1 + \varepsilon_1, a_2 + \varepsilon_2, \dots a_n + \varepsilon_n)$ von Null verschieden ist.

Man sieht also aus diesen Betrachtungen, wie eine a-fache Wurzel α von $f(x)$ bei stetiger Veränderung der Koeffizienten in a einfache Wurzeln auseinander strahlt, so daß man auch von diesem Gesichtspunkte berechtigt ist, die a-fache Wurzel als durch das Zusammenfallen von a einfachen Wurzeln entstanden zu betrachten.

Nehmen wir den Koeffizienten der höchsten Potenz von x nicht gleich 1 an, untersuchen also die Gleichung in der Form

$$(21) \quad a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0,$$

so sind, wenn $a_n = 0$ ist, eine oder mehrere der Wurzeln gleich Null, und die Gleichung (21) läßt sich durch x oder eine höhere Potenz von x dividieren, wodurch eine Gleichung von entsprechend niedrigerem Grade entsteht, in der das letzte Glied von Null verschieden ist. Wir wollen also von vornherein a_n von Null verschieden annehmen und nun in (21)

$$(22) \quad x = \frac{1}{y}$$

setzen. Durch Multiplikation mit y^n und Division mit a_n geht dann (21) über in:

$$(23) \quad y^n + \frac{a_{n-1}}{a_n} y^{n-1} + \cdots + \frac{a_1}{a_n} y + \frac{a_0}{a_n} = 0,$$

und wir können auf die Wurzeln dieser Gleichung den vorhin ausgesprochenen Satz anwenden, indem wir $a_0 = 0$ annehmen, so daß eine der Wurzeln von (23) verschwindet. Wir erlangen so den Satz:

Man kann in der Gleichung (21) a_0 so klein annehmen, daß eine ihrer Wurzeln über alle Grenzen groß wird, während die

anderen sich beliebig wenig von den Wurzeln der Gleichung $(n-1)$ten Grades:

$$a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0$$

unterscheiden.

Man drückt das auch so aus, daß mit verschwindendem a_0 eine der Wurzeln von (21) unendlich wird.

Wenn a_0 und a_1 oder a_0, a_1 und a_2 verschwinden, so werden zwei oder drei Wurzeln unendlich usf.

Es rechtfertigt sich hierdurch der bisweilen gebrauchte Ausdruck, daß eine Gleichung $n$ten Grades durch Unendlichwerden einer Wurzel in eine Gleichung $(n-1)$ten Grades übergehe.

Fünfter Abschnitt.

Kubische und biquadratische Gleichungen.

§ 34.

Die kubische Gleichung.

Man kann die kubische Gleichung

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0$$

zunächst auf eine einfachere Form bringen, indem man

$$x + \frac{a_1}{3} = y$$

setzt, wodurch man erhält:

(1) $$y^3 + ay + b = 0,$$

wenn

$$a = -\frac{a_1^2}{3} + a_2,$$

$$b = a_3 - \frac{a_1 a_2}{3} + \frac{2}{27} a_1^3$$

gesetzt wird.

Dieselbe Vereinfachung, nämlich daß das Glied mit der $(n-1)$ten Potenz der Unbekannten nicht mehr vorkommt, erreicht man bei einer Funktion $n$ter Ordnung $f(x)$ durch die Substitution $x + a_1/n = y$.

Führt man in (1) zwei neue Unbekannte u, v mittels der Gleichung

(2) $$y = u + v$$

ein, so erhält man:

(3) $$u^3 + v^3 + (3uv + a)(u + v) + b = 0.$$

Wir bestimmen eine der beiden Unbekannten u, v durch die andere mittels der Gleichung

(4) $$3uv = -a$$

und erhalten

(5) $$u^3 + v^3 = -b.$$

Aus (4) und (5) kann man u^3, v^3 durch eine Quadratwurzel bestimmen.

Man setze nämlich:

$$(u^3 - v^3)^2 = (u^3 + v^3)^2 - 4u^3v^3 \tag{6}$$

und folglich nach (4) und (5):

$$(u^3 - v^3)^2 = b^2 + \frac{4a^3}{27},$$

und wenn wir

$$R = \frac{b^2}{4} + \frac{a^3}{27} \tag{7}$$

setzen, so findet sich:

$$u^3 - v^3 = 2\sqrt{R},$$
$$u^3 + v^3 = -b,$$

also:

$$u^3 = -\frac{b}{2} + \sqrt{R}, \quad v^3 = -\frac{b}{2} - \sqrt{R}.$$

Wir geben dem $\sqrt{R}$ eines der beiden Zeichen und erhalten:

$$u = \sqrt[3]{-\frac{b}{2} + \sqrt{R}}, \quad v = \sqrt[3]{-\frac{b}{2} - \sqrt{R}},$$

und folglich erhält man die unter dem Namen der Cardanischen Formel bekannte Lösung der kubischen Gleichung:

$$y = \sqrt[3]{-\frac{b}{2} + \sqrt{R}} + \sqrt[3]{-\frac{b}{2} - \sqrt{R}}. \tag{8}$$

Die Multiplikation der beiden Ausdrücke u, v ergibt nach (4):

$$uv = \sqrt[3]{\frac{b^2}{4} - R} = -\frac{a}{3}, \tag{9}$$

wodurch eine der beiden Größen u, v durch die andere bestimmt ist.

Wenn R positiv ist, so ist $\sqrt{R}$ reell und kann positiv oder negativ genommen werden, wodurch u mit v vertauscht wird. u und v und folglich y haben reelle Werte. Ist aber R negativ, so sind u und v konjugiert imaginär und y ist wieder reell, kann aber nicht in reeller Form algebraisch dargestellt werden. Darum wird dieser Fall von den älteren Algebraikern der Casus irreducibilis genannt, auf den wir später zurückkommen.

Nun hat aber die kubische Gleichung nicht bloß eine, sondern drei Wurzeln. Um aus der einen die beiden anderen zu finden, setzen wir

$$\alpha = u + v$$

und dividieren $y^3 + ay + b$ durch $y - \alpha$. Der Quotient, dessen Wurzeln die beiden Wurzeln β, γ der kubischen Gleichung (1) sind, ist

$$y^2 + \alpha y + \alpha^2 + a = 0,$$

wovon man die Wurzeln erhält:

$$\beta, \gamma = \frac{-\alpha \pm \sqrt{-3\alpha^2 - 4a}}{2}.$$

Setzt man hierin $\alpha = u + v$, $a = -3uv$, ferner

$$\varepsilon = \frac{-1 + \sqrt{-3}}{2},$$

so folgt:

$$(10)\quad \varepsilon^2 = \frac{-1 - \sqrt{-3}}{2}, \quad \varepsilon + \varepsilon^2 = -1, \quad \varepsilon - \varepsilon^2 = \sqrt{-3}.$$

$$(11)\qquad \begin{aligned} \alpha &= u + v \\ \beta &= \varepsilon u + \varepsilon^2 v \\ \gamma &= \varepsilon^2 u + \varepsilon v. \end{aligned}$$

Setzt man $a = 0$ und $b = -1$, so bekommt man aus (10) die drei Wurzeln der Gleichung $y^3 - 1 = 0$. Es wird $R = \frac{1}{4}$, $\sqrt{R} = \frac{1}{2}$, $u = 0$, $v = 1$, also

$$\beta = \varepsilon^2, \quad \gamma = \varepsilon.$$

Es sind also ε und ε^2 die beiden imaginären dritten Wurzeln der Einheit, während 1 die reelle dritte ist.

Sind u und v reell, so ist α reell, β und γ sind aber konjugiert imaginär. Im Falle eines negativen R ist u zu v konjugiert imaginär und folglich α reell. Es sind aber auch εu und $\varepsilon^2 v$ konjugiert imaginär und folglich β gleichfalls reell, und ebenso ergibt sich in diesem Falle ein reeller Wert für γ.

Das Differenzenprodukt

$$\Delta = (\alpha - \beta)^2 (\alpha - \gamma)^2 (\beta - \gamma)^2$$

ist die Diskriminante der kubischen Gleichung. Setzen wir

$$\sqrt{\Delta} = (\alpha - \beta)(\alpha - \gamma)(\beta - \gamma),$$

so ist

$$\sqrt{\Delta} = 3\sqrt{-3}\,(u^3 - v^3)$$

und folglich

$$(12)\qquad \Delta = -4.27R = -27b^2 - 4a^3.$$

Wenn man in dem allgemeinen Ausdruck der Cardanischen Formel $u + v$ allen vorkommenden Wurzeln, den beiden Quadrat-

wurzeln und den beiden Kubikwurzeln alle möglichen Werte beilegt, so bekommt man nicht drei, sondern 36 verschiedene Werte, die sich erst durch die in (9) ausgedrückte Forderung und durch die zweite Forderung, daß $\sqrt{R}$ in u und v dasselbe Vorzeichen haben sollen, auf drei reduzieren.

Cayley hat darum der Cardanischen Formel einen Ausdruck gegeben, der von vornherein nur drei Werte enthält. Er setzt:

(13) $$u = \xi^2\eta, \quad v = \xi\eta^2,$$

woraus nach (4):

$$\xi\eta = \sqrt[3]{uv} = \sqrt[3]{\frac{-a}{3}}.$$

Dann folgt:

(14)
$$\xi = \sqrt[3]{\frac{3b}{2a} + \sqrt{\frac{9b^2}{4a^2} + \frac{a}{3}}},$$
$$\eta = \sqrt[3]{\frac{3b}{2a} - \sqrt{\frac{9b^2}{4a^2} + \frac{a}{3}}},$$

(15) $$y = \xi\eta(\xi + \eta),$$

und dieser Ausdruck erhält nur drei Werte, wenn man jeder der beiden Kubikwurzeln, unabhängig von der anderen, einen der drei Werte beilegt. Freilich ist auch hier daran festzuhalten, daß die Quadratwurzel in ξ und η dasselbe Zeichen bekommt.

§ 35.

Die biquadratische Gleichung.

Die biquadratische Gleichung oder Gleichung vierten Grades wird durch die Reduktion auf eine kubische Gleichung (die kubische Resolvente) gelöst. Der älteste Weg zur Herstellung einer solchen Resolvente ist der von Ferrari. Man nimmt die biquadratische Gleichung zunächst auch in der vereinfachten Form an:

(1) $$y^4 + ay^2 + by + c = 0$$

und setzt, ähnlich wie in der Cardanischen Formel:

(2) $$2y = u + v + w.$$

Setzt man zur Abkürzung

$$s = u^2 + v^2 + w^2,$$
$$t = v^2w^2 + w^2u^2 + u^2v^2,$$

so kommt:

$$\begin{aligned} 4y^2 &= s + 2(vw + wu + uv) \\ 16y^4 &= s^2 + 4s(vw + wu + uv) + 4t \\ &\quad + 8uvw(u + v + w), \end{aligned}$$

und danach ergibt sich aus (1):

$$\begin{aligned} &s^2 + 4t + 4as + 16c \\ &+ 8(uvw + b)(u + v + w) \\ &+ 4(s + 2a)(vw + wu + uv) = 0. \end{aligned}$$

Diese Gleichung ist erfüllt, wenn wir setzen:

$$\begin{aligned} s + 2a &= 0, \\ uvw + b &= 0, \\ s^2 + 4t + 4as + 16c &= 0, \end{aligned}$$

und die dritte dieser Gleichungen wird mit Hilfe der ersten

$$t = a^2 - 4c.$$

Demnach erhalten wir:

$$\begin{aligned} u^2 + v^2 + w^2 &= -2a, \\ v^2w^2 + w^2u^2 + u^2v^2 &= a^2 - 4c, \\ uvw &= -b. \end{aligned} \tag{3}$$

Daraus ergibt sich als kubische Resolvente für u^2, v^2, w^2:

$$z^3 + 2az^2 + (a^2 - 4c)z - b^2 = 0. \tag{4}$$

Die Quadratwurzeln aus den Wurzeln dieser kubischen Gleichung haben je zwei Werte. Die Vorzeichen sind aber durch die letzte der Gleichungen (3) miteinander verbunden, und man erhält daher folgende Ausdrücke für die Wurzeln der biquadratischen Gleichung:

$$\begin{aligned} 2\alpha &= u + v + w \\ 2\beta &= u - v - w \\ 2\gamma &= -u + v - w \\ 2\delta &= -u - v + w. \end{aligned} \tag{5}$$

Die Wurzel der kubischen Resolvente ist, durch die Wurzeln der biquadratischen Gleichung ausgedrückt:

$$u^2 = \tfrac{1}{4}(\alpha + \beta - \gamma - \delta)^2. \tag{6}$$

Man kann noch andere kubische Resolventen der biquadratischen Gleichung aufstellen, wobei wir diese gleich in der allgemeinen Form

$$x^4 - a_1x^3 + a_2x^2 - a_3x + a_4 = 0 \tag{7}$$

annehmen wollen. Setzt man für die linke Seite das Produkt

$$(x - \alpha)(x - \beta)(x - \gamma)(x - \delta),$$

so erhält man:

$$\begin{aligned} a_1 &= \Sigma\alpha \\ a_2 &= \Sigma\alpha\beta \\ a_3 &= \Sigma\alpha\beta\gamma \\ a_4 &= \alpha\beta\gamma\delta. \end{aligned} \tag{8}$$

Setzen wir

$$\begin{aligned} u &= \alpha\beta + \gamma\delta \\ u_1 &= \alpha\gamma + \beta\delta \\ u_2 &= \alpha\delta + \beta\gamma, \end{aligned} \tag{9}$$

so kann man, wenn u, u_1, u_2 als bekannt vorausgesetzt werden, die α, β, γ, δ folgendermaßen durch quadratische Gleichungen bestimmen.

Man bestimme aus

$$u = \alpha\beta + \gamma\delta, \quad \sqrt{u^2 - 4a_4} = \alpha\beta - \gamma\delta$$

$\alpha\beta$ und $\gamma\delta$ durch eine Quadratwurzel und ebenso $\alpha\gamma$ und $\beta\delta$ aus

$$u_1 = \alpha\gamma + \beta\delta, \quad \sqrt{u_1^2 - 4a_4} = \alpha\gamma - \beta\delta.$$

Welches Zeichen man den beiden Quadratwurzeln nun gibt, ist willkürlich und kommt auf eine Vertauschung der Wurzeln α, β, γ, δ hinaus. Die dritte, durch u_2 zu bestimmende Quadratwurzel ist nicht mehr willkürlich. Hat man aber die vier Produkte:

$$\alpha\beta, \quad \gamma\delta, \quad \alpha\gamma, \quad \beta\delta,$$

so kann man durch $\alpha\beta$ und $\gamma\delta$ aus den linearen Gleichungen

$$\begin{aligned} \alpha\beta(\gamma + \delta) + \gamma\delta(\alpha + \beta) &= a_3 \\ (\gamma + \delta) + (\alpha + \beta) &= a_1 \end{aligned}$$

die Unbekannten

$$(\gamma + \delta), \ (\alpha + \beta)$$

bestimmen, und ebenso aus den beiden anderen Produkten $\alpha\gamma$ und $\beta\delta$

$$(\alpha + \gamma), \ (\beta + \delta).$$

Aus $\alpha + \beta$ und $\alpha\beta$ kann man durch eine nochmalige Quadratwurzel α und β bestimmen, wodurch zugleich γ und δ eindeutig bestimmt sind.

Es bleibt noch übrig, die kubische Gleichung zu bilden, deren Wurzeln u, u_1, u_2 sind. Deren Koeffizienten sind die symmetrischen Grundfunktionen:

$$\begin{aligned}\Sigma u &= \Sigma\alpha\beta = a_2\\ \Sigma u u_1 &= \Sigma\alpha^2\beta\gamma = \Sigma\alpha\Sigma\alpha\beta\gamma - 4\alpha\beta\gamma\delta\\ &= a_1 a_3 - 4a_4\\ u u_1 u_2 &= \Sigma\alpha^3\beta\gamma\delta + \Sigma\alpha^2\beta^2\gamma^2\\ &= \alpha\beta\gamma\delta\Sigma\alpha^2 + \Sigma\alpha^2\beta^2\gamma^2.\end{aligned}$$

Um die beiden Summen $\Sigma\alpha^2$, $\Sigma\alpha^2\beta^2\gamma^2$ zu erhalten, bilde man aus (7) die Gleichung, deren Wurzeln ξ die Quadrate von α, β, γ, δ sind. Man setzt zu diesem Zweck in (7) $x = \sqrt{\xi}$ und schafft die Quadratwurzel durch Quadrierung heraus. Man erhält so:

$$(\xi^2 + a_2\xi + a_4)^2 - \xi(a_1\xi + a_3)^2 = 0,$$

und daraus findet man nach den Formeln (8):

$$\Sigma\alpha^2 = a_1^2 - 2a_2, \qquad \Sigma\alpha^2\beta^2\gamma^2 = a_3^2 - 2a_2a_4,$$

und folglich

$$u u_1 u_2 = a_1^2 a_4 - 4a_2a_4 + a_3^2,$$

und die kubische Resolvente für u lautet also:

$$u^3 - a_2u^2 + (a_1a_3 - 4a_4)u - (a_1^2a_4 - 4a_2a_4 + a_3^2) = 0 \tag{10}$$

Um das zweite Glied wegzuschaffen, substituiere man

$$3u = w + a_2 \tag{11}$$

und erhält für w die kubische Gleichung:

$$w^3 - 3Aw + B = 0, \tag{12}$$

deren Koeffizienten

$$\begin{aligned}A &= a_2^2 - 3a_1a_3 + 12a_4,\\ B &= 2a_2^3 + 27a_3^2 - 9a_1a_2a_3 + 27a_1^2a_4 - 72a_2a_4\end{aligned} \tag{13}$$

die erste und zweite Invariante der biquadratischen Form heißen.

Aus (9) und (11) ergibt sich:

$$w = (\alpha - \gamma)(\beta - \delta) + (\alpha - \delta)(\beta - \gamma),$$

und wenn man

$$\begin{aligned}U &= (\alpha - \beta)(\delta - \gamma)\\ U_1 &= (\alpha - \gamma)(\beta - \delta)\\ U_2 &= (\alpha - \delta)(\gamma - \beta)\end{aligned} \tag{14}$$

setzt, so ist

$$\begin{aligned}w &= U_1 - U_2, & 3U &= w_2 - w_1\\ w_1 &= U_2 - U, & 3U_1 &= w - w_2\\ w_2 &= U - U_1, & 3U_2 &= w_1 - w\end{aligned} \tag{15}$$

und die U sind selbst die Wurzeln einer kubischen Resolvente Es ist hier:

$$U + U_1 + U_2 = 0$$
$$U U_1 + U U_2 + U_1 U_2 = \tfrac{1}{3} \Sigma w_1 w_2 = -A$$
$$U U_1 U_2 = \sqrt{D},$$

wenn

$$D = (\alpha - \beta)^2 (\alpha - \gamma)^2 (\alpha - \delta)^2 (\beta - \gamma)^2 (\beta - \delta)^2 (\gamma - \delta)^2$$

die Diskriminante der biquadratischen Gleichung ist. Also ist die gesuchte Resolvente:

$$U^3 - A U + \sqrt{D} = 0. \tag{16}$$

Nach (15) ist

$$27^2 D = (w_2 - w_1)^2 (w - w_2)^2 (w_1 - w_2)^2$$

die Diskriminante der kubischen Gleichung (12), und diese ist nach § 32

$$+ 4 . 27 A^3 - 27 B^2,$$

woraus folgt:

$$27 D = 4 A^3 - B^2, \tag{17}$$

und die Diskriminante der biquadratischen Gleichung ist also rational durch die beiden Invarianten ausgedrückt. In dem langen Ausdruck, der sich durch Substitution von (13) daraus ergibt, hebt sich der Faktor 27 noch heraus.

Sechster Abschnitt.

Der Sturmsche Lehrsatz.

§ 36.

Das Sturmsche Problem.

Für eine numerische Berechnung der Wurzeln einer Gleichung ist zunächst immer die Frage zu beantworten: Wie viele reelle Wurzeln einer reellen numerischen Gleichung liegen zwischen zwei gegebenen reellen Zahlwerten a und b?

Ist dies entschieden, so handelt es sich weiter darum, die Grenzen a, b so weit einzuengen, daß nur noch eine Wurzel der gegebenen Gleichung zwischen ihnen liegt, und sie endlich einander so weit zu nähern, daß jede von ihnen als ein genäherter Wert dieser Wurzel betrachtet werden kann.

Nehmen wir $a = -\infty$, $b = +\infty$, oder doch a negativ, b positiv so groß an, daß jenseits dieser Grenzen keine Wurzeln mehr liegen können, so wird damit die Frage nach der Anzahl der reellen Wurzeln beantwortet, und darauf läßt sich auch das allgemeine Problem zurückführen.

Es möge sich zunächst darum handeln, die Anzahl der positiven Wurzeln einer Gleichung $f(x) = 0$ zu ermitteln. Setzen wir $x = y^2$, so wird jedem reellen Wert von y ein positiver Wert von x entsprechen, und umgekehrt entsprechen jedem positiven Wert von x zwei reelle, entgegengesetzte Werte von y. Die Anzahl der positiven Wurzeln von $f(x)$ ist also halb so groß, als die Anzahl der reellen Wurzeln von $f(y^2)$. Setzen wir ferner

$$y = \frac{x - a}{b - x},$$

so wird, während x von a bis b geht, y durch positive Werte von Null bis Unendlich gehen, und wenn wir durch diese Substitution $f(x)$ in $F(y)$ transformieren, so wird $f(x)$ ebenso viele Wurzeln

zwischen a und b haben, als $F(y)$ positive Wurzeln hat, mithin halb so viel, als $F(y^2)$ reelle Wurzeln hat. Unsere Aufgabe ist also dadurch in der Tat auf die Ermittelung der Anzahl der reellen Wurzeln einer gewissen anderen Gleichung zurückgeführt, deren Grad doppelt so groß ist, als der Grad der gegebenen Gleichung. Diese Zurückführung des Problems gibt aber seine Lösung nicht in der einfachsten Form, und wir müssen nach einer einfacheren Beantwortung der Frage suchen. Dazu führt die folgende Betrachtung:

Es seien α und β irgend zwei reelle Zahlen und $\alpha < \beta$. Es sei ferner $f(x)$ eine gegebene reelle ganze Funktion von x, von der ermittelt werden soll, wie viele ihrer Wurzeln in dem Intervall Δ

$$\alpha < x < \beta$$

liegen. Wir nehmen an, daß α, β nicht selbst zu den Wurzeln von $f(x) = 0$ gehören, daß also $f(\alpha)$ und $f(\beta)$ von Null verschieden sind. Außerdem wollen wir noch fürs erste annehmen, daß $f(x)$, wenigstens in dem Intervall Δ, keine mehrfache Wurzel habe, daß also für keinen Wert des Intervalls $f(x)$ und $f'(x)$ zugleich verschwinden sollen.

Wir nehmen an, daß wir auf irgend eine Weise eine Reihe von $m + 1$ reellen stetigen Funktionen herstellen können:

$$f(x),\ f_1(x),\ f_2(x),\ \ldots\ f_m(x), \qquad (\mathfrak{R})$$

denen folgende Eigenschaften zukommen:

1. Von den Funktionen f_ν sollen im Intervall Δ nicht zwei aufeinander folgende zugleich verschwinden.
2. Die letzte von ihnen, $f_m(x)$, soll im Intervall Δ überhaupt nicht verschwinden, also ein unveränderliches Zeichen behalten.
3. Wenn ein mittleres Glied, etwa $f_\nu(x)$, für irgend ein x im Intervall Δ verschwindet, so sollen für dieses x die beiden angrenzenden Funktionen $f_{\nu-1}(x)$ und $f_{\nu+1}(x)$ entgegengesetztes Vorzeichen haben.
4. Wenn $f(x)$ im Intervall Δ verschwindet, so soll $f_1(x)$ für diesen Wert von x dasselbe Vorzeichen haben wie $f'(x)$.

Eine solche Funktionenreihe $(\mathfrak{R})$ wollen wir eine Sturmsche Kette nennen. Wie man sie bilden kann, werden wir später

sehen; zunächst sollen aus der Definition Folgerungen gezogen werden.

Für jeden Wert von x, für den keine der Funktionen von ($\mathfrak{R}$) verschwindet, hat jede dieser Funktionen ein bestimmtes Vorzeichen. Wir zählen einen Zeichenwechsel, so oft beim Durchlaufen der Kette von links nach rechts auf ein positives Glied ein negatives oder auf ein negatives Glied ein positives folgt. Die Anzahl der so gezählten Zeichenwechsel (Variationen) für einen bestimmten Wert von x wollen wir mit $V(x)$ bezeichnen. Wenn ein mittleres Glied $f_\nu(x)$ verschwindet, so haben wir nach 3. beim Übergang von $f_{\nu-1}(x)$ zu $f_{\nu+1}(x)$ einen Zeichenwechsel zu zählen.

Sind α, β zwei Werte, für die keine der Funktionen f_ν verschwindet, und lassen wir nun x stetig wachsen von α bis β, so wird eine Änderung in der Zahl der Zeichenwechsel nur dann eintreten können, wenn eine der Funktionen von ($\mathfrak{R}$) ihr Zeichen wechselt, also durch den Wert Null hindurchgeht (wegen der vorausgesetzten Stetigkeit).

Ist dies aber eine mittlere Funktion, so ändert sich die Zahl der Zeichenwechsel nicht (nach 3.). Denn in

$$f_{\nu-1},\quad f_\nu,\quad f_{\nu+1}$$

findet vor und nach dem Durchgang von f_ν durch Null immer ein Zeichenwechsel statt, weil $f_{\nu+1}$ und $f_{\nu-1}$ entgegengesetzte Zeichen haben.

Wenn aber die erste Funktion $f(x)$ durch Null geht, so geht beim Durchgang wegen 4. ein Zeichenwechsel verloren. Denn geht $f(x)$ von negativen zu positiven Werten, so ist $f'(x)$ und mithin $f_1(x)$ positiv und die Vorzeichen ändern sich so:

$$\begin{array}{cc} f(x), & f_1(x) \\ - & + \\ + & +, \end{array}$$

und wenn $f(x)$ von positiven zu negativen Werten übergeht, so ist $f'(x)$ und $f_1(x)$ negativ, also die Zeichen so:

$$\begin{array}{cc} f(x), & f_1(x) \\ + & - \\ - & -, \end{array}$$

mithin ist in beiden Fällen ein Zeichenwechsel verloren gegangen.

Da nun $f_m(x)$ nach 2. sein Zeichen nicht wechselt, so folgt, daß $V(\alpha) - V(\beta)$ gleich der Anzahl der zwischen α und β gelegenen Wurzeln von $f(x) = 0$ ist, oder:

I. Die Anzahl der Wurzeln von $f(x) = 0$ zwischen α und β ist gleich dem Überschuß der Anzahl der Zeichenwechsel der Kette für $x = \alpha$ über die Anzahl der Zeichenwechsel für $x = \beta$.

Wenn für $x = \alpha$ oder $x = \beta$ eine oder einige der mittleren Funktionen von $(\mathfrak{K})$ verschwinden sollten, so ist es wegen 3. gleichgültig, ob wir diesen verschwindenden Wert durch einen positiven oder einen negativen ersetzen.

Um ein Beispiel zu geben, sei $a_{i,k}$, wenn i und k die Reihe der Indices 1, 2, 3, ... n durchlaufen, irgend ein System reeller Größen und

$$a_{i,k} = a_{k,i} \tag{1}$$

vorausgesetzt. Wir betrachten die symmetrische Determinante

$$L_n(x) = \begin{vmatrix} a_{1,1} - x, & a_{1,2} & \dots & a_{1,n} \\ a_{2,1}, & a_{2,2} - x & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1}, & a_{n,2} & \dots & a_{n,n} - x \end{vmatrix}, \tag{2}$$

die eine ganze rationale Funktion $n$ten Grades von x ist, in der x^n den Koeffizienten $(-1)^n$ hat. Die Gleichung $L_n(x) = 0$ soll der Gegenstand unserer Betrachtung sein.

Wir bilden die mit abwechselnden Zeichen genommene Kette der Haupt-Unterdeterminanten

$$L_n(x),\ -L_{n-1}(x),\ L_{n-2}(x) \dots,\ (-1)^{n-1} L_1(x),\ (-1)^n, \tag{3}$$

und wollen nun nachweisen, daß, wenn wir die Voraussetzung hinzufügen, daß von den Größen (3) keine zwei aufeinander folgende zugleich verschwinden, wir eine Sturmsche Kette vor uns haben.

Es ist dies eine einfache Folgerung aus der Formel, die wir schon im § 12 zu einem ähnlichen Zwecke benutzt haben, und die wir so darstellen können:

$$L_k S_k - T_k^2 = L_{k-1} L_{k+1}, \tag{4}$$

worin S_k und T_k ganze rationale Funktionen von x sind.

Wir haben nur zu zeigen, daß die Forderungen 1. bis 4. befriedigt sind. Davon ist aber 1. in die Voraussetzung aufgenommen, 2. ist erfüllt, da das letzte Element gleich ± 1 ist.

Aus der Formel (4) folgt, daß, wenn $L_k = 0$ ist, L_{k-1} und L_{k+1} entgegengesetzte Zeichen haben, daß also die Bedingung 3. erfüllt ist.

Es bleibt nur noch die Bedingung 4. übrig. Wenden wir aber die Formel (4) auf $k = n - 1$ an, so lautet sie:

$$\frac{\partial L_n}{\partial a_{n,n}} \frac{\partial L_n}{\partial a_{n-1,n-1}} - \left(\frac{\partial L_n}{\partial a_{n,n-1}}\right)^2 = L_n L_{n-2},$$

und daraus folgt, daß, wenn $L_n = 0$ ist,

$$\frac{\partial L_n}{\partial a_{n,n}}, \qquad \frac{\partial L_n}{\partial a_{n-1,n-1}}$$

gleiche Zeichen haben; und ebenso kann man schließen, daß alle Haupt-Unterdeterminanten von L_n

$$\frac{\partial L_n}{\partial a_{i,i}}$$

dasselbe Zeichen haben.

Nun ist aber die Derivierte von L_n

$$L_n'(x) = -\sum_{1,n}^{i} \frac{\partial L_n}{\partial a_{i,i}},$$

und hat also für einen Wert x, für den $L_n(x)$ verschwindet, das entgegengesetzte Zeichen, wie L_{n-1}, was eben die Forderung 4. verlangt. Daraus folgt der Satz:

II. Sind α, β zwei reelle Werte, $\alpha < \beta$, so ist die Anzahl der zwischen α und β gelegenen Wurzeln von $L_n(x)$ gleich dem Überschuß der Anzahl der Zeichenwechsel der Kette (3) für $x = \alpha$ über die Anzahl der Zeichenwechsel für $x = \beta$.

Die höchste Potenz von x, die in $L_k(x)$ vorkommt, ist

$$(-1)^k x^k,$$

und wenn der absolute Wert von x hinlänglich groß ist, so wird das Vorzeichen dieses Gliedes über das Vorzeichen von $L_k(x)$ entscheiden. Nehmen wir also für α einen genügend großen negativen, für β einen genügend großen positiven Wert, so finden in (3) für $x = \alpha$ lauter Zeichenwechsel, für $x = \beta$ lauter Zeichenfolgen statt. Es werden also n Wurzeln zwischen α und β liegen, und daraus folgt der Satz:

III. Die Gleichung $L_n(x) = 0$ hat lauter reelle Wurzeln.

Diese beiden Sätze sind aber nur von beschränkter Anwendbarkeit, solange wir uns nicht von der Voraussetzung frei machen können, daß in der Kette der L_k nicht zwei aufeinander folgende Glieder zugleich verschwinden sollen. Von dieser Beschränkung können wir den Satz aber durch folgende einfache Überlegung befreien.

Nehmen wir an, für irgend einen Wert von x verschwinde L_k, aber nicht L_{k-1}; wir können dann die Größen $a_{k+1,i}$, die in L_k nicht vorkommen, so bestimmen, daß L_{k+1} für diesen Wert von x nicht verschwindet; denn es kann L_{k+1} nicht identisch für alle $a_{k+1,i}$ verschwinden, weil es das Glied

$$-a_{k+1,k}^2 L_{k-1}$$

enthält (nach dem auf L_{k+1} angewandten Laplaceschen Determinantensatz § 3, XIV).

Hiernach können wir, wenn in der Kette

$$L_n, \quad -L_{n-1}, \quad L_{n-2} \dots \pm L_1, \quad \mp 1 \tag{5}$$

einige aufeinander folgende Glieder für irgend einen Wert von x verschwinden, durch Abänderung der $a_{i,k}$ eine andere Reihe

$$L_n', \quad -L_{n-1}', \quad L_{n-2}' \cdots \pm L_1', \quad \mp 1 \tag{6}$$

ableiten, in der keine zwei aufeinander folgenden Glieder für irgend einen Wert x (des Intervalls $\alpha \dots \beta$) verschwinden.

Zugleich können die $a_{i,k}'$, von denen die L' abhängen, so angenommen werden, daß sie sich von den $a_{i,k}$ um weniger als eine beliebig gegebene Größe ω unterscheiden.

Wenn nun die Zahlen α, β so angenommen sind, daß in der Reihe (5) kein Glied für $x = \alpha$ oder $x = \beta$ verschwindet, so können wir ω so klein annehmen, daß für $x = \alpha$ und $x = \beta$ entsprechende Glieder von (5) und (6) dasselbe Vorzeichen haben. Durch unseren Satz ist aber die Anzahl der Wurzeln von $L_n' = 0$ zwischen α und β durch die Zeichen der Reihe (6) bestimmt.

Nun läßt sich andererseits wieder ω so klein annehmen, daß die Wurzeln von $L_n' = 0$ von denen von $L_n = 0$ beliebig wenig unterschieden sind.

Es kann zwar eine Doppelwurzel von $L_n = 0$ in zwei einfache Wurzeln von $L_n' = 0$ übergehen; aber da $L_n' = 0$ keine imaginären Wurzeln hat, so sind diese reell; und dasselbe findet statt, wenn $L_n = 0$ mehrfache Wurzeln hat.

Hiernach behalten die Sätze II und III ihre Gültigkeit, wenn die Voraussetzung aufgegeben wird, daß in

der Kette der L_k keine aufeinander folgenden Glieder verschwinden; nur müssen die mehrfachen Wurzeln dabei nach ihrer Vielfachheit gezählt werden.

Hieraus können wir eine merkwürdige Beziehung der Gleichung $L_n = 0$ zu dem Trägheitsgesetz der quadratischen Formen herleiten, die in der Geometrie, bei der Bestimmung der Hauptachsen einer Fläche zweiten Grades wohl bekannt ist.

Wir betrachten neben der Funktion $L_n(x)$ die quadratische Form

$$\varphi(x_1, x_2 \dots x_n) = \Sigma a_{i,k} x_i x_k; \tag{7}$$

dann ist $L_n(\varepsilon)$ für ein beliebiges ε die Determinante der Form

$$\varphi' = \varphi - \varepsilon(x_1^2 + x_2^2 \cdots + x_n^2). \tag{8}$$

Die Gleichung $L_n(x) = 0$ möge P positive, N negative und R verschwindende Wurzeln haben. Wir nehmen ε positiv an, aber so klein, daß alle positiven Wurzeln von $L_n(x)$ größer als ε sind, und außerdem so, daß kein Glied der Reihe (3) für $x = \varepsilon$ verschwindet. Dann ist

$$L_n(\varepsilon),\ L_{n-1}(\varepsilon),\ \dots\ L_1(\varepsilon),\ 1 \tag{9}$$

eine Kette von Haupt-Unterdeterminanten für die Funktion φ', und wenn wir also mit π', ν', ϱ' die charakteristischen Zahlen von φ' bezeichnen, so ist π' (nach § 12) gleich der Anzahl der Zeichenfolgen, ν' gleich der Anzahl der Zeichenwechsel in der Kette (9) und $\varrho' = 0$, weil $L_n(\varepsilon)$ nicht verschwindet.

Wenn wir aber das Theorem II. anwenden, indem wir $\alpha = \varepsilon$ setzen, und β größer als alle Wurzeln von $L_n(x)$ annehmen, so daß in der Kette (3) für $x = \beta$ nur noch Zeichenfolgen vorkommen, so gibt uns die Anzahl der Zeichenwechsel in (3) für $x = \varepsilon$, die mit der Anzahl der Zeichenfolgen in (9) übereinstimmt, die Zahl P der positiven Wurzeln von $L_n(x) = 0$. Demnach ist, weil außerdem noch $\pi' + \nu' = P + N + R = n$ ist,

$$\pi' = P, \quad \nu' = N + R. \tag{10}$$

Wir nehmen nun die Funktionen φ und φ' in folgender Weise in ihre positiven und negativen Quadrate zerlegt an:

$$\varphi = y_1^2 + y_2^2 + \cdots + y_\pi^2 - z_1^2 - z_2^2 - \cdots - z_\nu^2. \tag{11}$$

$$\begin{aligned} \varphi' &= \varphi - \varepsilon(x_1^2 + x_2^2 + \cdots + x_n^2) \\ &= Y_1^2 + Y_2^2 + \cdots + Y_{\pi'}^2 - Z_1^2 - Z_2^2 \cdots - Z_{\nu'}^2. \end{aligned} \tag{12}$$

Wenn nun $\pi + \nu' < n$ ist, so können wir die $x_1, x_2 \ldots x_n$, zum Teil wenigstens von Null verschieden, so bestimmen, daß

$$y_1 = 0,\ y_2 = 0 \ldots,\ y_\pi = 0,\ Z_1 = 0,\ Z_2 = 0 \ldots,\ Z_{\nu'} = 0.$$

Dann aber ergibt die erste Darstellung (12) mit Hilfe von (11) einen negativen, die zweite einen positiven (oder verschwindenden) Wert von φ', so daß die Annahme $\pi + \nu' < n$ unstatthaft ist. Es ist also $\pi + \nu' \gtreqless n = \pi' + \nu'$, also nach (10):

$$\pi \gtreqless P.$$

Andererseits folgt aber aus § 12, (14)

$$P \gtreqless \pi,$$

was nur miteinander verträglich ist, wenn

$$\pi = P \tag{13}$$

ist. Ebenso können wir nun auch beweisen, indem wir statt eines positiven ein negatives ε zu Hilfe nehmen, daß $\nu = N$ und folglich $\varrho = R$ sein muß.

Wir sprechen dies noch als einen Satz aus:

IV. Die charakteristischen Zahlen π, ν, ϱ der quadratischen Form φ sind gleich den Anzahlen der positiven, negativen und verschwindenden Wurzeln der zugehörigen Determinante $L_n(x)$.

Eine in allen Fällen brauchbare Sturmsche Kette erhält man auf folgendem Wege[1]):

Wir beschränken uns hier auf die Betrachtung von Gleichungen ohne mehrfache Wurzeln, oder wir nehmen an, daß vor der Anwendung des darzulegenden Verfahrens $f(x)$ von jedem gemeinschaftlichen Faktor mit seiner Derivierten $f'(x)$ befreit sei.

Wenn wir dann

$$f_1(x) = f'(x) \tag{14}$$

annehmen, so ist sicher die Bedingung 4. befriedigt. Nun verfahren wir so, als ob es sich um die Aufsuchung des größten gemeinschaftlichen Teilers von $f(x)$ und $f_1(x)$ handle, indem wir dabei jedesmal das Vorzeichen des Restes umkehren; wir bilden also durch Division die Gleichungen

---

[1]) Sturm, Mém. sur la résolution des équations numériques. Mém. de l'académie de Paris. Sav. étrang. VI, 1835. Auszug in Bull. de Ferussac XI, 1829. Deutsch von Loewy in Ostwalds Klassikern, Nr. 143.

$$
\begin{aligned}
f &= q_1 f_1 - f_2 \\
f_1 &= q_2 f_2 - f_3 \\
&\cdots\cdots\cdots\cdots \\
f_{m-2} &= q_{m-1} f_{m-1} - f_m,
\end{aligned} \tag{15}
$$

worin die $q_1, q_2, \dots q_{m-1}$ und ebenso die $f_1, f_2, \dots f_m$ ganze rationale Funktionen von x sind; die Grade der Funktionen $f, f_1, f_2 \dots f_m$ nehmen ab und man kann daher die Operation so weit fortsetzen, daß f_m konstant ist, oder wenigstens in dem betrachteten Intervall nicht mehr verschwindet. Daß f_m nicht Null werden kann, ist eine Folge der Voraussetzung, daß f und f_1 ohne gemeinsamen Teiler sind.

Daß man dann in der Reihe

$$f, f_1, f_2 \dots f_m \tag{16}$$

wirklich eine Sturmsche Kette hat, ergibt sich unmittelbar, wenn man die Kriterien § 34, 1. bis 4. durchgeht. Denn wenn zwei aufeinander folgende der Funktionen (16) zugleich verschwinden, so verschwinden nach (15) auch alle nachfolgenden; dies ist aber unmöglich, weil f_m von Null verschieden ist.

Ist aber $f_\nu(x) = 0$, so folgt aus (15):

$$f_{\nu+1}(x) = -f_{\nu-1}(x),$$

womit alle Forderungen für eine Sturmsche Kette befriedigt sind.

Es ist, wie sich von selbst versteht, gestattet, die Funktionen der Reihe (16) mit positiven, z. B. konstanten Faktoren zu multiplizieren, ohne daß sie aufhören, eine Sturmsche Kette zu bilden.

Für $n = 2$ können wir demnach als die Sturmschen Funktionen folgende nehmen:

$$
\begin{aligned}
f(x) &= x^2 + ax + b \\
f_1(x) &= 2x + a \\
f_2(x) &= a^2 - 4b.
\end{aligned}
$$

§ 37.

Lösung des Sturmschen Problems durch Hurwitz.

Eine besonders einfache Lösung des Sturmschen Problems hat Hurwitz gegeben[1]).

Es sei

$$f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n \tag{1}$$

[1]) Mathematische Annalen, Bd. 46.

eine ganze Funktion $n$ten Grades der Variablen z. Es sei ferner $\Phi(z)$ eine zweite ganze Funktion von beliebigem Grade. Wir suchen nach § 18 die Entwickelung des Bruches $\Phi(z):f(z)$ nach fallenden Potenzen und bezeichnen den Teil dieser Entwickelung, der negative Potenzen von z enthält, mit

$$c_0 z^{-1} + c_1 z^{-2} + \cdots = \sum_{0\ldots}^{h} c_h z^{-h-1}. \tag{2}$$

Es sei nun weiter

$$\Theta(z) = t_0 + t_1 z + \cdots + t_{m-1} z^{m-1} \tag{3}$$

eine ganze Funktion von einstweilen noch beliebigem Grade $(m-1)$, deren Koeffizienten $t_0, t_1 \ldots t_{m-1}$ unabhängige Variable sind. Wir bilden die Entwickelung nach fallenden Potenzen von z für den Bruch

$$\frac{\Phi(z)}{f(z)}\,\Theta(z)^2 = \frac{\Phi(z)}{f(z)} \sum_{0,}^{i} \sum_{m-1}^{k} z^{i+k} t_i t_k,$$

wofür wir nach (2) erhalten:

$$\sum^{h} \sum^{i} \sum^{k} c_h z^{-h+i+k-1} t_i t_k,$$

oder, nach absteigenden Potenzen von z geordnet,

$$\sum^{\lambda} z^{-\lambda-1} \sum^{i,k} c_{\lambda+i+k} t_i t_k. \tag{4}$$

Bezeichnen wir den Rest der Division von $\Theta(z)$ durch $f(z)$ mit $\Theta_1(z)$ und setzen

$$\Theta_1(z) = t'_0 + t'_1 z + \cdots + t'_{n-1} z^{n-1},$$

so sind die $t'_0, t'_1, \ldots t'_{n-1}$ lineare homogene Funktionen der Variablen $t_0, t_1, \ldots t_{m-1}$. Es geben aber nach § 18 die beiden Brüche

$$\frac{\Phi(z)\,\Theta(z)^2}{f(z)} \quad \text{und} \quad \frac{\Phi(z)\,\Theta_1(z)^2}{f(z)} \tag{5}$$

bei der Entwickelung nach fallenden Potenzen von z dieselben Glieder mit negativen Exponenten, und daraus folgt, daß die in der Entwickelung (4) auftretenden quadratischen Formen der m Variablen t_i

$$T_\lambda = \sum^{i,k} c_{\lambda+i+k} t_i t_k \cdot$$

auch als quadratische Formen der n Variablen t'_i ausgedrückt werden können. Sind also π_λ, ν_λ, ϱ_λ die charakteristischen Zahlen von T_λ, so ist ϱ_λ mindestens $= m - n$. Von nun an wollen wir $m = n$ setzen, also unter $\Theta(z)$ eine Funktion

$(n-1)$ten Grades verstehen. Wir betrachten hauptsächlich die Funktion T_0, die wir auch kurz mit T bezeichnen, also die quadratische Form

$$T = \sum_{0,n-1}^{i,k} c_{i+k} t_i t_k, \tag{6}$$

und bezeichnen ihre charakteristischen Zahlen mit π, ν, ϱ. Dann ergibt sich aus dem soeben Bewiesenen zunächst:

1. Wenn $f(z)$ und $\Phi(z)$ nicht relativ prim sind, so ist $\varrho > 0$, also T keine definite Form.

Denn wenn wir den größten gemeinschaftlichen Teiler von f und Φ herausheben, so entstehen aus (5) Brüche von derselben Gestalt, in denen Θ dasselbe geblieben ist, während sich der Grad des Nenners erniedrigt hat. Ist n' dieser erniedrigte Grad, so ist $m - n'$ und folglich auch ϱ positiv. Wir können ferner noch beweisen:

2. Wenn $f(z)$ und $f'(z)$ einen gemeinschaftlichen Teiler haben, so ist T keine definite Form.

Denn wenn $f(z)$ und $f'(z)$ einen gemeinsamen Teiler haben, so läßt sich eine reelle Funktion $Q(z)$ (vom ersten oder zweiten Grad) so bestimmen, daß

$$f(z) = Q^2 f_1(z)$$

ist. Dann ist $Qf_1(z)$ von niedrigerem als $n$tem Grade, und wir können die Koeffizienten $t_0, t_1, \ldots t_{n-1}$ so bestimmen, daß

$$\Theta = Qf_1(z)$$

wird. Dann ist aber $\Theta(z)^2$ durch $f(z)$ teilbar, und $\Phi(z)\Theta(z)^2 : f(z)$ ist eine ganze Funktion. Es kommen dann also in der Entwickelung nach fallenden Potenzen keine negativen Potenzen mehr vor, und es wird nicht nur T, sondern alle Funktionen T_λ gleich Null.

Es gibt also von Null verschiedene Werte der Variablen t_i, für die T verschwindet, was bei definiten Formen nicht möglich ist.

Wenn win annehmen, daß die Funktion $f(z)$ keine mehrfachen Faktoren hat, so können wir, wenn die Wurzeln von $f(z)$ bekannt sind, die Funktionen T_λ leicht in eine Summe von Quadraten zerlegen. Es ist nämlich dann, wenn das Summenzeichen S sich auf die Wurzeln

$$x = x_1, x_2, \ldots x_n$$

von $f(z)$ erstreckt, und $G(z)$ eine ganze Funktion bedeutet, durch Partialbruchzerlegung

$$\frac{\Phi(z)\Theta(z)^2}{f(z)} = S\frac{\Phi(x)\Theta(x)^2}{(z-x)f'(x)} + G(z), \tag{7}$$

und wenn wir hierin $1:(z-x)$ nach § 18 nach fallenden Potenzen von z entwickeln, so ergibt sich:

$$T_\lambda = S\frac{x^\lambda \Phi(x)\Theta(x)^2}{f'(x)}. \tag{8}$$

Die Funktionen

$$\begin{aligned}
\Theta(x_1) &= t_0 + t_1 x_1 + \cdots + t_{n-1} x_1^{n-1} \\
\Theta(x_2) &= t_0 + t_1 x_2 + \cdots + t_{n-1} x_2^{n-1} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\Theta(x_n) &= t_0 + t_1 x_n + \cdots + t_{n-1} x_n^{n-1},
\end{aligned} \tag{9}$$

deren Determinante gleich dem Differenzenprodukte der x_i und folglich von Null verschieden ist, sind voneinander linear unabhängige, lineare Funktionen der t_i und durch (8) ist also T_λ in eine Summe von n Quadraten zerlegt. Setzen wir

$$y_i = \sqrt{\frac{\Phi(x_i)}{f'(x_i)}}\,\Theta(x_i), \tag{10}$$

so ergibt sich (für $\lambda = 0$):

$$T = y_1^2 + y_2^2 + \cdots + y_n^2. \tag{11}$$

Unter der Voraussetzung reeller Koeffizienten von f und Φ lassen sich nun die für die Trägheit der Funktion T charakteristischen Zahlen π, ν, ϱ bestimmen.

Zunächst ist ersichtlich, daß y_i^2 aus der Summe (11) dann, und nur dann, wegfällt, wenn $\Phi(x_i) = 0$ ist, also wenn x_i eine gemeinschaftliche Wurzel von f und Φ ist. Es ist also ϱ gleich dem Grade des größten gemeinschaftlichen Teilers von f und Φ, und gleich Null, wenn f und Φ relativ prim sind.

Ist x_1, x_2 ein imaginäres Paar, und $\Phi(x_i)$ von Null verschieden, so ist

$$y_1 = P + Qi, \qquad y_2 = P - Qi$$

und folglich

$$y_1^2 + y_2^2 = 2(P^2 - Q^2).$$

Dieses Paar gibt also einen Beitrag von einer Einheit zu π sowohl als zu ν.

Für ein reelles x ist aber y^2 positiv oder negativ, je nachdem $\Phi(x)$ und $f'(x)$ gleiche oder verschiedene Zeichen haben.

Hieraus ergibt sich folgende Bestimmung für die charakteristischen Zahlen π, ν, ϱ der Form T:

(12)
ϱ ist gleich der Anzahl der verschwindenden $\Phi(x)$.

π ist gleich der Anzahl der imaginären Paare mit nicht verschwindenden $\Phi(x)$, vermehrt um die Anzahl der reellen Wurzeln mit positivem $\Phi(x) f'(x)$.

ν ist gleich der Anzahl der imaginären Paare mit nicht verschwindenden $\Phi(x)$, vermehrt um die Anzahl der reellen Wurzeln mit negativem $\Phi(x) f'(x)$.

Wir machen hiervon einige besondere Anwendungen.

Der Fall, daß T eine definite Form ist, daß also ϱ und $\nu = 0$ oder ϱ und $\pi = 0$ sind, kann nur dann eintreten, wenn Φ und f relativ prim sind, wenn $f(x)$ nur voneinander verschiedene reelle Wurzeln hat, und wenn außerdem $\Phi(x) f'(x)$ für alle Wurzeln von $f(x)$ dasselbe Zeichen hat. Sind α und β zwei der Größe nach aufeinander folgende von diesen Wurzeln, so haben $f'(\alpha)$ und $f'(\beta)$ entgegengesetzte Vorzeichen. Es müssen daher auch $\Phi(\alpha)$ und $\Phi(\beta)$ entgegengesetzte Vorzeichen haben. Es muß also zwischen α und β eine ungerade Zahl von Wurzeln von $\Phi(x) = 0$ liegen. Unter diesen Voraussetzungen ist auch umgekehrt T eine definite Form, die im Zeichen mit $\Phi(x) f'(x)$ übereinstimmt. Φ kann also, wenn T eine definite Form sein soll, nicht von niedrigerem Grade sein als $n - 1$. Ist Φ auch nicht von höherem Grade, so muß zwischen je zwei Wurzeln von $f(x)$ eine und nur eine Wurzel von $\Phi(x)$ liegen, oder wie wir uns auch ausdrücken können, die Wurzeln von $f(x)$ sind durch die Wurzeln von $\Phi(x)$ voneinander getrennt.

Die Form T wird speziell eine positive sein, wenn wir außerdem noch annehmen, daß c_0 positiv ist, und dies wird dann eintreten, wenn die Koeffizienten der höchsten Potenzen von z in $f(z)$ und $\Phi(z)$ dasselbe Zeichen haben. Da wir nun in § 12, XXIII. die notwendige und hinreichende Bedingung für eine positive Form aufgestellt haben, so können wir jetzt folgenden Satz aussprechen, wenn wir

$$\begin{vmatrix} c_0, c_1 \dots\dots c_\nu \\ c_1, c_2 \dots\dots c_{\nu+1} \\ \dots\dots\dots\dots\dots \\ c_\nu, c_{\nu+1} \dots c_{2\nu} \end{vmatrix} = C_\nu$$

setzen:

V. Die notwendige und hinreichende Bedingung dafür, daß die Gleichung $n$ten Grades $f(x) = 0$ nur voneinander verschiedene reelle Wurzeln hat, die durch die Wurzeln der Gleichung $(n-1)$ten Grades $\Phi(x) = 0$ voneinander getrennt sind, ist, wenn die Koeffizienten der höchsten Potenzen von f und Φ von einerlei Zeichen sind, die, daß die Determinanten

$$C_0, C_1, C_2 \ldots C_{n-1}$$

alle positiv sind.

Wenn wir $\Phi(z) = f'(z)$ setzen, so ist $\Phi(z) : f'(z) = 1$, also immer positiv. Zugleich gehen in diesem Falle die $c_0, c_1, c_2 \ldots$ (nach § 18) in die Potenzsummen $s_0, s_1, s_2 \ldots$ über. Setzen wir dann

$$\begin{vmatrix} s_0, s_1 \ldots\ldots s_\nu \\ s_1, s_2 \ldots\ldots s_{\nu+1} \\ \ldots\ldots\ldots\ldots \\ s_\nu, s_{\nu+1} \ldots s_{2\nu} \end{vmatrix} = D_\nu,$$

so ist (wenn $a_0 = 1$ angenommen wird) D_{n-1} die Diskriminante von f und es folgt der Satz:

VI. Die notwendige und hinreichende Bedingung dafür, daß die Gleichung $n$ten Grades $f(x) = 0$ nur reelle voneinander verschiedene Wurzeln hat, besteht darin, daß die Determinanten

$$D_0, D_1, D_2 \ldots D_{n-1}$$

positiv sind.

Nimmt man $\Phi(x) = (x - a) f'(x)$ an, worin a eine beliebige reelle Zahl ist, die nicht zu den Wurzeln von $f(x)$ gehört, so ist, wenn $f(x)$ nur einfache Wurzeln hat, $\varrho = 0$, und, mag nun $f(x)$ imaginäre Wurzeln haben oder nicht, es ist $\pi - \nu$ die Anzahl der reellen Wurzeln, die größer als a sind, vermindert um die Anzahl der unter a liegenden Wurzeln. Die Differenz $\pi - \nu$ nimmt also jedesmal um zwei Einheiten ab, wenn a wachsend durch eine Wurzel von $f(x)$ hindurchgeht. Nimmt man zwei reelle Zahlen $a < b$, die nicht zu den Wurzeln von $f(x)$ gehören, so ergibt sich hieraus für die Zahl der zwischen a und b gelegenen Wurzeln $\pi_a - \pi_b$ oder $\nu_b - \nu_a$, und dies kann man nach § 12 auch so ausdrücken, daß die Anzahl der zwischen a und b

gelegenen Wurzeln von $f(x)$ gleich ist der Anzahl der in der Reihe

$$C_0, C_1, C_2 \ldots C_n$$

beim Übergang der Variablen x von a zu b verlorenen Zeichenwechsel.

§ 38.

Abschätzung der Wurzeln.

Die Funktionen der Sturmschen Ketten lösen zwar vollständig und ausnahmslos das Problem der Bestimmung der Anzahl der Wurzeln einer Gleichung zwischen gegebenen Grenzen, aber ihre wirkliche Berechnung ist meist schwierig und oft unausführbar. Man kennt eine Reihe von Sätzen zur Abschätzung der Zahl der Wurzeln zwischen gegebenen Grenzen, die viel einfachere Regeln liefern, aber freilich auch die Frage nicht vollständig beantworten, sondern nur eine Maximalzahl geben, worüber die Anzahl der Wurzeln nicht hinausgehen kann. Diese Regeln sind oft für die Anwendung sehr nützlich und ausreichend und sie dürfen daher hier nicht fehlen.

Wir betrachten zunächst ein Verfahren, das unter dem Namen des Budan-Fourierschen Theorems bekannt ist[1]).

Betrachten wir an Stelle einer Sturmschen Kette die Reihe der Ableitungen einer reellen Funktion $f(x)$ vom $n$ten Grade:

$$(1) \qquad f(x), f'(x), f''(x) \ldots f^{(n)}(x),$$

worin also $f^{(n)}(x)$ eine Konstante ist, die wir von Null verschieden und positiv voraussetzen. Sie ist, von dem positiven Zahlenfaktor $\Pi(n)$ abgesehen, der Koeffizient von x^n in $f(x)$.

Es seien α und $\beta > \alpha$ zwei reelle Zahlen, die das Intervall (α, β) bestimmen, in dem die Anzahl der reellen Wurzeln von $f(x)$ gezählt werden sollen.

Wir wollen zunächst annehmen, daß in dem Intervall (α, β) nicht zwei Glieder der Reihe (1) zugleich verschwinden, und daß für $x = \alpha$, $x = \beta$ kein Glied der Reihe verschwindet.

Lassen wir nun x von α bis β stetig wachsen, so kann in der Vorzeichenfolge der Reihe (1) nur dann eine Änderung

---

[1]) Wie G. Darboux in den Oeuvres de Fourier nachgewiesen hat, verdient das Theorem nur nach Fourier benannt zu werden. Vgl. die deutsche Ausgabe „Die Auflösung der bestimmten Gleichungen" von Fourier in Nr. 127 von Ostwalds Klassikern und die Anmerkungen von Loewy.

eintreten, wenn eine der Funktionen durch Null geht. Wenn $f^{(\nu)}(x)$ für $x = \xi$ durch Null geht, so geht $f^{(\nu)}(x)$, je nachdem $f^{(\nu+1)}(x)$ positiv oder negativ ist, von negativen zu positiven oder von positiven zu negativen Werten über, und es geht also zwischen $f^{(\nu)}$ und $f^{(\nu+1)}$ beim Durchgang durch ξ ein Zeichenwechsel verloren. Dies gilt auch, wenn $f^{(\nu)}$ die Funktion $f(x)$ selbst ist. Wenn aber $f^{(\nu)}$ eine der Derivierten ist, so geht ihr eine Funktion $f^{(\nu-1)}$ voran, und da $f^{(\nu-1)}$ nach Voraussetzung für $x = \xi$ nicht Null ist, und $f^{(\nu)}$ beim Durchgang durch ξ sein Zeichen wechselt, so findet zwischen $f^{(\nu-1)}$ und $f^{(\nu)}$ beim Durchgang durch ξ entweder ein Verlust oder ein Gewinn von einem Zeichenwechsel statt. Wenn also ein inneres Glied der Reihe (1) durch Null geht, so bleibt die Anzahl der Zeichenwechsel ungeändert, oder es gehen zwei Zeichenwechsel verloren.

Geht aber $f(x)$ selbst durch Null, so geht ein Zeichenwechsel verloren.

Daraus folgt das Theorem:

VII. Die Anzahl der zwischen α und β gelegenen Wurzeln von $f(x)$ ist höchstens so groß, wie die Zahl der zwischen α und β verlorenen Zeichenwechsel, und wenn sie kleiner ist, so ist der Unterschied eine gerade Zahl.

Mit Benutzung einer Formel können wir auch sagen:

Ist $V(x)$ die Anzahl der Zeichenwechsel (Variationen), die die Reihe (1) für irgend einen Wert x darbietet, so ist die Anzahl der zwischen α und β gelegenen Wurzeln von $f(x)$

$$V(\alpha) - V(\beta) - 2h, \tag{2}$$

worin h eine nicht negative ganze Zahl ist.

Der Beweis des Satzes VII bedarf noch einer Ergänzung für den Fall, daß in der Reihe (1) mehrere aufeinander folgende Glieder zugleich verschwinden.

Es mögen also für einen Wert ξ von x zwischen α und β in der Reihe

$$f^{(\nu)}(x), \quad f^{(\nu+1)}(x), \ldots f^{(\nu+\mu-1)}(x), \quad f^{(\nu+\mu)}(x) \tag{3}$$

alle Glieder, mit Ausnahme des letzten, verschwinden, und das letzte $f^{(\nu+\mu)}(x)$ möge etwa einen positiven Wert haben.

Wir grenzen um ξ zwei Intervalle δ_1, δ_2 ab, so daß alle Werte von x im Intervall δ_1 kleiner, im Intervall δ_2 größer

als ξ sind, und nehmen diese Intervalle so klein, daß die Funktionen (3) außer in ξ darin nicht verschwinden, also auch $f^{(\nu+\mu)}(x)$ positiv bleibt.

Da nun, wenn $f^{(\nu+\mu)}(x)$ positiv ist, $f^{(\nu+\mu-1)}(x)$ mit x zugleich wächst, so ist

$$f^{(\nu+\mu-1)}(x) \text{ in } \delta_1 \text{ negativ, in } \delta_2 \text{ positiv.}$$

Daraus folgt, daß $f^{(\nu+\mu-2)}(x)$ in δ_1 abnimmt, in δ_2 wächst, also in beiden Intervallen positiv ist, und so schließen wir weiter auf die Vorzeichenfolge in der Reihe (3):

(4)

	$f^{(\nu)}(x)$,	$f^{(\nu+1)}(x)$,	$f^{(\nu+2)}(x)$	...	$f^{(\nu+\mu-1)}(x)$,	$f^{(\nu+\mu)}(x)$
δ_1,	$(-1)^{\mu}$	$(-1)^{\mu-1}$	$(-1)^{\mu-2}$	...	$-$	$+$
δ_2,	$+$	$+$	$+$		$+$	$+$

d. h. in der Reihe (3) werden beim Durchgang durch ξ aus lauter Zeichenwechseln Zeichenfolgen, und es gehen in der Reihe (3) μ Zeichenwechsel verloren.

Ist $f^{(\nu+\mu)}(\xi)$ negativ, so sind alle Zeichen in (4) die entgegengesetzten, und der Schluß bleibt derselbe.

Wenn nun $\nu = 0$, d. h. $f^{(\nu)}(x)$ die ursprüngliche Funktion $f(x)$ selbst ist, die dann in ξ eine μfache Wurzel hat, so findet also beim Durchgang durch ξ auch in der Reihe (1) ein Verlust von μ Zeichenwechseln statt.

Ist aber $\nu > 0$ und die dem $f^{(\nu)}(x)$ vorangehende Funktion $f^{(\nu-1)}(x)$ in ξ von Null verschieden, so haben wir folgende Zeichen:

1. μ gerade:

	a) $f^{(\nu-1)}(x)$,	$f^{(\nu)}(x)$,	b) $f^{(\nu-1)}(x)$,	$f^{(\nu)}(x)$
δ_1,	$+$	$+$	$-$	$+$
δ_2,	$+$	$+$	$-$	$+$

2. μ ungerade:

	a) $f^{(\nu-1)}(x)$,	$f^{(\nu)}(x)$,	b) $f^{(\nu-1)}(x)$,	$f^{(\nu)}(x)$
δ_1,	$+$	$-$	$-$	$-$
δ_2,	$+$	$+$	$-$	$+$

Es geht also bei geradem μ zwischen $f^{(\nu-1)}$ und $f^{(\nu)}$ kein Zeichenwechsel verloren, bei ungeradem μ geht ein Zeichenwechsel verloren oder es wird einer gewonnen. Es ist also der Verlust an Zeichenwechseln beim Durchgang durch ξ in der Reihe

(5) $$f^{(\nu-1)},\ f^{(\nu)},\ f^{(\nu+1)} \ \ldots\ f^{(\nu+\mu)}$$

bei geradem μ gleich μ, bei ungeradem μ gleich $\mu \pm 1$, also immer eine gerade Zahl und nie negativ.

Wir können diesen Ergebnissen einen übersichtlichen analytischen Ausdruck geben, der ihre Bedeutung besser erkennen läßt.

Wir bezeichnen mit $V(x)$ wie früher die Anzahl der Zeichenwechsel in der Reihe (1) für einen bestimmten Wert von x, jedoch mit der näheren Bestimmung, daß, wenn für einen Wert von x einige der Glieder der Reihe verschwinden, diese bei Abzählung der Zeichenwechsel einfach übergangen werden. Ist in der Reihe (1) nur ein einziges von Null verschiedenes Glied vorhanden, so ist die Anzahl der Zeichenwechsel ebenso wie die der Zeichenfolgen $= 0$ zu setzen. Wir können dann $V(x)$ als eine Funktion von x auffassen, die sich aber nur um ganze Zahlen ändern kann, also unstetig ist für die Werte von x, für welche einige Glieder der Reihe (1) verschwinden. Wir wollen dann mit $V(x-0)$ und $V(x+0)$ die Werte der Funktion $V(x)$ unmittelbar vor und unmittelbar nach einer solchen Stelle bezeichnen. Dann zeigt die Betrachtung von (5), daß in allen Fällen

$$V(\xi + 0) = V(\xi) \tag{6}$$

ist, und daß, wenn ξ eine μ-fache Wurzel von $f(x)$ ist,

$$V(\xi - 0) = V(\xi) + \mu + 2h, \tag{7}$$

worin h eine nicht negative ganze Zahl ist. Diese Formeln (6), (7) gelten auch dann noch, wenn in der Reihe (1) für $x = \xi$ mehrere Reihen verschwindender Funktionen wie (4) vorkommen.

Markieren wir in dem Intervall $\alpha \leqq x \leqq \beta$ die in endlicher Anzahl ν vorhandenen Unstetigkeitspunkte von $V(x)$:

$$\alpha,\ \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_\nu = \beta, \tag{8}$$

dann ist damit in jedem der Teilintervalle $(\alpha_i + 0) \ldots (\alpha_{i+1} - 0)$ die Funktion $V(x)$ ungeändert, also:

$$V(\alpha_i + 0) - V(\alpha_{i+1} - 0) = 0,$$

und es ergibt sich, wenn wir mit μ_i die Anzahl der in α_i vereinigten Wurzeln von $f(x)$ und mit h_i nicht negative ganze Zahlen bezeichnen, nach (6) und (7):

$$0 = V(\alpha_i) - V(\alpha_{i+1}) - \mu_{i+1} - 2h_{i+1}. \tag{9}$$

Summiert man diese Gleichungen für $i = 0, 1, \ldots \nu - 1$ und setzt

$$N = \Sigma \mu_{i+1}, \quad H = \Sigma h_{i+1},$$

so erhält man die Verschärfung des Theorems VII:

$$V_\alpha - V_\beta = N + 2H.$$

VIII. Hierin ist N die Anzahl aller Wurzeln (jede nach ihrer Vielfachheit gerechnet) in dem Intervall $\alpha < x \leqq \beta$, d. h. die nach β fallenden Wurzeln mitgezählt, die nach α fallenden ausgeschlossen[1]).

Das Budan-Fouriersche Theorem gibt zwar nicht, wie der Sturmsche Satz, eine vollständig sichere Entscheidung über die Zahl der Wurzeln in einem Intervall, es kann aber doch in manchen Fällen den Sturmschen Satz ersetzen; denn hat man das Intervall (α, β) so weit eingeschränkt, daß kein oder nur ein Zeichenwechsel von α bis β verloren geht, so folgt mit Sicherheit, daß im ersten Falle keine, im zweiten eine und nur eine Wurzel im Intervall liegt.

Die Abgrenzung der Wurzeln kann also durch den Fourierschen Satz nur dann vollständig gegeben werden, wenn nicht beim Durchgang durch einen Wert ξ gleichzeitig mehrere Zeichenwechsel verloren gehen. Dies Verhalten kann, wie klein auch das Intervall (α, β) gewählt sein mag, nicht mit Sicherheit erkannt werden, und man wird also, wenn nach einer angemessenen Einengung des Intervalls die Abgrenzung nicht gelungen ist, doch zu einem anderen Verfahren, in letzter Instanz zum Sturmschen Satze greifen müssen.

Ein komplizierteres Verfahren, das etwas mehr leistet, hat Newton gegeben, auf das hier nicht eingegangen werden soll.

Das Fouriersche Theorem gibt uns auch eine obere Grenze, d. h. eine positive Zahl, die größer ist als der absolute Wert sämtlicher reeller Wurzeln einer gegebenen Gleichung.

Nehmen wir den Koeffizienten der höchsten Potenz von x in $f(x)$ gleich 1 (oder wenigstens positiv) an, so kann man die positive Zahl α immer so groß annehmen, daß

$$f(\alpha),\ f'(\alpha),\ f''(\alpha) \ldots f^{(n)}(\alpha) \tag{10}$$

alle positiv werden. Dann geht in der Reihe der Funktionen

$$f(x),\ f'(x),\ f''(x) \ldots f^{(n)}(x)$$

zwischen $x = \alpha$ und $x = \infty$ kein Zwischenwechsel mehr verloren, und es kann also auch nach dem Theorem VIII keine Wurzel von $f(x)$ zwischen α und ∞ liegen.

---

[1]) Hurwitz: Über den Satz von Budan-Fourier. Mathematische Annalen, Bd. 71.

Will man aus dem gleichen Satze für die negativen Wurzeln eine untere Grenze haben, so nehme man die positive Zahl β so an, daß

$$(11) \qquad (-1)^n f(-\beta),\ (-1)^{n-1} f'(-\beta) \dots f^{(n)}(-\beta)$$

alle positiv werden, dann hat die Gleichung $f(x) = 0$ sicher keine negative Wurzel unter $-\beta$.

Ein Mittel zur Abschätzung der Zahl der positiven Wurzeln gibt der Cartesische Lehrsatz. Wir können ihn einfach als speziellen Fall des Fourierschen Theorems betrachten, indem wir $\alpha = 0$ und $\beta = \infty$ oder wenigstens so groß annehmen, daß die Funktionen

$$(12) \qquad f(x),\ f'(x),\ f''(x) \dots f^{(n)}(x)$$

für $x > \beta$ alle dasselbe Zeichen haben. Dies Zeichen stimmt überein mit dem Zeichen des Koeffizienten der höchsten Potenz von x in $f(x)$ und kann positiv angenommen werden.

Ist

$$f(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_{n-1} x + a_n,$$

so erhalten die Funktionen (12), von positiven Zahlenfaktoren abgesehen, für $x = 0$ die Werte

$$(13) \qquad a_n,\ a_{n-1},\ a_{n-2} \dots a_0,$$

und daraus ergibt sich also das Theorem:

IX. Die Anzahl der positiven Wurzeln der Gleichung $f(x) = 0$ ist gleich oder um eine gerade Zahl kleiner als die Anzahl der Zeichenwechsel in der Reihe der Koeffizienten von $f(x)$, wobei etwa verschwindende Koeffizienten einfach zu übergehen sind und die Wurzel $x = 0$, wenn sie vorhanden ist, nicht mitzählt.

Um auch die Anzahl der negativen Wurzeln abzuschätzen, kann man entweder das Fouriersche Theorem auf das Intervall $(-\infty, 0)$ anwenden, oder man ersetzt x durch $-x$, d. h. man ändert die Vorzeichen von $a_1, a_3, a_5 \dots$ und wendet dann das Theorem IX an.

Siebenter Abschnitt.

Genäherte Berechnung der Wurzeln.

§ 39.

Interpolation.

Mit der Abgrenzung der Intervalle, in denen nur je eine Wurzel einer algebraischen Gleichung liegt, ist die Möglichkeit gegeben, die reellen Wurzeln mit beliebiger Genauigkeit numerisch zu berechnen, indem man das Intervall mehr und mehr einengt, z. B. fortgesetzt halbiert. Teilt man ein Intervall von der Größe 1 fortgesetzt in zehn Teile, so liefert jeder neue Schritt eine weitere Dezimalstelle.

Wenn einmal erst eine Wurzel von allen übrigen abgesondert ist, so hat man bei der weiteren Einengung des Intervalls immer nur das Vorzeichen der Funktion $f(x)$ selbst zu berücksichtigen. Die dazu nötigen Rechnungen können sehr erleichtert werden durch die Anwendung der Interpolationsformel, die wir im § 17 kennen gelernt haben.

Wir haben nämlich dort eine Formel hergeleitet, nach der man eine Funktion $n$ten Grades, die wir jetzt mit $\varphi(\xi)$ bezeichnen wollen, bestimmen kann, wenn $n+1$ aufeinander folgende Werte $\varphi(0)$, $\varphi(1)$, ... $\varphi(n)$ gegeben sind. Diese Formel ist, wenn

$$(1)\qquad B_\nu^\xi = \frac{\xi(\xi-1)\dots(\xi-\nu+1)}{1.2.3\dots\nu}$$

$$(2)\qquad \begin{aligned} \Delta_\xi &= \varphi(\xi+1) - \varphi(\xi),\\ \Delta'_\xi &= \Delta_{\xi+1} - \Delta_\xi\\ &\dots\dots\dots\dots\dots\dots\dots\dots\\ \Delta_\xi^{(n-1)} &= \Delta_{\xi+1}^{(\nu-2)} - \Delta_\xi^{(n-2)} \end{aligned}$$

gesetzt wird,

$$(3)\qquad \varphi(\xi) = \varphi(0) + \Delta_0 B_1^{(\xi)} + \Delta'_0 B_2^{(\xi)} + \dots + \Delta_0^{(n-1)} B_n^{(\xi)}.$$

Der Wert der Formel liegt darin, daß die Reihe der Differenzen Δ_0, Δ'_0, Δ''_0 ... in vielen Fällen so rasch abnehmende Zahlenwerte bietet, daß man sich mit einigen der ersten Glieder der Reihe (3) begnügen kann.

Ist nun $f(x)$ die zu untersuchende Funktion und (α, β) das Intervall, in dem eine der gesuchten Wurzeln liegt, so setzen wir, indem wir unter m irgend eine geeignete ganze Zahl verstehen,

$$\delta = \frac{\beta - \alpha}{m}, \quad x = \alpha + \delta\,\xi,$$

und wenn wir also

$$\Delta_x = f(x + \delta) - f(x)$$
$$\Delta'_x = \Delta_{x+\delta} - \Delta_x$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

setzen, und die Formel (3) auf $\varphi(\xi) = f(x)$ anwenden, so folgt:

$$(4)\quad f(x) = f(\alpha) + \frac{\Delta_\alpha}{\delta}(x - \alpha) + \frac{\Delta'_\alpha}{\delta^2}\frac{(x-\alpha)(x-\alpha-\delta)}{2} + \cdots$$

Man könnte ebensogut auch von dem anderen Endpunkt β des Intervalls ausgehen, und müßte dann in (3)

$$x = \beta - \delta\,\xi$$

setzen. Man würde dann erhalten:

$$\Delta_x = f(x - \delta) - f(x)$$
$$\Delta'_x = \Delta_{x-\delta} - \Delta_x$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(5)\quad f(x) = f(\beta) - \frac{\Delta_\beta}{\delta}(x - \beta) + \frac{\Delta'_\beta}{\delta^2}\frac{(x-\beta)(x-\beta+\delta)}{2} + \cdots$$

Wenn man die Gleichung

$$y = f(x)$$

als Gleichung einer Kurve in einem rechtwinkligen Koordinatensystem x, y deutet, so lassen sich die Verhältnisse geometrisch veranschaulichen.

Wenn man z. B. die Gleichung (4) auf das Intervall von α bis $(\alpha + \delta)$ anwendet und sich mit der Berücksichtigung der ersten Differenz begnügt, so heißt das in der Sprache der Geometrie, daß man den zwischen den beiden Kurvenpunkten α, $f(\alpha)$ und $\alpha + \delta$, $f(\alpha + \delta)$ verlaufenden Kurvenbogen durch die Sehne ersetzt. Berücksichtigt man auch die zweite Differenz, so wird der durch die drei aufeinander folgenden Punkte mit den

Abszissen α, $\alpha + \delta$, $\alpha + 2\delta$ gehende Kurvenbogen durch den Bogen einer Parabel ersetzt, die durch dieselben Punkte geht und deren Achse der Ordinatenachse parallel ist; und diese Parabel wird sich der Kurve noch enger anschließen als die Sehne. Je kleiner das Intervall δ ist, um so weniger werden die höheren Differenzen von Einfluß sein.

Wie weit man also bei der Annäherung zu gehen hat, das hängt nicht nur von der Genauigkeit ab, mit der man $f(x)$ zu kennen wünscht, sondern wesentlich auch von der Dichtigkeit der Werte α, $\alpha + \delta$, $\alpha + 2\delta$ ..., für die die Funktion bekannt ist. Auf diesen Sätzen beruht die Einrichtung unserer Tabellenwerke, besonders der Logarithmentafeln. Es handelt sich dabei freilich nicht um ganze rationale Funktionen; allein bei den stetigen Funktionen überhaupt gelten hier dieselben Gesetze. Man findet in den Tafeln daher auch neben den Werten $f(\alpha)$, $f(\alpha + \delta)$, $f(\alpha + 2\delta)$ ... die Werte der ersten oder der beiden ersten Differenzen angegeben. Bei den gebräuchlichen siebenstelligen Tafeln genügt die erste Differenz. In der zehnstelligen Tafel „Thesaurus logarithmorum“ von Vega sind auch die zweiten Differenzen angegeben und müssen bei ganz scharfen Rechnungen berücksichtigt werden.

Unsere Interpolationsformeln lassen sich mit Nutzen anwenden, um die Wurzeln der Gleichungen zu berechnen, oder, genauer ausgedrückt, die auf anderem Wege gefundenen Näherungswerte zu verbessern.

Wir können die Aufgabe so formulieren, daß zu einem gegebenen, zwischen $f(\alpha)$ und $f(\alpha + \delta)$ gelegenen Wert von $f(x)$ der zugehörige Wert von x gefunden werden soll.

Wir betrachten $x = \alpha$ als einen ersten Näherungswert. Setzen wir nun

$$f(x) - f(\alpha) = \Delta,$$

so ist, wenn die Funktion $f(x)$ zwischen $x = \alpha$ und $x = \alpha + \delta$ nur wächst oder nur abnimmt, Δ ein gegebener Wert von demselben Zeichen wie $\Delta_\alpha = f(\alpha + \delta) - f(\alpha)$ und absolut kleiner als Δ_α. Setzen wir noch $x - \alpha = u$, so gibt die Formel (4):

$$(6) \qquad \Delta = \frac{\Delta_\alpha u}{\delta} + \frac{\Delta'_\alpha}{\delta^2} \frac{u(u-\delta)}{2} + \cdots$$

Bleiben wir zunächst bei der ersten Differenz stehen, so ergibt sich als erste Korrektion:

(7) $$u = \delta \frac{\Delta}{\Delta_\alpha},$$

und wenn wir nun

$$u = \delta \frac{\Delta}{\Delta_\alpha} + u'$$

setzen, so folgt aus (6), wenn man im zweiten Gliede u' wegläßt,

$$0 = \frac{\Delta_\alpha}{\delta} u' + \frac{\Delta'_\alpha \Delta (\Delta - \Delta_\alpha)}{2 \Delta_\alpha^2},$$

also als zweite Korrektion:

(8) $$u' = \delta \frac{\Delta'_\alpha \Delta (\Delta_\alpha - \Delta)}{2 \Delta_\alpha^3}.$$

Die erste Korrektion (7) erhält man dadurch, daß man den zwischen α und $\alpha + \delta$ verlaufenden Kurvenbogen durch die Sehne ersetzt, wie oben, die zweite Korrektion (8) dadurch, daß man den Bogen zwischen α, $\alpha + \delta$, $\alpha + 2\delta$ durch eine Parabel ersetzt, die durch dieselben Punkte geht, die aber jetzt ihre Achse mit der x-Achse parallel hat.

Nehmen wir als Beispiel die Gleichung

$$f(x) = x^3 - 2x - 2 = 0,$$

die zwischen 1,7 und 1,8 eine reelle Wurzel hat.

Man berechnet

$x = 1{,}7,$	$f(x) = -\ 0{,}487,$	$\Delta_\alpha = 0{,}719,$	$\Delta'_\alpha = 0{,}108$
$1{,}8,$	$= +\ 0{,}232,$	$= 0{,}827,$	
$1{,}9,$	$=\ 1{,}059.$		

Da $f(x) = 0$ sein soll, so ist

$$\Delta = 0{,}487$$

zu setzen und die erste Korrektion zu $\alpha = 1{,}7$ ist nach (7):

$$u = 0{,}067\,73,$$

die zweite Korrektion ergibt nach (8):

$$u' = 0{,}001\,64,$$

also:

$$\alpha + u + u' = 1{,}769\,37.$$

Der auf andere Weise berechnete genauere Wert ist 1,76929... Wir haben also ein in den ersten drei Dezimalen genaues Resultat. Wir haben aber hier kein anderes Mittel, um die Genauigkeit von vornherein zu schätzen, als die Abnahme der Differenzen Δ, Δ', Δ'' ... Ist Δ' so klein, daß es außerhalb der Grenzen

der beabsichtigten Genauigkeit fällt, so gibt die Berücksichtigung der ersten Differenz ein genaues Resultat. Die hier auseinandergesetzte Vorschrift zur Wurzelberechnung wird die Regula falsi genannt.

Die einfachste, für die erste Annäherung geeignete Form dieser Vorschrift ist die:

Liegt zwischen α und β eine Wurzel x der Gleichung $f(x)=0$ und ist $f(\alpha) = -a$, $f(\beta) = b$, so ist

$$x = \frac{b\alpha + a\beta}{a + b} \tag{9}$$

ein genäherter Wert von x. Dies ist nur eine andere Schreibweise für die Formel (7).

§ 40.

Die Newtonsche Näherungsmethode.

Eine Methode, die zur genäherten Berechnung der Wurzeln einer Gleichung meist besser ist als die Interpolation, rührt von Newton her und wurde von Fourier ausgebildet und genauer untersucht. Sie besteht in folgendem. Es sei

$$f(x) = 0 \tag{1}$$

die aufzulösende Gleichung und es sei ein Wert $x = \alpha$ gefunden, den man als eine gewisse Annäherung an eine Wurzel betrachten kann. Wir setzen in (1)

$$x = \alpha + h,$$

und erhalten:

$$f(\alpha + h) = f(\alpha) + hf'(\alpha) + \frac{h^2}{1.2} f''(\alpha) + \cdots$$

Wenn man nun h aus der Gleichung bestimmt:

$$f(\alpha) + hf'(\alpha) = 0,$$

was voraussetzt, daß $f'(\alpha)$ von Null verschieden ist, so wird

$$f(\alpha + h) = \frac{h^2}{1.2} f''(\alpha) + \cdots,$$

und wird also, wenn h eine kleine Zahl ist, da nur das Quadrat und höhere Potenzen von h vorkommen, einen kleinen Wert haben. Es wird also unter den geeigneten Voraussetzungen

$$\alpha' = \alpha - \frac{f(\alpha)}{f'(\alpha)} \tag{2}$$

als eine bessere Annäherung an den wahren Wert der Wurzel zu betrachten sein.

Ersetzt man dann α durch α', so wird man in

$$\alpha'' = \alpha' - \frac{f(\alpha')}{f'(\alpha')}$$

eine noch bessere Näherung erhalten usf.

Es bleiben hier aber noch folgende beiden Fragen zu beantworten:

1. Unter welchen Voraussetzungen ist α' wirklich ein besserer Wert als α?
2. Wie kann man den Grad von Genauigkeit schätzen, den man so erreicht?

Diese Fragen hat Fourier beantwortet; er macht aber dabei folgende Voraussetzung:

Es ist eine Wurzel von $f(x)$ in einem Intervall (α, β) eingeschlossen, das keine zweite Wurzel enthält.

In dem Intervall (α, β) ist $f'(x)$ und $f''(x)$ von Null verschieden.

Was die letztere Voraussetzung betrifft, daß $f''(x)$ im Intervall von Null verschieden ist, so ist sie nur gemacht, um einfacher auszudrückende Bedingungen für die Anwendung der Methode zu erhalten. An sich ist ihre Brauchbarkeit bei genügender Einengung des Intervalls davon nicht abhängig. Wenn aber $f(x)$ und $f''(x)$ keinen gemeinsamen Teiler haben, also nicht zugleich verschwinden, so kann man die Fouriersche Voraussetzung durch Einengung des Intervalls immer erfüllen, und wenn $f(x)$ und $f'(x)$ einen gemeinsamen Teiler haben, so kann man diesen zuvor absondern und dann auf die einzelnen Faktoren von $f(x)$ die Näherungsmethoden anwenden.

Am einfachsten übersieht man die Verhältnisse in der Geometrie, wenn man $y = f(x)$ als Gleichung einer ebenen Kurve in einem rechtwinkligen Koordinatensystem deutet.

Die Gleichung

$$y = f(\alpha) + (x - \alpha) f'(\alpha) \tag{3}$$

ist die Gleichung der Kurventangente in dem Punkte α, $f(\alpha)$ und die Newtonsche Näherungsmethode kommt also darauf hinaus, daß man die Kurve in dem Intervall (α, β) in erster Annäherung durch die Tangente in einem der Endpunkte ersetzt, anstatt wie bei der Interpolationsmethode durch ihre Sehne.

Wenn der zweite Differentialquotient in dem Intervall (α, β) verschwindet, also die Kurve einen Wendepunkt hat, so kann der Fall eintreten, daß beide Endtangenten aus dem Intervall hinausführen; dann ist die Newtonsche Methode also nicht anwendbar (Fig. 4). Es kann aber auch, wenn das Intervall (α, β) schon genügend eingeengt ist, der andere Fall eintreten, daß

Fig. 4.

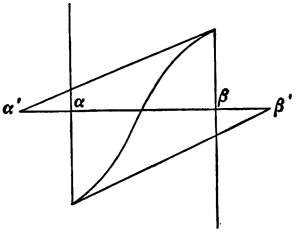
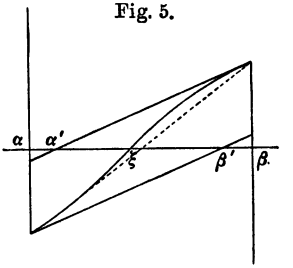


Fig. 5.



beide Tangenten nach inneren Punkten des Intervalls führen (Fig. 5). Indessen wird in diesen Fällen immer die Regula falsi eine bessere Annäherung geben.

Wir wollen aber jetzt annehmen, daß $f'(x)$ und $f''(x)$ in dem Intervall (α, β) nicht verschwinden. Dann ändert die Kurve den Sinn ihrer Krümmung nicht, und dann hat gewiß immer

Fig. 6.

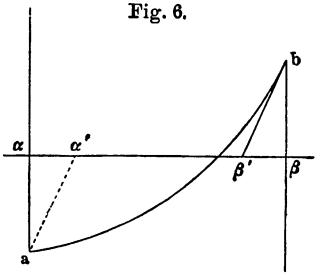
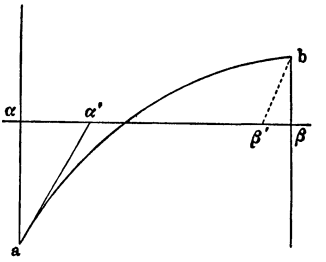


Fig. 7.



eine der beiden Endtangenten ihren Schnittpunkt mit der x-Achse im Inneren des Intervalls. Dies trifft sicher zu, wenn die Tangente in dem Endpunkte des Intervalls genommen wird, in dem $f(x)$ und $f''(x)$ dasselbe Vorzeichen haben. Die beiden Fig. 6 und 7 veranschaulichen das Verhältnis bei positivem $f'(x)$ und bei positivem und negativem $f''(x)$.

Es ist also im ersten Falle β', im zweiten α' ein besserer Annäherungswert, als β und α.

Will man gleichzeitig die andere Grenze verschieben, um ein neues engeres Intervall zu bekommen, so kann man nach Fourier von dem anderen Endpunkte, also im Falle der Fig. 6, in a die zu $b\beta'$ parallele gerade Linie ziehen, und erhält den Punkt α' als unteren Grenzpunkt des neuen Intervalls. Es ist dann im ersten Falle

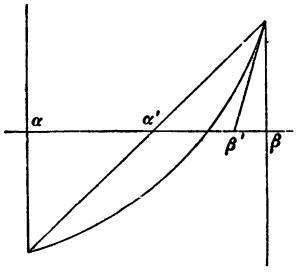
$$\alpha' = \alpha - \frac{f(\alpha)}{f'(\beta)}, \quad \beta' = \beta - \frac{f(\beta)}{f'(\beta)},$$

im zweiten Falle

$$\alpha' = \alpha - \frac{f(\alpha)}{f'(\alpha)}, \quad \beta' = \beta - \frac{f(\beta)}{f'(\alpha)},$$

und (α', β') ist ein engeres Intervall, in dem die gesuchte Wurzel liegt.

Fig. 8.



Es ist kaum nötig, die beiden anderen Fälle, in denen $f'(x)$ im Intervall negativ ist, noch besonders zu betrachten.

Man kann aber auch, um zwei neue Grenzen zu erhalten, mit noch besserem Erfolge die Newtonsche Methode mit der Interpolationsmethode verbinden, wie die Fig. 8 zeigt.

Man erhält dann als die beiden neuen Grenzen:

$$\alpha' = \alpha - \frac{f(\alpha)(\beta - \alpha)}{f(\beta) - f(\alpha)} = \frac{\alpha f(\beta) - \beta f(\alpha)}{f(\beta) - f(\alpha)},$$

$$\beta' = \beta - \frac{f(\beta)}{f'(\beta)}.$$

Wir fassen die Resultate in folgendem Satze zusammen:

Ist ein Intervall (α, β) abgegrenzt, in dem $f(x)$ einmal, $f'(x)$ und $f''(x)$ gar nicht das Zeichen wechseln, und ist β der Endpunkt des Intervalls, in dem $f(\beta)$ und $f''(\beta)$ dasselbe Vorzeichen haben, gleichviel, ob β kleiner oder größer als α ist, so erhält man ein engeres Intervall (α', β') von denselben Eigenschaften, wenn man

$$(4) \qquad \alpha' = \alpha - \frac{f(\alpha)(\beta - \alpha)}{f(\beta) - f(\alpha)}, \quad \beta' = \beta - \frac{f(\beta)}{f'(\beta)}$$

setzt.

Um diese Resultate der geometrischen Anschauung analytisch zu beweisen, nehmen wir an, es sei im Intervall $f'(x)$ und $f''(x)$ positiv, also $\alpha < \beta$, und

$$f(\alpha) < 0, \quad f(\beta) > 0.$$

Wir beschränken uns auf die Betrachtung dieses einen Falles, da die drei anderen genau in derselben Weise zu behandeln sind.

Wir bilden die Funktion

$$\varphi(x) = \frac{f(\beta) - f(x)}{\beta - x}$$

und deren erste Derivierte

$$\varphi'(x) = \frac{-(\beta - x) f'(x) + f(\beta) - f(x)}{(\beta - x)^2}.$$

Der Zähler dieses Ausdruckes

$$\psi(x) = -(\beta - x) f'(x) + f(\beta) - f(x)$$

verschwindet für $x = \beta$, und seine Derivierte ist

$$\psi'(x) = -(\beta - x) f''(x),$$

also im Intervall negativ. Daraus folgt, daß $\psi(x)$ im Intervall mit wachsendem x abnimmt und daher, weil es für den oberen Grenzwert $x = \beta$ verschwindet, positiv bleibt. Folglich ist auch $\varphi'(x)$ positiv und $\varphi(x)$ wächst im Intervall mit wachsendem x beständig. Wir haben demnach:

$$(5) \qquad \frac{f(\beta) - f(\alpha)}{\beta - \alpha} < \frac{f(\beta) - f(x)}{\beta - x} < f'(\beta)$$

$$\alpha < x < \beta.$$

Setzen wir nun

$$\alpha' = \alpha - \frac{f(\alpha)(\beta - \alpha)}{f(\beta) - f(\alpha)}, \quad \beta' = \beta - \frac{f(\beta)}{f'(\beta)},$$

so wird

$$(6) \qquad \beta' - \alpha' = f(\beta) \left\{ \frac{\beta - \alpha}{f(\beta) - f(\alpha)} - \frac{1}{f'(\beta)} \right\},$$

und nach (5) folgt, daß diese Differenz positiv ist, also

$$\alpha < \alpha' < \beta' < \beta.$$

Setzt man dann in (5)

$$x = \alpha' \quad \text{und} \quad x = \beta'$$

und beachtet, daß

$$\beta - \alpha' = \frac{(\beta - \alpha) f(\beta)}{f(\beta) - f(\alpha)}, \quad \beta - \beta' = \frac{f(\beta)}{f'(\beta)},$$

so ergibt sich:

$$f(\alpha') < 0, \quad f(\beta') > 0.$$

Es ist also die gesuchte Wurzel zwischen α' und β' enthalten, und das Intervall (α', β') ist kleiner als (α, β).

Setzen wir in diesen Ausdrücken α', β' an Stelle von α, β, so erhalten wir ein neues, noch engeres Intervall usf.

Was die Konvergenz dieses Verfahrens betrifft, so können wir uns darüber folgendermaßen vergewissern. Wir setzen nach (6):

$$(7) \quad \frac{\beta' - \alpha'}{\beta - \alpha} = \Phi(\alpha, \beta) = \frac{f(\beta)\{f(\alpha) - f(\beta) + (\beta - \alpha) f'(\beta)\}}{(\beta - \alpha) f'(\beta)\{f(\beta) - f(\alpha)\}}.$$

Zähler und Nenner sind hier durch $(\beta - \alpha)^2$ teilbar, und wenn wir den gemeinsamen Faktor $(\beta - \alpha)^2$ wegheben und ξ, η für α, β setzen, so erhalten wir einen Ausdruck von der Form

$$\Phi(\xi, \eta) = \frac{\psi_1(\xi, \eta)}{\psi_2(\xi, \eta)},$$

worin ψ_1 und ψ_2 ganze rationale Funktionen der beiden Variablen ξ, η sind.

Ist x die im Intervall (α, β) gelegene Wurzel, so werden die beiden Funktionen ψ_1 und ψ_2 nach unseren Voraussetzungen über $f(x)$ nicht gleich Null, wenn

$$(8) \quad \alpha \leqq \xi < x, \quad x < \eta \leqq \beta.$$

Die Funktion $\Phi(\xi, \eta)$ ist für $\xi = \alpha$, $\eta = \beta$, aber auch für jedes andere Wertpaar in dem Intervall (α, β), wofür $f(\xi)$ negativ, $f(\eta)$ positiv ist, kleiner als 1, da jedes solche Intervall (ξ, η) an Stelle von (α, β) genommen werden könnte.

Da nun $\Phi(\xi, \eta)$ eine stetige Funktion der beiden Veränderlichen ξ, η ist, solange diese in dem Bereich (8) bleiben, so muß in diesem Bereich die Funktion $\Phi(\xi, \eta)$ einen Maximumwert haben, und dieser muß kleiner als 1 sein, weil $\Phi(\xi, \eta)$ auch noch in den Grenzfällen $\xi = x$, $\eta = x$ kleiner als 1 bleibt.

Es läßt sich also ein positiver echter Bruch Θ angeben, so daß

$$\Phi(\xi, \eta) < \Theta$$

ist, solange ξ, η dem Bereich (8) angehören. Dann folgt aus (7):

$$\beta' - \alpha' < (\beta - \alpha)\Theta.$$

Ebenso folgt, wenn wir auf dieselbe Weise von dem Intervall (α', β') zu einem engeren Intervall (α'', β'') fortschreiten,

$$\beta'' - \alpha'' < (\beta' - \alpha')\,\Theta',$$

worin Θ' dieselbe Bedeutung für α', β' hat, wie Θ für α, β. Da aber α', β' dem Bereich (8) angehören, so ist Θ' nicht größer als Θ und statt Θ' kann auch Θ gesetzt werden. Es folgt also:

$$\beta'' - \alpha'' < (\beta - \alpha)\,\Theta^2,$$

und so schließen wir weiter

$$\beta^{(\nu)} - \alpha^{(\nu)} < (\beta - \alpha)\,\Theta^{\nu}.$$

Die Intervalle nehmen also mindestens so stark ab, wie die Glieder einer fallenden geometrischen Progression.

Als Beispiel mag die Gleichung dienen:

$$x^3 - 2x^2 - 2 = 0,$$

die eine Wurzel zwischen

$$\alpha = 2{,}25 \quad \text{und} \quad \beta = 2{,}36$$

hat.

Man erhält aus den Formeln (4) für

$$\beta' = 2{,}359\,31 \ldots \quad \alpha' = 2{,}359\,298 \ldots,$$

so daß ein in der vierten Dezimale genauer Wert der Wurzel

$$2{,}3593$$

ist. Für diesen Wert selbst ist, wie eine genauere Rechnung ergibt, $f(x)$ noch negativ, so daß er für α' genommen werden kann. Der nächste Schritt der Annäherung ergibt

$$2{,}359\,304.$$

Auf demselben Prinzip, das der Newtonschen Auflösungsmethode zugrunde liegt, beruhen auch die Methoden von W. G. Horner und Joseph Horner, die wir hier nur erwähnen können. Bei beiden ist das Ziel eine zweckmäßige Anordnung der Rechnung, und, besonders in der ersten, ein dem dekadischen Zahlensystem angepaßter Algorithmus, nach Analogie der Divisionsregeln der niederen Arithmetik. Ein schnelles und sicheres Rechnen nach diesen Methoden erfordert viel Übung[1]).

---

[1]) W. G. Horner, A new method of solving numerical equations of all orders, by continuous approximation. Communicated by Davies Gilbert. Philosophical Transactions 1819. Joseph Horner, Approximation to the roots of algebraic equations in a series of aliquot parts. Quarterly Journal of Mathematics, Vol. III, 1860.

§ 41.

Die Näherungsmethode von Daniel Bernoulli und verwandte Methoden.

Die Methode zur genäherten Auflösung einer Gleichung, die von Daniel Bernoulli herrührt[1]), beruht darauf, daß, wenn man eine Reihe reeller Größen hat, die Potenzen der größten unter ihnen um so mehr die gleich hohen Potenzen der übrigen überwiegen werden, je höher die Potenzen sind.

Sind $\alpha, \beta, \gamma \ldots$ beliebige reelle oder komplexe Größen, so jedoch, daß der absolute Wert von α größer ist, als der absolute Wert aller übrigen, daß also die absoluten Werte der Brüche $\beta:\alpha, \gamma:\alpha \ldots$ echte Brüche sind, so ist

$$(1) \qquad \frac{\alpha^m + \beta^m + \gamma^m + \cdots}{\alpha^{m-1} + \beta^{m-1} + \gamma^{m-1} + \cdots} = \alpha \frac{1 + \left(\frac{\beta}{\alpha}\right)^m + \left(\frac{\gamma}{\alpha}\right)^m + \cdots}{1 + \left(\frac{\beta}{\alpha}\right)^{m-1} + \left(\frac{\gamma}{\alpha}\right)^{m-1} + \cdots},$$

und je größer m wird, um so mehr wird sich dieser Ausdruck, wie die zweite Darstellung zeigt, der Grenze α nähern.

Sind $\alpha, \beta, \gamma \ldots$ die Wurzeln einer algebraischen Gleichung, so ist die linke Seite von (1) der Quotient der $m$ten und $(m-1)$ten Potenzsumme, und wir erhalten also den Satz:

> Der Quotient der $m$ten und $(m-1)$ten Potenzsumme nähert sich mit wachsendem m der absolut größten unter den Wurzeln.

Nimmt man m negativ an, und setzt α absolut kleiner als $\beta, \gamma \ldots$ voraus, so folgt auf die gleiche Weise:

> Der Quotient der $-m$ten und $-(m+1)$ten Potenzsumme nähert sich mit wachsendem m der absolut kleinsten unter den Wurzeln.

Da man die Potenzsummen als symmetrische Funktionen durch die Koeffizienten berechnen kann, so braucht man nur ein hinlänglich großes m zu nehmen, um einen angenäherten Wert der absolut größten und absolut kleinsten Wurzel zu erhalten.

---

[1]) D. Bernoulli, Commentarii Petropolit., Bd. III.

Wenn aber unter den Wurzeln der Gleichung solche vorkommen, die denselben absoluten Wert haben, und dieser der größte oder der kleinste ist, so ist diese Methode nicht anwendbar, also z. B. nicht auf reelle Gleichungen, bei denen ein Paar imaginärer Wurzeln vorkommt, das die reellen an absolutem Wert übertrifft.

Wenn aber die größte Wurzel zwar einzeln vorkommt, aber die nächstfolgende nicht viel übertrifft, so wird das Verfahren nur langsam einige Genauigkeit geben.

Jacobi hat die Bernoullische Methode nach einer Richtung ergänzt[1]).

Nehmen wir an, es seien

$$x_1, \quad x_2, \quad \ldots x_k$$

unter den Wurzeln $x_1, x_2, \ldots x_n$ solche, deren absolute Werte größer sind, als die absoluten Werte der folgenden $x_{k+1}, x_{k+2}, \ldots x_n$ und setzen

$$\begin{aligned} p_m &= x_1^m + x_2^m + \cdots + x_k^m \\ p_{m+1} &= x_1^{m+1} + x_2^{m+1} + \cdots + x_k^{m+1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ p_{m+k} &= x_1^{m+k} + x_2^{m+k} + \cdots + x_k^{m+k}, \end{aligned} \tag{2}$$

so werden, wenn wir

$$(x - x_1)(x - x_2) \ldots (x - x_k) = x^k + A_1 x^{k-1} + \cdots + A_k \tag{3}$$

setzen, die Koeffizienten $A_1, A_2, \ldots A_k$ den Gleichungen genügen:

$$\begin{aligned} p_{m+k} &+ A_1 p_{m+k-1} + A_2 p_{m+k-2} + \cdots + A_k p_m &&= 0 \\ p_{m+k+1} &+ A_1 p_{m+k} + A_2 p_{m+k-1} + \cdots + A_k p_{m+1} &&= 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ p_{m+2k-1} &+ A_1 p_{m+2k-2} + A_2 p_{m+2k-3} + \cdots + A_k p_{m+k-1} &&= 0. \end{aligned} \tag{4}$$

Wenn man den Gleichungen (4) noch die Gleichung

$$x^k + A_1 x^{k-1} + A_2 x^{k-2} + \cdots + A_k = 0$$

zugesellt und die $A_1, A_2, \ldots A_k$ eliminiert, so ergibt sich die Gleichung $k$ten Grades:

$$\begin{vmatrix} x^k, & x^{k-1}, & \ldots 1 \\ p_{m+k}, & p_{m+k-1}, & \cdots p_m \\ p_{m+k+1}, & p_{m+k}, & \cdots p_{m+1} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ p_{m+2k-1}, & p_{m+2k-2} & \cdots p_{m+k-1} \end{vmatrix} = 0,$$

[1]) Observatiunculae etc. Werke, Bd. 3, S. 280.

deren Wurzeln $x_1, x_2, \ldots x_k$ sind. Die Größen $p_m, p_{m+1} \ldots$ können aber mit um so größerer Genauigkeit durch die entsprechenden Potenzsummen aller Wurzeln ersetzt werden, je größer m ist.

So bekommt man z. B. für $k = 2$:

$$x^2 + \frac{p_m p_{m+3} - p_{m+1} p_{m+2}}{p_{m+1}^2 - p_m p_{m+2}} x + \frac{p_{m+2}^2 - p_{m+1} p_{m+3}}{p_{m+1}^2 - p_m p_{m+2}} = 0,$$

eine Gleichung, die man anwenden kann, wenn bei einer reellen Gleichung ein Paar konjugierter Wurzeln den absolut größten Wert haben.

Wenn man die Bernoullische Methode auf die Summe der negativen Potenzen anwendet, so erhält man, wie schon bemerkt, eine Annäherung an die absolut kleinste Wurzel. Man kann aber, durch Verlegung des Anfangspunktes, jede Wurzel zur absolut kleinsten machen, wozu freilich die Kenntnis eines bis zu einem gewissen Grade genäherten Wertes nötig ist.

Eine Formel, die für alle Fälle ausreicht, hat Fr. Meyer gegeben[1]). Es seien $\alpha_1, \alpha_2, \alpha_3, \ldots \alpha_n$ die Wurzeln der Gleichung $f(x) = 0$, die reell oder imaginär sein können. Man suche einen Punkt p in der x-Ebene, so daß sich um p als Mittelpunkt ein Kreis beschreiben läßt, der nur einen, aber einen beliebigen der Punkte $\alpha_1, \alpha_2, \ldots \alpha_n$ enthält, etwa den Punkt α_1, daß also die absoluten Werte von

$$\frac{p - \alpha_1}{p - \alpha_2}, \quad \frac{p - \alpha_1}{p - \alpha_3}, \ldots \frac{p - \alpha_1}{p - \alpha_n} \tag{5}$$

echte Brüche sind. Solche Punkte existieren immer; sie müssen nötigenfalls nach der Sturmschen Methode gefunden werden. Wählt man außerdem noch eine beliebige Funktion $\varphi(x)$, die mit $f(x)$ keinen gemeinsamen Teiler hat, z. B. $\varphi(x) = 1$, und bildet die symmetrischen Funktionen

$$y_m = \frac{\frac{\alpha_1 \varphi(\alpha_1)}{(p - \alpha_1)^m} + \frac{\alpha_2 \varphi(\alpha_2)}{(p - \alpha_2)^m} + \cdots + \frac{\alpha_n \varphi(\alpha_n)}{(p - \alpha_n)^m}}{\frac{\varphi(\alpha_1)}{(p - \alpha_1)^m} + \frac{\varphi(\alpha_2)}{(p - \alpha_2)^m} + \cdots + \frac{\varphi(\alpha_n)}{(p - \alpha_n)^m}}, \tag{6}$$

so nähern sich diese mit wachsendem m der Grenze α_1, und zwar um so schneller, je kleiner die absoluten Werte der Brüche (5) bereits sind.

---

[1]) Mathematische Annalen, Bd. 33.

Die Richtigkeit hiervon zeigt sich sofort, wenn man die Formel (6) in der Weise schreibt:

$$y_m = \frac{\alpha_1 \varphi(\alpha_1) + \alpha_2 \varphi(\alpha_2) \left(\frac{p-\alpha_1}{p-\alpha_2}\right)^m + \cdots + \alpha_n \varphi(\alpha_n) \left(\frac{p-\alpha_1}{p-\alpha_n}\right)^m}{\varphi(\alpha_1) + \varphi(\alpha_2) \left(\frac{p-\alpha_1}{p-\alpha_2}\right)^m + \cdots + \varphi(\alpha_n) \left(\frac{p-\alpha_1}{p-\alpha_n}\right)^m}$$

Auf einem ähnlichen Gedanken, wie die Bernoullische Näherung, beruht auch ein Verfahren, das von Gräffe angegeben und von Encke weiter ausgebildet ist[1]), und das sich besonders zu einer praktischen Durchführung der numerischen Rechnungen eignet.

Sind, wie oben, $\alpha, \beta, \gamma \ldots$ die Wurzeln einer Gleichung und ist α die absolut größte unter ihnen, so daß keine der anderen der Wurzel α an absolutem Wert gleichkommt, so hat

$$\sqrt[m]{\alpha^m + \beta^m + \gamma^m + \cdots} \tag{7}$$

mit unbegrenzt wachsendem m den Wert α zur Grenze; es ist sogar die Konvergenz gegen α noch besser, als bei dem im vorigen Paragraphen betrachteten Quotienten

$$\frac{\alpha^m + \beta^m + \gamma^m + \cdots}{\alpha^{m-1} + \beta^{m-1} + \gamma^{m-1} + \cdots},$$

wie man erkennt, wenn man nach dem polynomischen Lehrsatz:

$$\sqrt[m]{\alpha^m + \beta^m + \gamma^m + \cdots} = \alpha \sqrt[m]{1 + \left(\frac{\beta}{\alpha}\right)^m + \left(\frac{\gamma}{\alpha}\right)^m + \cdots}$$

$$= \alpha \left[1 + \frac{1}{m}\left(\frac{\beta}{\alpha}\right)^m + \frac{1}{m}\left(\frac{\gamma}{\alpha}\right)^m + \cdots\right]$$

setzt.

Die absolut kleinste unter den Wurzeln erhält man nach demselben Prinzip als Grenzwert von

$$\frac{1}{\sqrt[m]{\frac{1}{\alpha^m} + \frac{1}{\beta^m} + \frac{1}{\gamma^m} + \cdots}}, \tag{8}$$

und wenn p ein beliebiger Wert ist, so erhält man die dem Wert p am nächsten liegende Wurzel als Grenzwert von

---

[1]) Gräffe, Die Auflösung der höheren numerischen Gleichungen, als Beantwortung einer von der Berliner Akademie gestellten Preisfrage, Zürich 1837. Encke, Crelles Journal, Bd. 22 (1841).

$$(9) \qquad p + \frac{1}{\sqrt[m]{\frac{1}{(\alpha - p)^m} + \frac{1}{(\beta - p)^m} + \cdots}}.$$

Wenn man die Funktion $f(x)$ durch die Substitution

$$y = \frac{1}{x}, \quad y = \frac{1}{x - p}$$

transformiert, so gehen die Ausdrücke (8), (9) aus dem Ausdruck (7) hervor, auf den wir uns daher jetzt beschränken wollen.

Wendet man die Formel (7) auf eine reelle Gleichung an, so ist $\alpha^m + \beta^m + \gamma^m + \cdots$ eine reelle Zahl. Existiert eine absolut größte Wurzel, so muß sie reell sein, da eine imaginäre Wurzel bei einer reellen Gleichung auch die konjugiert imaginäre Zahl, die den gleichen absoluten Wert hat, nach sich zieht. Es ist also, wenn α die absolut größte Wurzel ist, die $m$te Wurzel in (7) reell zu nehmen. Ist aber m eine gerade Zahl, so muß man, um über das Vorzeichen zu entscheiden, noch wissen, ob α positiv oder negativ ist. Nötigenfalls wird darüber durch Einsetzen des gefundenen Näherungswertes in die gegebene Gleichung entschieden.

Ebenso wird man, wenn man die Formel (7) auf eine imaginäre Gleichung anwendet, entscheiden müssen, welche der verschiedenen $m$ten Wurzeln die richtige Annäherung gibt.

Die Ausdrücke

$$s_m = \alpha^m + \beta^m + \gamma^m + \cdots$$

lassen sich besonders einfach dann berechnen, wenn man für m die aufeinander folgenden Potenzen von 2, also $m = 2, 4, 8, 16 \ldots$ setzt, und liefern dann sehr gute Resultate.

Es ist s_2 der negative Koeffizient der $(n-1)$ten Potenz der Unbekannten in der Gleichung, deren Wurzeln die Quadrate der Wurzeln von $f(x)$ sind. Diese Gleichung wird aber leicht auf folgende Weise gebildet.

Man fasse in $f(x)$ die Glieder mit geraden und mit ungeraden Potenzen von x zusammen und setze demnach

$$f(x) = \varphi(x^2) + x\psi(x^2).$$

Es ist dann

$$f_1(x) = \varphi(x)^2 - x\psi(x)^2 = 0$$

die Gleichung, deren Wurzeln die Quadrate der Wurzeln von $f(x)$ sind. Behandelt man $f_1(x)$ ebenso, so erhält man eine

Gleichung, deren Wurzeln die vierten Potenzen der Wurzeln von $f(x)$ sind usf.

Die Ausführung dieser Rechnung ist sehr einfach und führt oft nach wenigen Schritten zu einer guten Näherung.

Wir betrachten ein Beispiel.

Wir nehmen die Gleichung

$$x^3 - 2x - 2 = 0, \tag{10}$$

die eine reelle positive und zwei imaginäre Wurzeln hat. Die reelle Wurzel ist größer als 1,7, und da das Produkt aller drei Wurzeln gleich 2 und $(1{,}7)^3 > 2$ ist, so ist der absolute Wert der beiden imaginären Wurzeln kleiner als die reelle Wurzel. Also ist die reelle Wurzel die absolut größte und die Gräffesche Näherungsmethode muß auf sie führen. Man bekommt nun die Gleichungen, deren Wurzeln die zweite, vierte, achte ... Potenz der Gleichung (4) sind:

$$\begin{aligned} x^3 - 4x^2 + 4x - 4 &= 0 \\ x^3 - 8x^2 - 16x - 16 &= 0 \\ x^3 - 96x^2 - 16^2 &= 0 \end{aligned}$$

und

$$x = \sqrt[8]{96} = 1{,}7692\ldots$$

ist bereits ein in den vier ersten Dezimalen genauer Wert. Der nächste Schritt würde kein anderes Resultat ergeben (weil die erste Potenz der Unbekannten in der letzten Gleichung fehlt); aber der darauf folgende ergibt den in der sechsten Dezimale genauen Wert

$$x = \sqrt[32]{85032960} = 1{,}769293\ldots$$

§ 42.

Trigonometrische Auflösung kubischer Gleichungen.

Wir besprechen nun noch einige auf Gleichungen von speziellen Formen anwendbare Methoden der numerischen Auflösung. Das Ziel dieser Methoden ist, die allgemein verbreiteten Tafeln der trigonometrischen Funktionen und der Logarithmen für die Auflösung von Gleichungen nutzbar zu machen. Wir wenden uns zunächst zur Betrachtung der kubischen Gleichungen, die wir immer in der reduzierten Form

$$x^3 + ax + b = 0 \tag{1}$$

annehmen, worin a und b reelle Zahlen sind. Da die Vertauschung von x mit $-x$ gleichbedeutend ist mit der Vertauschung von b mit $-b$, so können wir uns auf die Annahme beschränken, daß b negativ sei, und es bleiben dann noch drei verschiedene Fälle zu betrachten. Wir setzen $b = -g$ und, je nachdem a positiv oder negativ ist, $a = \pm e$. Dann haben wir

$$\text{a)}\quad x^3 + ex - g = 0$$

$$\text{b)}\quad x^3 - ex - g = 0, \qquad \frac{e^3}{27} < \frac{g^2}{4}$$

$$\text{c)}\quad x^3 - ex - g = 0, \qquad \frac{e^3}{27} > \frac{g^2}{4}.$$

Die Grenzfälle, daß eine der Größen e, g, $4e^3 - 27g^2$ verschwindet, schließen wir aus.

Um die Cardanische Formel anzuwenden, setzen wir

$$R = \frac{g^2}{4} \pm \frac{e^3}{27}, \tag{2}$$

wo das obere Zeichen im Falle a), das untere in den Fällen b) und c) gilt. In den Fällen a) und b) ist nur eine Wurzel reell, in dem Falle c) (dem Casus irreducibilis) sind alle drei Wurzeln reell. Die Cardanische Formel gibt:

$$x = \sqrt[3]{\frac{g}{2} + \sqrt{R}} + \sqrt[3]{\frac{g}{2} - \sqrt{R}}. \tag{3}$$

Wir verfahren nun so in den drei Fällen:

a) Wir führen einen Winkel ϑ ein, den wir zwischen 0 und 90° wählen können, durch die Gleichung:

$$\frac{g}{2} = \sqrt{\frac{e^3}{27}}\operatorname{cotg}\vartheta, \tag{4}$$

also

$$\sqrt{R} = \sqrt{\frac{e^3}{27}}\,\frac{1}{\sin\vartheta},$$

$$x = \sqrt{\frac{e}{3}}\left(\sqrt[3]{\operatorname{cotg}\vartheta + \frac{1}{\sin\vartheta}} + \sqrt[3]{\operatorname{cotg}\vartheta - \frac{1}{\sin\vartheta}}\right)$$

$$= \sqrt{\frac{e}{3}}\left(\sqrt[3]{\operatorname{cotg}\frac{\vartheta}{2}} - \sqrt[3]{\operatorname{tg}\frac{\vartheta}{2}}\right),$$

und wenn wir noch einen Winkel φ aus

$$\operatorname{tg}\varphi = \sqrt[3]{\operatorname{tg}\frac{\vartheta}{2}} \tag{5}$$

bestimmen, so ergibt sich:

$$x = 2\sqrt{\frac{e}{3}}\operatorname{cotg} 2\varphi. \tag{6}$$

Die Winkel ϑ, φ und zuletzt x findet man aus den logarithmisch trigonometrischen Tafeln nach den Formeln (4), (5), (6).

Um die imaginären Wurzeln zu erhalten, ersetzt man $\operatorname{tg}\varphi$ durch $\varrho\operatorname{tg}\varphi$, wenn ϱ eine imaginäre dritte Einheitswurzel bedeutet.

b) Im zweiten Falle bestimmen wir den Winkel ϑ, gleichfalls im ersten Quadranten, aus

$$\frac{g}{2} = \sqrt{\frac{e^3}{27}}\frac{1}{\sin\vartheta} \tag{7}$$

$$\sqrt{R} = \sqrt{\frac{e^3}{27}}\operatorname{cotg}\vartheta$$

$$x = \sqrt{\frac{e}{3}}\left(\sqrt[3]{\operatorname{cotg}\frac{\vartheta}{2}} + \sqrt[3]{\operatorname{tg}\frac{\vartheta}{2}}\right),$$

$$\operatorname{tg}\varphi = \sqrt[3]{\operatorname{tg}\frac{\vartheta}{2}}, \tag{8}$$

$$x = 2\sqrt{\frac{e}{3}}\frac{1}{\sin 2\varphi}. \tag{9}$$

c) Im letzten Falle endlich, wo drei reelle Wurzeln vorhanden sind, setzen wir:

$$\frac{g}{2} = \sqrt{\frac{e^3}{27}}\cos\vartheta \tag{10}$$

$$\sqrt{R} = i\sqrt{\frac{e^3}{27}}\sin\vartheta$$

$$x = \sqrt{\frac{e}{3}}\left(\sqrt[3]{\cos\vartheta + i\sin\vartheta} + \sqrt[3]{\cos\vartheta - i\sin\vartheta}\right),$$

oder nach dem Moivreschen Satze:

$$x = 2\sqrt{\frac{e}{3}}\cos\frac{\vartheta}{3}. \tag{11}$$

Nimmt man ϑ wieder im ersten Quadranten, so erhält man für die beiden anderen gleichfalls reellen Wurzeln:

$$-2\sqrt{\frac{e}{3}}\cos\frac{\pi+\vartheta}{3}, \quad -2\sqrt{\frac{e}{3}}\cos\frac{\pi-\vartheta}{3}.$$

Alle diese Formeln sind für die logarithmische Rechnung eingerichtet.

§ 43.

Die Gaußsche Methode der Auflösung trinomischer Gleichungen.

Gauß hat eine Methode angegeben, um die Wurzeln einer Gleichung, die nur drei Glieder enthält, in einfacher Weise numerisch aufzulösen. Solche Gleichungen kommen häufig vor und umfassen als Spezialfälle alle quadratischen und die reduzierten kubischen Gleichungen.

Es wird zunächst von einer solchen Gleichung, deren allgemeine Form

$$x^{m+n} + ax^m + b = 0 \tag{1}$$

ist, nur die positive Wurzel, wenn sie existiert, gesucht; die etwaige negative ergibt sich, wenn man x durch $-x$ ersetzt.

Nach den Vorzeichen von a, b hat man drei Fälle zu unterscheiden, da, wenn beide Vorzeichen positiv sind, keine positive Wurzel vorhanden ist. Wir betrachten also, indem wir mit e und g positive Zahlen bezeichnen, die drei Fälle:

$$\begin{aligned} &\text{a)}\quad x^{m+n} + ex^m - g = 0,\\ &\text{b)}\quad x^{m+n} - ex^m - g = 0,\\ &\text{c)}\quad x^{m+n} - ex^m + g = 0. \end{aligned}$$

In den beiden ersten Fällen haben wir nach dem Cartesischen Lehrsatz je eine positive Wurzel, im dritten können zwei oder keine positive Wurzel vorhanden sein.

Wir setzen nun mit Gauß

$$\lambda = \frac{g^n}{e^{m+n}},$$

und suchen die drei Gleichungen a), b), c) durch passende Substitutionen auf die Form

$$\sin^2\Theta + \cos^2\Theta = 1$$

zurückzuführen. Wir setzen:

$$\begin{aligned} &\text{a)}\quad \frac{x^{m+n}}{g} = \sin\Theta^2,\ \frac{ex^m}{g} = \cos\Theta^2,\quad \lambda = \frac{\sin\Theta^{2m}}{\cos\Theta^{2m+2n}},\\ &\text{b)}\quad gx^{-m-n} = \sin\Theta^2,\ ex^{-n} = \cos\Theta^2,\quad \lambda = \frac{\sin\Theta^{2n}}{\cos\Theta^{2m+2n}},\\ &\text{c)}\quad \frac{x^n}{e} = \sin\Theta^2,\ \frac{gx^{-m}}{e} = \cos\Theta^2,\quad \lambda = \sin\Theta^{2m}\cos\Theta^{2n}. \end{aligned} \tag{2}$$

Die letzte Gleichung ergibt die Unterscheidung der beiden Fälle, in denen die Gleichung c) zwei oder keine positive Wurzel hat

Man erhält nämlich für das Maximum von $\sin \Theta^{2m} \cos \Theta^{2n}$ nach den Regeln der Differentialrechnung

$$\frac{m^m n^n}{(m+n)^{m+n}},$$

das für $\operatorname{tg} \Theta^2 = m:n$ erreicht wird. Wenn also λ unter dieser Grenze liegt, so haben wir in c) zwei reelle Wurzeln, sonst keine. In den Formeln (2) ist λ eine gegebene Größe, und man hat nur aus den Tafeln den Winkel Θ zu suchen, der diesen Gleichungen genügt. Wenn man noch gar keine Kenntnis über die ungefähre Lage dieses Winkels hat, so ist es zweckmäßig, für die erste Annäherung eine etwa von Grad zu Grad fortschreitende, auf zwei Dezimalen abgekürzte Tafel zu benutzen, um den so gefundenen Wert mit Hilfe genauerer Tafeln zu verbessern.

Gauß benutzt nicht die trigonometrischen Tafeln, sondern die von ihm zuerst eingeführten Additions- und Subtraktionslogarithmen. Wir wollen dies an einem der Fälle in der Kürze zeigen. Die Einzelheiten für die praktische Anwendung der Methode sind in der Gaußschen Abhandlung zu suchen[1]).

Die Tafeln der Additions- und Subtraktionslogarithmen, wie sie zuerst von Gauß eingeführt und berechnet sind, und wie sie sich jetzt auch in den gebräuchlichen Tabellenwerken finden, geben zu drei Zahlen, die größer als 1 sind:

$$a, \quad b = 1 + \frac{1}{a}, \quad c = 1 + a,$$

die Briggschen Logarithmen:

$$A, \quad B, \quad C.$$

Ist Θ ein Winkel im ersten Oktanten, so kann man setzen:

$$(3) \quad a = \operatorname{cotg} \Theta^2, \quad b = \frac{1}{\cos \Theta^2}, \quad c = \frac{1}{\sin \Theta^2}, \quad 0 < \Theta < 45^0,$$

und wenn Θ im zweiten Oktanten liegt:

$$(4) \quad a = \operatorname{tg} \Theta^2, \quad b = \frac{1}{\sin \Theta^2}, \quad c = \frac{1}{\cos \Theta^2}, \quad 45^0 < \Theta < 90^0.$$

Wir wollen dies auf den Fall b) anwenden; dabei ist zu unterscheiden, ob λ kleiner oder größer als 2^m ist, weil davon

[1]) Beiträge zur Theorie der algebraischen Gleichungen, zweite Abteilung (1849). Gauß' Werke, Bd. III, S. 85.

abhängt, ob Θ im ersten oder zweiten Oktanten liegt. Es sind also wieder zwei Fälle zu unterscheiden:

$$\alpha) \quad \lambda < 2^m,$$

$$\lambda = \frac{b^m}{a^n} = \frac{c^m}{a^{m+n}} = \frac{b^{m+n}}{c^n},$$

$$x^{m+n} = g c, \quad x^n = e b, \quad x^m = \frac{g a}{e}. \tag{5}$$

$$\beta) \quad \lambda > 2^m,$$

$$\lambda = a^{m+n} b^m = a^n c^m = \frac{c^{m+n}}{b^n},$$

$$x^{m+n} = g b, \quad x^n = e c, \quad x^m = \frac{g}{e a}.$$

Im Falle α) würde man also

$$\log \lambda = m B - n A \tag{6}$$

setzen und danach aus der Tafel die zusammengehörigen Werte von A und B aufsuchen. Hat man diese gefunden, so ergibt sich:

$$m \log x = A + \log g - \log e. \tag{7}$$

Um den Gebrauch dieser Formeln an einem Beispiel zu erläutern, wollen wir die Gleichung betrachten:

$$x^3 - 2x - 2 = 0,$$

also $e = g = 2$, $m = 1$, $n = 2$, $\lambda = \frac{1}{2}$ setzen. Die Formel (6) gibt also

$$2A - B = 0{,}3010300 \tag{8}$$

und (7)

$$\log x = A. \tag{9}$$

Um den ersten Näherungswert von A zu finden, benutzt man einen kleinen Auszug aus der Tafel:

$$\left| \begin{array}{ll} A = 0, & B = 0{,}301 \\ A = 0{,}1, & B = 0{,}254 \\ A = 0{,}2, & B = 0{,}212 \\ A = 0{,}3, & B = 0{,}176 \end{array} \right|.$$

Es ergibt sich daraus:

A	$2A - B$	Fehler
0,2	0,188	$-0{,}113$
0,3	0,424	$+0{,}123$

Man kann hierauf die Interpolationsmethode anwenden, um einen genaueren Wert von A zu erhalten, indem man den Gesamtfehler von 0,236 nach Verhältnis der Teilfehler auf beide Werte von A verteilt, also

$$A = 0{,}2 + \frac{0{,}1}{0{,}236}\, 0{,}113 = 0{,}247 \ldots$$

setzt.

Mit Hilfe der siebenstelligen Tafel von Zach erhält man nun:

A	B	$2A - B$	Fehler
0,247	0,1948581	0,2991419	−0,0018881
0,248	0,1944969	0,3015031	+0,0004731

und durch abermalige Anwendung der Interpolation

$$A = 0{,}2477996$$

oder

$$x = 1{,}769292.$$

Eine etwas andere Anordnung der Rechnung ist den „Tafeln zur Berechnung der reellen Wurzeln sämtlicher trinomischer Gleichungen“ von Gundelfinger (Leipzig 1896) zugrunde gelegt.

Dort wird z. B. die Gleichung b) in die Form gesetzt:

$$\frac{x^{m+n}}{g} = \frac{e}{g}\, x^m + 1.$$

Ist dann

$$\frac{x^{m+n}}{g} = 10^B, \qquad \frac{e}{g}\, x^m = 10^A,$$

so folgt

$$1 + 10^A = 10^B \tag{10}$$

und

$$\log x = \frac{B + \log g}{m+n} = \frac{A + \log g - \log e}{m}, \tag{11}$$

also

$$A - \frac{m}{m+n} B = \log e - \frac{n}{m+n} \log g. \tag{12}$$

Die Tafeln von Gundelfinger enthalten nun außer den zusammengehörigen Werten von A, B noch die Werte von $A - \mu B$ für ein echtgebrochenes μ von $\mu = 0{,}05$, $\mu = 0{,}1$, $\mu = 0{,}15$, ..., so daß man, wenn man die rechte Seite von (12) berechnet hat, aus der Tafel direkt genäherte Werte von A, B findet, die dann nach (11) genäherte Werte von $\log x$ ergeben.

Die Methode von Gauß ist auch auf Gleichungen mit mehr als zwei Gliedern ausgedehnt worden[1]).

Gauß hat auch für die Berechnung der imaginären Wurzeln einer trinomischen Gleichung ein Verfahren angegeben, das wir hier noch kurz besprechen wollen. Wir machen die Annahme reeller Koeffizienten, obwohl die Methode auch auf den allgemeinen Fall anwendbar ist. Wir wollen die Gleichung in die Form setzen:

$$x^{m+n} - e x^m - g = 0, \tag{13}$$

brauchen uns aber hier nicht auf positive Werte von e und g zu beschränken.

Wir setzen

$$x = r e^{i\vartheta},$$

und erhalten, indem wir den imaginären Teil in (13) für sich Null setzen,

$$r^n = \frac{e \sin m\vartheta}{\sin (n + m)\vartheta}; \tag{14}$$

solcher Gleichungen erhalten wir aber noch zwei, wenn wir (13) mit x^{-m} und mit x^{-m-n} multiplizieren. Dies gibt

$$r^{n+m} = \frac{-g \sin m\vartheta}{\sin n\vartheta}, \tag{15}$$

$$r^m = \frac{-g \sin (n + m)\vartheta}{e \sin n\vartheta}, \tag{16}$$

und von diesen drei Gleichungen folgt jede aus den beiden anderen. Wenn man r aus zwei von ihnen eliminiert und wie oben

$$\lambda = \frac{g^n}{e^{n+m}}$$

setzt, so ergibt sich:

$$\lambda = (-1)^n \frac{\sin n\vartheta^n \sin m\vartheta^m}{\sin (n + m)\vartheta^{n+m}}. \tag{17}$$

Man kann, da man von den beiden konjugierten Wurzeln nur die eine zu berechnen braucht, ϑ auf den ersten Quadranten beschränken, muß aber dann unter Umständen auch r negativ annehmen.

Wie man nun aus der Formel (17) mittels der trigonometrischen Tafeln den Winkel ϑ und dann aus einer der drei

[1]) A. Wiener, Schlömilchs Zeitschr., Bd. 31; R. Mehmke, ebenda, Bd. 36.

Formeln (14), (15), (16) den zugehörigen Wert von r berechnet, das wollen wir nun an dem vorhin betrachteten Beispiel

$$x^3 - 2x - 2 = 0$$

noch zeigen.

Hier wird die Gleichung (17)

$$\frac{1}{2} = \frac{\sin 2\vartheta^2 \sin\vartheta}{\sin 3\vartheta^3} \tag{18}$$

oder

$$3\log\sin 3\vartheta - 2\log\sin 2\vartheta - \log\sin\vartheta = 0{,}301\,030\,0, \tag{19}$$

woraus zunächst ersichtlich ist, daß der Winkel ϑ in dem Intervall von 0 bis 60° liegt. Man findet zunächst nach wenigen Versuchen, daß ϑ zwischen 33° und 34° liegt, und wenn man dann für diese beiden Werte die Differenz

$$3\log\sin 3\vartheta - 2\log\sin 2\vartheta - \log\sin\vartheta - 0{,}301\,030\,0 \tag{20}$$

berechnet, so erhält man zunächst auf drei Dezimalen

$$0{,}026, \qquad -0{,}013,$$

und hieraus ergibt sich durch Interpolation der genauere Wert

$$\vartheta = 33°\,40'.$$

Hierauf berechnet man die Differenz (20) mit etwas größerer Genauigkeit, etwa auf fünf Stellen für einige Winkel in der Nähe der Werte 30° 40′, von Minute zu Minute fortschreitend, und findet aus der Vorzeichenänderung, daß ϑ zwischen

$$33°\,41' \text{ und } 33°\,42'$$

liegt. Wenn man nun für diese beiden Werte die Rechnung auf sieben Stellen durchführt, so ergibt sich wieder durch Interpolation der genauere Wert

$$\vartheta = 33°\ 41'\ 20{,}6''.$$

Aus den Formeln (15) oder (16) sieht man, daß hier r negativ ist, und man erhält r sehr einfach aus einer dieser Gleichungen. Man findet die Briggschen Logarithmen:

$$\log(-r) = 0{,}026\,614\,8$$
$$\log\cos\vartheta = 9{,}920\,154\,7 - 10$$
$$\log\sin\vartheta = 9{,}744\,046\,8 - 10.$$

Also sind die beiden imaginären Wurzeln:

$$-0{,}884\,646 \pm 0{,}589\,740\,i\,[1]).$$

[1]) Lüroth, Circ. mat. di Palermo. Tomo XXVII.

Achter Abschnitt.

Gruppen.

§ 44.

Definition der Gruppen.

Es sind hauptsächlich zwei große allgemeine Begriffe, von denen die moderne Algebra beherrscht wird. Die Existenz und Bedeutung dieser Begriffe konnte allerdings erst erkannt werden, nachdem die Algebra bis zu einem gewissen Grad fertig und zum Eigentum der Mathematiker geworden war. Erst dann konnte in ihnen das verbindende und führende Prinzip erkannt werden.

Es sind das die Begriffe der Gruppen und des Körpers, zu deren Erklärung wir jetzt fortschreiten. Der allgemeinere Begriff ist der der Gruppe, mit dem wir also beginnen.

Es sei irgend ein allgemeiner Begriff Z gegeben, für den es Einzeldinge, Repräsentanten in beliebiger Anzahl gibt. Wir können z. B. an die Zahl oder die rationale Funktion denken. Außerdem sei uns die Kraft gegeben, nach irgend einer Regel einem ersten und zweiten Repräsentanten von Z, einen dritten, den wir möglicherweise erst bilden müssen, in unzweideutiger Weise zuzuordnen. Die Repräsentanten heißen auch die Elemente der Menge Z.

Bei Zahlen denke man z. B. an eine der Rechenregeln, Multiplikation oder Addition. Diese Zuordnung nennen wir die Komposition des ersten und zweiten Elementes zu dem dritten. Zur Bezeichnung der Repräsentanten bedient man sich der Buchstaben und für die Komposition der arithmetischen Zeichen der Multiplikation und Gleichheit:

$$ab = c,$$

wobei indessen zu betonen ist, daß man nicht notwendig an die wirkliche Multiplikation denken muß. Die Komposition könnte

z. B. auch Addition oder Division sein, so daß nicht notwendig ab dasselbe Element c wie ba gibt.

Wir nehmen nun aus Z einen engeren Begriff P heraus, von dem wir fordern:

I. Sind a, b irgend zwei Elemente von P, so ist auch $ab = c$ ein Element von P.

II. Sind von den Elementen a, b, c aus P irgend zwei beliebig gegeben, so kann man das dritte immer und nur auf eine Weise so bestimmen, daß

$$ab = c$$

ist.

Sind a, b, c drei Elemente von P, so gehören nach I. auch (ab) und (bc) zu P und ebenso $(ab)c$ und $a(bc)$. Wir setzen als dritte Forderung voraus:

III. Das assoziative Gesetz

$$a(bc) = (ab)c,$$

und nennen P eine Gruppe oder, genauer gesagt, die Elemente Z, die den Forderungen 1, 2, 3 genügen, der Gruppe P angehörig.

Ist z. B. Z das Reich der Zahlen und die Kompositionsregel die wirkliche Multiplikationsregel, so bildet das System der rationalen positiven Zahlen (ohne die Null) eine Gruppe. Ist die Kompositionsregel die Addition oder die Subtraktion, so müssen die Null und die negativen Zahlen hinzugenommen werden. Es würde dann aber auch eine Gruppe zustande kommen, wenn man sich auf die ganzen Zahlen beschränken würde.

Der Gruppenbegriff bezieht sich also nicht auf die Elemente als solche, sondern es gehört noch ein bestimmtes Kompositionsgesetz dazu.

Aus dem assoziativen Gesetz III. folgt durch den Schluß der vollständigen Induktion, daß man immer zu demselben Resultat kommt, wenn man in einer beliebigen Reihe von Elementen aus P in endlicher Anzahl, a, b, c, d ... zuerst zwei benachbarte Elemente komponiert, dann wieder zwei benachbarte usw., bis die ganze Reihe auf ein Element reduziert ist, das mit $abcd$ ... bezeichnet wird. So ist z. B.:

$$\begin{aligned} abcd &= (ab)cd = [(ab)c]d = (ab)(cd) \\ &= a(bc)d = [a(bc)]d = a[(bc)d] \\ &= ab(cd) = (ab)(cd) = a[b(cd)]\,^{1}). \end{aligned}$$

Wir ziehen nun aus dieser Definition zunächst einige ganz allgemeine Folgerungen.

Nach II. gibt es für jedes gegebene b in P ein Element e gleichfalls in P, das der Bedingung

$$eb = b \tag{1}$$

genügt, und dies e ist von b unabhängig; denn aus (1) folgt für jedes c

$$ebc = bc,$$

und bc kann nach II. jedes Element in P bedeuten. Ebenso gibt es ein Element e', das für jedes b der Bedingung

$$be' = b \tag{2}$$

genügt. Dies Element e' ist aber von e nicht verschieden; denn setzen wir $b = e'$ in (1) und $b = e$ in (2), so folgt

$$ee' = e', \quad ee' = e,$$

also

$$e = e'.$$

Das Element e ändert nichts, wenn es mit irgendwelchen Elementen aus P komponiert wird, und wird die Einheit der Gruppe genannt. In vielen Fällen kann es ohne Mißverständnis geradezu mit „1“ bezeichnet werden.

Zu jedem Element a gibt es nach II. ein bestimmtes Element a^{-1}, das der Bedingung

$$a^{-1}a = e \tag{3}$$

genügt. Aus (3), (1) und (2) folgt

$$a^{-1}aa^{-1} = ea^{-1} = a^{-1} = a^{-1}e,$$

und folglich nach 3.

$$aa^{-1} = e. \tag{4}$$

Die beiden Elemente a, a^{-1} heißen zueinander entgegengesetzt oder reziprok. Sind a, b zwei Elemente der Gruppe und

$$c = ab, \tag{5}$$

so ist

$$c^{-1} = b^{-1}a^{-1}. \tag{6}$$

Es ist ein besonderer Fall, wenn bei der Komposition der Elemente einer Gruppe P das kommutative Gesetz gilt, d. h. wenn für je zwei Elemente a, b der Gruppe

$$ab = ba$$

ist. In den oben als Beispiele angeführten Fällen, in denen Z das Zahlenreich ist, wo die Komposition nach den gewöhnlichen Rechenregeln geschieht, tritt dieser Fall immer ein. Wir werden aber alsbald Fälle anderer Art kennen lernen.

IV. Gruppen, in denen das kommutative Gesetz gilt, heißen kommutative Gruppen oder auch Abelsche Gruppen.

Wenn sich die Elemente zweier Gruppen

$$a, b, c, d \ldots$$

und

$$a', b', c', d' \ldots$$

in der Weise gegenseitig eindeutig entsprechen, daß immer, wenn $ab = c$ ist, auch $a'b' = c'$ wird, so heißen die Gruppen isomorph, und es gilt der evidente Satz, daß zwei mit einer dritten isomorphe Gruppen untereinander isomorph sind. Man kann hiernach die untereinander isomorphen Gruppen zu einer Klasse von Gruppen zusammenfassen, die selbst wieder eine Gruppe ist, deren Elemente die Gattungsbegriffe sind, die man erhält, wenn man die entsprechenden Elemente der einzelnen isomorphen Gruppen zu einem Allgemeinbegriff zusammenfaßt. Die einzelnen untereinander isomorphen Gruppen sind dann als verschiedene Repräsentanten eines Gattungsbegriffes aufzufassen.

Eine Gruppe, die nur eine bestimmte endliche Anzahl von Elementen enthält, heißt eine endliche Gruppe, und wenn n die Anzahl der zu einer solchen Gruppe gehörigen Elemente ist, so heißt n der Grad der Gruppe.

Bei einer endlichen Gruppe P kann man die Forderung II. durch die einfachere ersetzen, daß ab nur dann gleich $a'b$ oder ab gleich ab' sein kann, wenn $a = a'$ oder $b = b'$ ist, weil alsdann xb oder ax zugleich mit x die ganze Gruppe P durchläuft.

Wenn ein System Q von Elementen aus P die Eigenschaft hat, daß das Kompositum je zweier Elemente aus Q wieder in Q enthalten ist, so ist Q für sich eine Gruppe und wird ein Teiler von P genannt.

Gibt es in P noch wenigstens ein Element, das nicht zu Q gehört, so heißt Q ein echter Teiler von P.

Das Einheitselement „1“ bildet für sich eine Gruppe, die als Teiler in jeder anderen Gruppe enthalten ist.

Es sei nun P irgend eine endliche oder unendliche Gruppe und

$$Q = Q_1 = a_1, a_2, a_3, \ldots$$

sei ein echter Teiler von P. Ist b ein nicht in Q enthaltenes Element von P, so sind die Elemente

$$Q_2 = a_1 b, \quad a_2 b, \quad a_3 b, \ldots$$

alle voneinander verschieden und alle nicht in Q enthalten; denn wenn etwa $a_2 b$ in Q enthalten wäre, so müßte auch, da Q eine Gruppe ist, $a_2^{-1} a_2 b = b$ in Q enthalten sein, gegen die Voraussetzung.

Ist mit Q_1 und Q_2 die ganze Gruppe P noch nicht erschöpft, so nehmen wir eines der noch übrigen Elemente, das wir mit c bezeichnen können, und bilden

$$Q_3 = a_1 c, \quad a_2 c, \quad a_3 c, \ldots,$$

und überzeugen uns leicht, daß die Elemente von Q_3 alle nicht nur voneinander, sondern auch von den Elementen von Q_1 und Q_2 verschieden sind. Denn wäre etwa $a_2 c = a_3 b$, so würde folgen, daß $c = a_2^{-1} a_3 b$ sein müßte; es ist aber $a_2^{-1} a_3$ in Q_1 enthalten, also mit einem der Elemente $a_1, a_2, a_3 \ldots$ identisch, und es wäre also c gegen die Voraussetzung in Q_2 enthalten. So fahren wir fort, die Systeme $Q_1, Q_2, Q_3, \ldots$ zu bilden, indem wir der Übereinstimmung der Bezeichnung wegen unter Q_1 die Gruppe Q selbst verstehen.

Diese Systeme $Q_2, Q_3, \ldots$ nennen wir die zu Q gehörigen Nebengruppen und bezeichnen sie durch

$$Q_2 = Qb, \quad Q_3 = Qc, \ldots$$

Es sind nun zwei Fälle möglich: entweder die Bildung der Nebengruppen $Q_1, Q_2, Q_3, \ldots$ geht ohne Ende weiter, oder es ist nach einer endlichen Zahl von Bildungen dieser Art die ganze Gruppe P erschöpft. Der letzte Fall, der uns hier hauptsächlich beschäftigt, tritt dann immer ein, wenn P eine endliche Gruppe ist. Wir nehmen jetzt eine endliche Zahl von Nebengruppen an und bezeichnen die letzte von ihnen mit

$$Q_j = Qg,$$

so daß wir auch symbolisch

$$P = Q_1 + Q_2 + Q_3 + \cdots + Q_j \tag{7}$$

setzen können.

Bisweilen werden wir auch das ganze System $Q_1, Q_2, Q_3, \ldots$ als ein System von Nebengruppen bezeichnen, und also die Darstellung (7) die Zerlegung von P in ein System von Nebengruppen nennen.

Ist die Anzahl der Nebengruppen zu Q endlich, so heißt ihre Anzahl, also die Zahl j, der Index des Teilers Q von P, und Q ein Teiler von P von endlichem Index. Dieser Index wird (nach Dedekind) durch das Symbol

$$j = (P, Q) \tag{8}$$

bezeichnet.

Wählt man aus jeder der Nebengruppen $Q_1, Q_2, \ldots Q_j$ ein Element $a, b, \ldots g$ beliebig aus, so erhält man ein volles Repräsentantensystem der Gruppe P nach Q, und man kann setzen:

$$P = Qa + Qb + Qc + \cdots + Qg.$$

Aus der Bildungsweise der Nebengruppen folgt noch, daß, wenn b, c irgend zwei Elemente aus P sind, die beiden Systeme Qb und Qc entweder ganz identisch sind oder kein gemeinsames Element enthalten. Denn ist $a_1 b = a_2 c$, worin a_1, a_2 zwei Elemente aus Q sind, so ist auch für jedes andere Element a aus Q:

$$a a_1 b = a a_2 c,$$

und wenn a die ganze Gruppe Q durchläuft, so durchläuft auch jedes der beiden $a a_1$ und $a a_2$ die ganze Gruppe Q; folglich sind Qb und Qc identisch.

Ist also e irgend ein Element aus P, so ist das System Qe mit einer der Nebengruppen $Qa, Qb, \ldots Qg$ identisch. Wenn zwei Nebengruppen Qb und Qc kein gemeinschaftliches Element enthalten, so enthalten die beiden Systeme $b^{-1}Q$ und $c^{-1}Q$, die wir gleichfalls Nebengruppen nennen, kein gemeinschaftliches Element. Denn ist $b^{-1}a_1 = c^{-1}a_2$ ein gemeinschaftliches Element der beiden letzten Systeme, so ist $a_1^{-1}b = a_2^{-1}c$ ein gemeinschaftliches Element von Qb und Qc. Zu $b^{-1}Q$ gehören alle zu den Elementen von Qb reziproken Elemente von P, und hieraus ergeben sich, wenn Q einen endlichen Index hat, die beiden gleichzeitig bestehenden Zerlegungen von P in j Nebengruppen:

$$\begin{aligned} P &= Q + Qb + Qc + \cdots + Qg \\ P &= Q + b^{-1}Q + c^{-1}Q + \cdots + g^{-1}Q. \end{aligned} \tag{9}$$

Ist R ein Teiler von P von endlichem Index, und Q ein anderer Teiler von P, der seinerseits R als Teiler enthält, so wird eine der Nebengruppen R_1 von R entweder ganz in Q enthalten sein, oder kein Element mit Q gemein haben. Es ist daher auch R ein Teiler von Q von endlichem Index $(Q, R) = k$. Ist ferner R_1 eine Nebengruppe zu R, und Q_1 eine Nebengruppe zu Q, so wird R_1 entweder ganz in Q_1 enthalten sein, oder kein einziges Element von Q_1 ist in R_1 enthalten. Es zerfällt also jede Nebengruppe Q_1 in eine endliche Zahl Nebengruppen R_1 und die Anzahl (P, Q) der Q_1 ist also gleichfalls endlich. Ist die Zerlegung von Q in die Nebengruppen nach R

$$Q = R_1 + R_2 + \cdots + R_k,$$

so erhält man irgend eine der Nebengruppen Q_1, wenn man ein Element a aus P passend auswählt, in der Form

$$Q_1 = R_1 a + R_2 a + \cdots + R_k a,$$

worin die Nebengruppen $R_1 a$, $R_2 a$, ... $R_k a$ alle voneinander verschieden sind. Es zerfällt daher jedes Q_1 in gleichviel Nebengruppen nach R, und wir erhalten den Satz:

$$(P, R) = (P, Q)\,(Q, R). \tag{10}$$

Wenn P eine endliche Gruppe vom Grade n ist, so können wir für R die aus dem einzigen Elemente 1 bestehende Gruppe nehmen, und dann wird (P, R) gleich dem Grade n von P. Ebenso wird (Q, R) gleich dem Grade m von Q, und wenn wir den Index des Teilers Q von P mit j bezeichnen, so geht (10) in die Form über:

$$n = jm,$$

und wir erhalten den Satz:

V. Der Grad einer endlichen Gruppe ist durch den Grad eines jeden seiner Teiler teilbar, und der Quotient beider Zahlen ist der Index des Teilers.

Die Nebengruppen haben nicht die Merkmale einer Gruppe. Denn damit zwei Elemente aus Q_1 bei der Zusammensetzung wieder ein Element von Q_1 ergeben, müßte etwa

$$a_1 b\, a_2 b = a_3 b,$$

also $a_1 b a_2 = a_3$, $b = a_1^{-1} a_3 a_2^{-1}$ sein. Es wäre also b der Voraussetzung entgegen in Q enthalten. Die Benennung Nebengruppe ist also nur uneigentlich zu verstehen.

Zwei Teiler Q, Q' von P haben immer das Element 1 miteinander gemein. Sie können aber auch noch andere Elemente gemeinschaftlich haben, und diese gemeinschaftlichen Elemente bilden eine Gruppe. Denn gehören die Elemente a und b sowohl zu Q als zu Q', so gilt wegen des Gruppencharakters von Q und Q' dasselbe von dem Kompositum ab. Und wenn a in Q und in Q' vorkommt, so ist auch a^{-1} in beiden enthalten. Diese gemeinschaftliche Gruppe nennen wir den größten gemeinschaftlichen Teiler von Q und Q' oder auch, nach einem Vorschlage von Study, mit einem geometrischen Anklange, den Durchschnitt von Q und Q'. Ebenso folgt, daß die Elemente, die irgend einer Anzahl von Teilern von P gemeinsam sind, eine Gruppe bilden, die wir ebenso als den größten gemeinschaftlichen Teiler oder den Durchschnitt aller dieser Gruppen bezeichnen.

In jeder Gruppe können wir nach folgendem Verfahren Teiler bilden.

Ein Teiler ist immer das Einheitselement für sich.

Bezeichnen wir die wiederholte Zusammensetzung eines Elementes mit sich selbst durch Potenzen, mit a^0 das Einheitselement, und mit a^{-r} die $r$te Potenz des Elementes a^{-1}, dann ist a^{-r} das mit a^r reziproke Element, und die Reihe der Elemente

$$\ldots a^{-2}, a^{-1}, 1, a, a^2, a^3 \ldots,$$

die alle der Gruppe P angehören, bildet eine Gruppe. Ist die Gruppe P endlich vom Grade n, so kann diese Reihe nicht lauter verschiedene Elemente enthalten. Ist also ein Element, das zum zweiten Male wiederkehrt,

$$a^\mu = a^{\mu+\alpha},$$

so folgt, daß $a^\alpha = 1$ sein muß, und wir nehmen an, daß α die kleinste positive Zahl ist, die dieser Bedingung genügt. Es ist dann, wenn $m = q\alpha$ ein beliebiges Vielfaches von α ist, $a^m = 1$, und umgekehrt: so oft $a^m = 1$ ist, muß m durch α teilbar sein; denn sonst könnte man $m = q\alpha + \alpha'$ setzen, worin α' positiv und kleiner als α ist, und es wäre $a^{\alpha'} = 1$, gegen die Voraussetzung. Es ist immer und nur dann

$$a^\mu = a^{\mu'},$$

wenn $\mu \equiv \mu' \pmod{\alpha}$.

Die Reihe

$$A = 1, a, a^2 \ldots a^{\alpha-1}$$

enthält dann α voneinander verschiedene Elemente von P, wobei alle Potenzen von a vertreten sind. Die Elemente A bilden aber

offenbar eine Gruppe vom Grade α, weil sich die Exponenten bei der Zusammensetzung einfach addieren. Diese Gruppe ist ein Teiler von P und also ist α ein Teiler von n. Es ist also für jedes Element $a^n = 1$.

Die Gruppe A heißt die Periode des Elementes a und α wird auch der Grad des Elementes a genannt.

Ist P eine Gruppe und Q ein Divisor von P mit den Elementen $a_1, a_2, a_3, \ldots$, ferner b ein nicht in Q enthaltenes Element von P, so ist die Nebengruppe Qb keine Gruppe. Dagegen bildet das System $b^{-1}Qb$ sicher eine Gruppe, weil

$$b^{-1}a_1 b \,.\, b^{-1}a_2 b = b^{-1}a_1 a_2 b$$

ist, und diese Gruppe ist mit Q isomorph. Die Gruppe $b^{-1}Qb$ heißt die durch b aus Q transformierte Gruppe. Gehört b selbst zu Q, so ist $b^{-1}Qb$ mit Q identisch. Nimmt man für b die verschiedenen Elemente von P, so erhält man eine ganze Schar solcher Gruppen, die wir die zu Q konjugierten Teiler von P oder auch kurz konjugierte Gruppen nennen. Ersetzt man b durch ein Element $b_1 = ab$ der Nebengruppe Qb, so ist $b_1^{-1}Qb_1$ mit $b^{-1}Qb$ identisch. Wenn also Q ein Teiler von P von endlichem Index ist, so ist die Anzahl der verschiedenen zu Q konjugierten Teiler jedenfalls endlich.

Es kann vorkommen, daß alle konjugierten Teiler miteinander identisch sind. Wir führen folgende Definition ein:

VI. Wenn Q ein Teiler von P ist, der mit seinen sämtlichen konjugierten Teilern identisch ist, so heißt Q ein Normalteiler von P[1]).

Die aus dem einzigen Element 1 gebildete Gruppe ist ein Normalteiler von jeder Gruppe. Wir erhalten ferner einen Normalteiler in dem größten gemeinschaftlichen Teiler R der sämtlichen mit irgend einem Teiler Q von P konjugierten Teiler.

Denn es ist schon oben bewiesen, daß R als der Durchschnitt mehrerer Teiler von P eine Gruppe ist.

Sind nun

$$Q, Q', Q'' \ldots \tag{11}$$

die zu Q konjugierten Teiler von P, und ist b irgend ein Element von P, dann ist das System der Gruppen

$$b^{-1}Qb, \quad b^{-1}Q'b, \quad b^{-1}Q''b \ldots \tag{12}$$

[1]) Auch ausgezeichnete oder invariante Untergruppe genannt.

von dem System (11) nicht verschieden. Wenn aber R der Durchschnitt der Gruppen (11) ist, so ist $b^{-1}Rb$ der Durchschnitt von (12) und folglich ist R mit $b^{-1}Rb$ identisch, d. h. R ist ein Normalteiler von P.

Ist N ein Normalteiler irgend einer Gruppe P, und b ein beliebiges Element in P, so ist

$$b^{-1}Nb = N \quad \text{oder} \quad Nb = bN. \tag{13}$$

Ist der Index (P, N) endlich und gleich μ, so können wir die μ Elemente $b_1, b_2, \ldots b_\mu$ so wählen, daß die Zerlegung von P in die Nebengruppen

$$\begin{aligned} P &= Nb_1 + Nb_2 + Nb_3 + \cdots + Nb_\mu \\ &= b_1N + b_2N + b_3N + \cdots + b_\mu N \end{aligned}$$

ergibt. Es ist also

$$N = N_1, \quad N_2 = Nb_2 = b_2N, \quad N_3 = Nb_3 = b_3N \ldots, \quad N_\mu = Nb_\mu = b_\mu N \tag{14}$$

das System der Nebengruppen.

Wir erwähnen noch folgenden Satz, dessen Richtigkeit sich unmittelbar aus dem Isomorphismus transformierter Gruppen ergibt:

VII. Ist Q ein Teiler von P, R ein Teiler von Q, so ist, wenn Q' und R' aus Q und R durch dasselbe Element von P transformiert sind, auch R' ein Teiler von Q', und wenn R Normalteiler von Q ist, so ist auch R' Normalteiler von Q'.

Es gilt ferner noch folgender Satz:

VIII. Ist Q ein Teiler von P, N ein Teiler von Q, und zugleich Normalteiler von P, so ist N auch Normalteiler von Q.

Das ist selbstverständlich, weil die den Normalteiler von N definierende Relation

$$a^{-1}Na = N \tag{15}$$

für jedes Element a von P, also auch für jedes Element von Q gilt.

Man darf aber nicht umgekehrt schließen, daß ein Normalteiler von Q auch immer Normalteiler von P sein müsse. Denn die Bedingung (15) könnte zwar für jedes Element a aus Q befriedigt sein, ohne für jedes Element aus P zu gelten.

Jede Gruppe hat sich selbst und das Einheitselement zu Normalteilern.

IX. Eine Gruppe, die außer diesen beiden keinen anderen Normalteiler hat, heißt einfach[1]).

§ 45.

Komposition der Teile.

Es sei jetzt P eine endliche oder nicht endliche Gruppe, und A, B irgend zwei Reihen von Elementen aus P (Gruppen oder nicht). Wir verstehen unter dem symbolischen Produkte AB die Elemente, die man erhält, wenn man je ein Element a von A mit je einem Element b von B nach der in P geltenden Vorschrift zu ab komponiert. Diese Art der Zusammensetzung von A und B zu AB wollen wir die Komposition der Teile (von P) nennen. Wir unterscheiden hier zwischen einem Teil und einem Teiler von P, so daß ein Teiler immer eine Gruppe sein soll was bei einem Teil nicht notwendig ist; auch bei einem Teil heißt die Anzahl der Elemente der Grad.

Man kann ebenso drei und mehr Teile $A, B, C \ldots$ von P komponieren, und es gilt bei der Komposition der Teile das assoziative Gesetz, wie unmittelbar daraus folgt, daß dieses Gesetz in P gilt. Danach ist die Bedeutung eines Kompositums $ABC\ldots$ eindeutig bestimmt.

In der Zusammensetzung AB kann einer der Teile A, B auch aus einem einzigen Elemente bestehen, und dann stimmt die Bezeichnung AB mit der oben für die Nebengruppen gebrauchten Bezeichnung überein.

1. Die notwendige und hinreichende Bedingung, daß A eine Gruppe oder ein Teiler von P sei, besteht hiernach in der Gleichung

$$AA = A. \tag{1}$$

Denn diese Gleichung besagt nichts weiter, als daß das Kompositum irgend zweier Elemente von A wieder in A enthalten sei.

---

[1]) Vgl. über die Definition der Gruppe: Huntington Simplified Definition of a Group. Bull. of the American mathematical Society, 1902.

2. Bezeichnen wir mit A^{-1} das System der zu den Elementen von A reziproken Elemente, so ist, wenn A eine Gruppe ist,

$$A^{-1} = A.$$

3. Ist A ein Normalteiler von P, so besteht für jeden beliebigen Teil B von P die Gleichung

$$AB = BA. \tag{2}$$

Denn ist b irgend ein Element aus B (also auch aus P), so ist für einen Normalteiler A von P nach § 44, (12) $Ab = bA$, woraus sich (2) ergibt.

4. Ist die Gruppe A ein Teiler von P, und B ein Teil von A, so sind auch AB und BA Teile von A. Ist auch B eine Gruppe, so ist

$$AB = A, \quad BA = A. \tag{3}$$

Denn wenn B eine Gruppe und Teiler der Gruppe A ist, und wenn a in A, b in B und also auch in A enthalten ist, so ist auch ab in A enthalten. A enthält also jedes Element von AB.

Weil aber zweitens B als Gruppe auch das Element 1 enthält, so enthält AB auch jedes Element von A, und also ist AB mit A identisch. Ebenso sieht man, daß BA mit A identisch ist.

Andererseits folgt aus jeder der Gleichungen (3), da A als Gruppe das Einheitselement enthält, daß jedes Element von B in A enthalten, also B ein Teiler von A ist.

5. Ist P eine endliche Gruppe und A ein Teiler von P, also eine Gruppe, so kann man ein System B, das im allgemeinen keine Gruppe ist, in P so auswählen, daß die Formel

$$P = AB$$

die Zerlegung von P in die Nebengruppen von A darstellt.

6. Bei der Komposition der Teile bildet das System der Nebengruppen zu einem Normalteiler von P selbst eine Gruppe, in der der Normalteiler N die Einheit bildet, und

$$Nb, \quad Nb^{-1}$$

entgegengesetzte Elemente sind.

Denn ist N ein Normalteiler von P mit den Nebengruppen

(4) $N_1 = N, \quad N_2 = Nb_2 = b_2 N, \quad N_3 = Nb_3 = b_3 N \ldots,$

so ist, da $NN = N$ [nach (1)],

(5) $$N_h N_k = Nb_h Nb_k = Nb_h b_k = Nb_l = N_l.$$

Bezeichnet man Nb_h^{-1} mit N_h^{-1}, so ist hiernach

(6) $$N_h N_h^{-1} = N, \quad N_k = N_h^{-1} N_l, \quad N_h = N_l N_k^{-1}.$$

Die Gruppe der Größen N_k nennen wir die Gruppe der Nebengruppen oder die zu N komplementäre Gruppe in bezug auf P. Wir bezeichnen sie nach dem Vorgange von Hölder[1]) mit P/N. Hat N einen endlichen Index (P, N) in bezug auf P, so ist die Gruppe P/N endlich, auch wenn P selbst nicht endlich ist, und der Grad von P/N ist gleich dem Index (P, N).

Sind A und B Gruppen in P, so ist ihr Durchschnitt D gleichfalls eine Gruppe. Das System AB wird aber nicht immer eine Gruppe sein. Wenn die Gruppen A, B und folglich D endlich sind, und von den Graden α, β, δ, so ist AB gleichfalls endlich und vom Grade $\alpha\beta:\delta$. Denn sind a, a_1 Elemente in A und b, b_1 Elemente in B, so ist dann und nur dann

$$ab = a_1 b_1,$$

wenn

$$a_1^{-1} a = b_1 b^{-1} = d,$$

in D enthalten ist. Dann ist aber

$$a_1 = ad^{-1}, \quad b_1 = db.$$

Wenn wir also a die Elemente von A, b die Elemente von B durchlaufen lassen, so wird in der Form ab jedes Element von AB, und jedes genau δmal dargestellt. Die Zahl der in AB enthaltenen verschiedenen Elemente ist also $\alpha\beta:\delta$. Hieraus leiten wir folgenden Satz ab:

7. Sind A, B endliche Gruppen in P, so ist AB dann und nur dann eine Gruppe, wenn

(7) $$AB = BA$$

ist.

Damit nämlich AB eine Gruppe sei, ist notwendig und hinreichend, daß zu je zwei Elementenpaaren a, b; a_1, b_1 aus A, B

[1]) Mathematische Annalen, Bd. 34. Das Zeichen erinnert an einen Quotienten, mit dem ja die komplementäre Gruppe eine gewisse Analogie hat.

sich ein drittes Elementenpaar a_2, b_2 bestimmen läßt, so daß

$$ab\,a_1 b_1 = a_2 b_2 \tag{8}$$

ist. Daraus folgt aber

$$b\,a_1 = a^{-1} a_2 b_2 b_1^{-1},$$

d. h. es muß ba_1 in AB enthalten sein. Da b und a_1 in ihrer Gruppe beliebig sind, so ist BA ein Teil von AB. Da aber die Grade AB und BA übereinstimmend $= \alpha\beta : \delta$ sind, so folgt, daß AB mit BA identisch ist. Ist umgekehrt die Bedingung 7. erfüllt, so kann jedes ba in die Form ab gesetzt werden, woraus sich die Relation (8.) ergibt. Damit ist also das Theorem 7. bewiesen.

Wir beschäftigen uns zunächst fast ausschließlich mit endlichen Gruppen. Wenn P eine solche Gruppe vom Grade n ist, und N ein Normalteiler von P vom Grade ν und Index μ, so ist

$$n = \mu\nu,$$

und die komplementäre Gruppe P/N ist vom Grade μ. Wir wollen diese oder eine damit isomorphe Gruppe mit Q bezeichnen und ihre Elemente, die den Nebengruppen N_1, N_2, N_3 ... entsprechen, mit A, B, C, ... Jedem dieser Elemente entsprechen ν Elemente der Gruppe P, etwa

dem Elemente A die Elemente $a_1, a_2 \ldots a_\nu$,
dem Elemente B die Elemente $b_1, b_2 \ldots b_\nu$,
.....

Dies Entsprechen ist dann derart, daß jedes zusammengesetzte Element ab dem zusammengesetzten Elemente AB entspricht. Denn die Elemente A und B entsprechen den Nebengruppen Na und Nb, und es ist, weil N ein Normalteiler ist,

$$NaNb = Nab.$$

Es gilt also hier bei der Zusammensetzung der A, B, ... einerseits und der a, b, ... andererseits dasselbe Gesetz, wie bei den isomorphen Gruppen, nur mit dem Unterschiede, daß jedem Elemente A nicht bloß ein Element a, sondern mehrere entsprechen. Diese Tatsache gibt Anlaß, den Begriff des Isomorphismus zu erweitern, wie es durch folgende Definition geschieht:

Man nennt eine endliche Gruppe P mit den Elementen a, b, c, ... (mehrstufig) isomorph mit einer Gruppe Q mit den Elementen A, B, C, ..., wenn

beide Gruppen so aufeinander bezogen werden können, daß jedem der Elemente A, B, C, ... ein oder mehrere der Elemente a, b, c, ... entsprechen, und zwar so, daß jedes der Elemente von P einem und nur einem der Elemente von Q entspricht, und daß, wenn a und A, b und B einander entsprechen, ab dem Elemente AB entspricht.

Es läßt sich zunächst beweisen, daß jedem der Elemente A, B, C, ... eine gleiche Zahl von Elementen a, b, c, ... entspricht und daß also der Grad von Q ein Teiler des Grades von P ist. Denn es sei A das Einheitselement in Q, dem die Elemente a_1, a_2, ... a_ν in P entsprechen mögen. Diese letzteren Elemente müssen dann eine Gruppe vom Grade ν bilden, die wir mit N bezeichnen wollen, weil nach der Definition des Isomorphismus das Element aa_1 dem Elemente $AA = A$ entsprechen muß. Ist dann b ein dem Elemente B entsprechendes Element von P, so müssen wiederum alle Elemente Nb demselben Elemente B aus Q entsprechen. Denn jedes ab muß dem Elemente $AB = B$ entsprechen. Sind b_1, b zwei dem Elemente B entsprechende Elemente, so entspricht $b_1 b^{-1} = a$ dem Elemente A, und ist also in N enthalten, also ist $b_1 = ab$ in Nb enthalten. Durch Nb sind also alle Elemente, die dem Elemente B entsprechen, erschöpft, und jedem Elemente von Q entsprechen ν Elemente von P. Der Isomorphismus heißt ν-stufig. Jedes Element $b^{-1}ab$ entspricht aber nach 7. gleichfalls dem Einheitselemente A und also ist $b^{-1}Nb = N$, d. h. N ist Normalteiler von P, und Q ist einstufig isomorph mit P/N.

Wir sehen also, daß es einen anderen mehrstufigen Isomorphismus als den zwischen der komplementären Gruppe eines Normalteilers und der Gesamtgruppe nicht gibt. Man hat den Begriff des Isomorphismus noch dahin erweitert, daß man zwei Gruppen, die zu einer dritten Gruppe μ- und ν-stufig isomorph sind, μ-ν-stufig isomorph zueinander nennt. Bei weitem der wichtigste ist der einstufige Isomorphismus, den wir daher als Isomorphismus schlechtweg bezeichnen, während ein mehrstufiger Isomorphismus immer durch einen Zusatz kenntlich gemacht werden soll[1]).

[1]) Vgl. C. Jordan, Traité des substitutions. Netto, Substitutionentheorie. Die Bezeichnung „μ-stufig“ rührt von Netto her. Nach Jordan

§ 46.

Zerlegung einer Gruppe nach zwei Teilern.

Es seien jetzt Q und R irgend zwei Teiler einer endlichen Gruppe P. Die Elemente von P sollen mit a, die von Q mit b und die von R mit c bezeichnet sein, so daß sowohl die b als die c unter den a enthalten sind.

Wir zerlegen zunächst, indem wir

$$(P, R) = j$$

setzen und die Elemente $a_1, a_2 \dots a_j$ passend aus P auswählen, P in die Nebengruppen nach R:

$$P = Ra_1 + Ra_2 + \cdots + Ra_j. \tag{1}$$

Nun betrachten wir einen (nach der Komposition der Teile gebildeten) Komplex

$$P_k = Ra_k Q \tag{2}$$

und bemerken, daß, wenn in diesem Komplex ein Element aus einer Nebengruppe Ra vorkommt, zugleich alle Elemente dieser Nebengruppe darin enthalten sind. Denn unter dieser Voraussetzung muß a die Form haben:

$$a = ca_kb,$$

und dann ist auch jedes Element $c'a$ in derselben Form enthalten, weil $c'c$ in R enthalten ist. Um die Anzahl der in P_k enthaltenen Nebengruppen Ra zu ermitteln, betrachten wir zunächst die Elemente e aus Q, die der Bedingung

$$Ra_k = Ra_ke \tag{3}$$

genügen. Diese Gleichung besagt aber, daß a_ke in Ra_k oder daß e in $a_k^{-1}Ra_k$ enthalten ist. Demnach ist die Bedingung (3) erfüllt für alle Elemente e, die der Gruppe Q_k (Durchschnitt von Q mit $a_k^{-1}Ra_k$) angehören. Daraus ergibt sich allgemein, daß zwei Nebengruppen Ra_kb, Ra_kb_1 dann und nur dann einander gleich sind, wenn $b_1b^{-1} = e$ in Q_k enthalten, also $b_1 = eb$ ist, und daß also, wenn man b in Ra_kb die ganze Gruppe Q durchlaufen läßt, jede der Nebengruppen Ra, die überhaupt darunter vorkommt,

heißt der einstufige Isomorphismus „holoedrischer", der mehrstufige „meroedrischer" Isomorphismus. Der einstufige Isomorphismus wird bisweilen auch als Äquivalenz bezeichnet und jede Art von Isomorphismus als Homomorphismus.

gleich oft, nämlich so oft, als der Grad von Q_k beträgt, erzeugt wird, also:

1. Die Anzahl der in P_k vorkommenden Nebengruppen Ra ist gleich dem Index (Q, Q_k), und die Anzahl der Elemente oder der Grad von P_k ist, wenn r den Grad von R bedeutet, gleich $r(Q, Q_k)$.

Ist nun a_h ein Element in P, so wird das System

$$P_h = Ra_hQ$$

nur dann ein Element mit P_k gemein haben, wenn a_h von der Form ca_kb ist, und dann ist P_h mit P_k identisch. Hieraus ergibt sich, daß man eine bestimmte Anzahl von Elementen $a_1, a_2, a_3 \ldots$ aus P so auswählen kann, daß keine zwei der Systeme

$$P_1 = Ra_1Q, \quad P_2 = Ra_2Q, \quad P_3 = Ra_3Q \ldots$$

ein Element gemein haben, während sie in ihrer Gesamtheit alle Elemente von P enthalten, so daß man

$$P = Ra_1Q + Ra_2Q + Ra_3Q + \cdots \tag{4}$$

setzen kann. Man kann die Elemente $a_1, a_2, \ldots$ aus den in der Formel (1) vorkommenden entnehmen.

Jedes der Systeme $P_1, P_2, P_3, \ldots$ enthält eine nach 1. zu bestimmende Anzahl von Nebengruppen Ra, und da die Gesamtzahl dieser Nebengruppen $= (P, R)$ ist, so ergibt sich, wenn $Q_1, Q_2, Q_3 \ldots$ die Durchschnitte von Q mit den Gruppen

$$a_1^{-1}Ra_1, \quad a_2^{-1}Ra_2, \quad a_3^{-1}Ra_3, \ldots$$

bedeuten, die Relation

$$(P, R) = (Q, Q_1) + (Q, Q_2) + (Q, Q_3) + \cdots \tag{5}$$

und die Zahlen (Q, Q_1), (Q, Q_2), (Q, Q_3), ... sind Teiler des Grades q der Gruppe Q.

In dieser Betrachtung kann man die beiden Teiler R und Q von P auch miteinander vertauschen und aus $a = ca_kb$ folgt $a^{-1} = b^{-1}a_k^{-1}c^{-1}$. Wenn also a in Ra_kQ vorkommt, so ist a^{-1} in $Qa_k^{-1}R$ enthalten und umgekehrt. Es ist folglich:

$$P = Qa_1^{-1}R + Qa_2^{-1}R + Qa_3^{-1}R + \cdots^1). \tag{6}$$

---

[1]) Die Zerlegung einer Gruppe P nach zweien ihrer Teiler rührt von Dedekind her. Göttinger Nachrichten 1894.

§ 47.

Die Kompositionsreihe und der Satz von C. Jordan.

Eine endliche Gruppe P heißt einfach, wenn sie außer sich selbst und der Einheit keinen Normalteiler hat, zusammengesetzt, wenn noch andere Normalteiler vorhanden sind.

Ein Normalteiler, der keinen anderen Normalteiler über sich hat, der also nicht Teil eines anderen echten Normalteilers von P ist, heißt ein größter Normalteiler von P[1]).

Ist P_1 ein größter Normalteiler von P, P_2 ein größter Normalteiler von P_1, P_3 ein größter Normalteiler von P_2 usf., so heißt die Reihe von Gruppen

$$P, P_1, P_2, P_3 \ldots 1, \tag{1}$$

die notwendig mit der aus dem Einheitselemente allein gebildeten Gruppe 1 endigt, eine Kompositionsreihe der Gruppe P.

Wir wollen die Grade der Gruppen (1) mit $n, n_1, n_2, n_3 \ldots 1$ bezeichnen, und die Quotienten $n:n_1$, $n_1:n_2$, $n_2:n_3 \ldots$, also die Indices von P_1 in bezug auf P, P_2 in bezug auf P_1 usf. mit $\nu_1, \nu_2, \nu_3 \ldots$ Es ist dann nach der in § 44 eingeführten Bezeichnung der Indices

$$\nu_1 = (P, P_1), \quad \nu_2 = (P_1, P_2), \quad \nu_3 = (P_2, P_3), \ldots$$

und die Reihe der ganzen Zahlen

$$\nu_1, \nu_2, \nu_3, \ldots \tag{2}$$

nennen wir eine Indexreihe der Gruppe P.

Die größten Normalteiler P_1, P_2, $P_3, \ldots$ können für eine gegebene Gruppe P im allgemeinen auf mehrere verschiedene Arten ausgewählt werden, und danach können auch die Gradzahlen $n_1, n_2, n_3, \ldots$ und die Indices $\nu_1, \nu_2, \nu_3, \ldots$ verschieden ausfallen. Es gilt aber der folgende schöne und wichtige Satz von C. Jordan[2]):

1. Satz von der Konstanz der Indexreihe. Wie auch die Kompositionsreihe einer Gruppe P gewählt sein mag, die Indexreihe ist, von der Anordnung abgesehen, immer dieselbe.

---

[1]) Ausgezeichnete Maximaluntergruppe nach anderer Bezeichnung.

[2]) C. Jordan, Traité des substitutions, Paris 1870. Netto, Substitutionentheorie, Leipzig 1882. Hölder, Zurückführung einer beliebigen algebraischen Gleichung auf eine Kette von Gleichungen; Mathematische Annalen, Bd. 34 (1888). Pierpont, On the invariance of the factors of composition; American Journal of Mathematics, Vol. XVIII.

Der Beweis beruht auf der Betrachtung der im § 45 eingeführten komplementären Gruppen.

Es seien Q und Q_1 Normalteiler von P und zugleich Q_1 ein Teiler (also nach § 44, 3. Normalteiler) von Q.

Dann gilt der Satz:

2. Die Gruppe Q/Q_1 ist ein Normalteiler der Gruppe P/Q_1, und der Index von Q/Q_1 in bezug auf P/Q_1 ist gleich dem Index von Q in bezug auf P. In Zeichen:

$$(P/Q_1,\ Q/Q_1) = (P, Q). \tag{3}$$

Um seine Richtigkeit einzusehen, braucht man nur die Bildungsweise der einzelnen Gruppen genauer zu betrachten.

Man zerlegt Q in die Nebengruppen zu Q_1, d. h. man setzt, wenn $b_1, b_2, \dots b_{\nu_1}$ passend bestimmte Elemente in Q sind und $\nu_1 = (Q, Q_1)$ der Index von Q_1 in bezug auf Q ist,

$$Q = Q_1 b_1 + Q_1 b_2 + \cdots Q_1 b_{\nu_1}. \tag{4}$$

Die Elemente der Gruppe Q/Q_1 sind:

$$Q_1 b_1,\ Q_1 b_2, \dots Q_1 b_{\nu_1}. \tag{5}$$

Wenn wir P in die Nebengruppen zu Q_1 zerlegen, so kommen darunter sicher die Elemente (5) vor, und folglich ist Q/Q_1 ein Teiler von P/Q_1.

Es ist noch nachzuweisen, daß dieser Teiler ein Normalteiler ist.

Bezeichnen wir mit a irgend ein Element von P, so sind alle Elemente von P/Q_1 in der Form $Q_1 a$ enthalten, und $Q_1 a^{-1}$ ist zu $Q_1 a$ reziprok (§ 45, 6.). Wir haben also nur zu zeigen, daß, wenn b irgend ein Element in Q ist, jedes Element

$$Q_1 a^{-1} Q_1 b Q_1 a \tag{6}$$

in Q/Q_1, d. h. in der Form $Q_1 b$ enthalten ist.

Weil aber Q_1 Normalteiler von Q und von P ist, so ist $b Q_1 = Q_1 b$, $a^{-1} Q_1 = Q_1 a^{-1}$, und daher:

$$Q_1 a^{-1} Q_1 b Q_1 a = Q_1 a^{-1} b a, \tag{7}$$

und weil $a^{-1} b a$ zugleich mit b in Q vorkommt, so ist $Q_1 a^{-1} b a$ mit einem der Elemente (5) identisch.

Bezeichnen wir die Grade von P, Q, Q_1 mit n, q, q_1, so sind die Grade von P/Q, P/Q_1, Q/Q_1:

$$(P, Q) = \frac{n}{q}, \qquad (P, Q_1) = \frac{n}{q_1}, \qquad (Q, Q_1) = \frac{q}{q_1},$$

und der Index:

$$(P/Q_1,\ Q/Q_1) = \frac{n}{q_1} : \frac{q}{q_1} = \frac{n}{q(P,\ Q)},$$

was die Formel (3) ist.

Daneben besteht folgender Satz:

3. Ist Q_1 ein Normalteiler von P vom Grade q_1, und hat die Gruppe P/Q_1 selbst einen Teiler M vom Grade m, so hat P einen Teiler Q vom Grade $q = mq_1$. Zugleich ist Q_1 ein Normalteiler von Q, und es ist $M = Q/Q_1$. Ist M Normalteiler von P/Q_1, so ist auch Q Normalteiler von P.

Es sei μ der Index von Q_1 in bezug auf P, und P sei in die Nebengruppen zerlegt:

$$P = Q_1 e_1 + Q_1 e_2 + \cdots + Q_1 e_\mu, \tag{8}$$

so daß $e_1, e_2, \ldots e_\mu$ passend gewählte Elemente in P und

$$Q_1 e_1,\ Q_1 e_2 \ldots Q_1 e_\mu \tag{9}$$

die Elemente von P/Q_1 sind. Wenn nun in der Gruppe P/Q_1 ein Teiler M enthalten ist, so muß dieser aus einem Teile der Elemente (9) bestehen, etwa aus

$$Q_1 b_1,\ Q_1 b_2 \ldots Q_1 b_m, \tag{10}$$

und wenn dies System eine Gruppe bilden soll. so muß für irgend zwei Elemente b_i, b_k das Produkt

$$Q_1 b_i Q_1 b_k = Q_1 b_i b_k \tag{11}$$

wieder in (10) enthalten, also etwa $= Q_1 b_h$ sein. Wenn also das System $b_1, b_2, \ldots b_m$ mit B bezeichnet wird, so läßt sich nach der Komposition der Teile die Relation (11) so schreiben:

$$Q_1 B Q_1 B = Q_1 B, \tag{12}$$

und so besagt sie nach § 45, 1., daß die in

$$Q = Q_1 B \tag{13}$$

enthaltenen Elemente von P eine Gruppe bilden, die die Gruppe Q_1 als Teiler enthält und selbst ein Teiler von P ist. Daß Q_1 Normalteiler von Q ist, folgt aus § 44, 3., weil Q_1 als Normalteiler von P vorausgesetzt ist, und aus der Darstellung (10) der Gruppe M ergibt sich, daß $M = Q/Q_1$ ist.

Es ist noch zu zeigen, daß, wenn M Normalteiler von P/Q_1 ist, auch Q Normalteiler von P ist. Dies folgt so:

Ist e irgend ein Element in P und $Q_1 b_k$ ein Element in M, so folgt aus der Annahme, daß M Normalteiler von P/Q_1 ist:

$$Q_1 e^{-1} Q_1 b_k Q_1 e = Q_1 b_h,$$

was wir auch nach der Komposition der Teile durch die Formel

$$Q_1 e^{-1} Q_1 B Q_1 e = Q_1 B \tag{14}$$

ausdrücken können. Da nun Q_1 Normalteiler von P ist, so kann Q_1 mit jedem der anderen Faktoren vertauscht werden, so daß aus (14) folgt:

$$e^{-1} Q_1 B e = Q_1 B,$$

oder

$$e^{-1} Q e = Q, \tag{15}$$

wodurch ausgedrückt ist, daß Q Normalteiler von P ist.

Aus dem Theorem 2 und 3 ergibt sich das Korollar:

4. Ein Normalteiler Q von P ist dann und nur dann ein größter Normalteiler, wenn die komplementäre Gruppe P/Q einfach ist.

Denn wenn Q_1 nicht größter Normalteiler von P ist, so gibt es einen Normalteiler Q über Q_1 und nach 2. ist P/Q_1 nicht einfach. Ist andererseits P/Q_1 nicht einfach, so ist nach 3. Q_1 nicht größter Normalteiler.

Wenn Q und Q' zwei Normalteiler von P sind, so ist auch das symbolische Produkt QQ' ein Normalteiler.

Denn es ist, weil Q ein Normalteiler ist,

$$QQ' = Q'Q,$$

und dies ist nach § 45, 7. das Kennzeichen, daß QQ' eine Gruppe ist. Daß es ein Normalteiler von P ist, ergibt sich dann nach § 45, 3. aus

$$QQ'B = QBQ' = BQQ',$$

wenn B irgend ein Teil von P ist.

Ist nun Q ein größter Normalteiler von P, und Q' ein nicht in Q enthaltener Normalteiler von P, so ist QQ' ein Normalteiler von P, der sowohl Q als Q' in sich enthält und also von Q verschieden ist. Folglich ist, da es über Q keinen Normalteiler von P gibt, außer P selbst,

$$QQ' = P, \tag{16}$$

und ebenso muß auch

$$Q'Q = P \tag{17}$$

sein. Diese Sätze gelten, wenn nur eine der beiden Gruppen Q, Q' größter Normalteiler von P ist. Wir nehmen jetzt an, daß sie beide diese Eigenschaft haben, und verstehen unter R den Durchschnitt von Q und Q'. Diese Gruppe R ist dann ein Normalteiler von P sowohl als von Q und Q'. Denn ist a irgend ein Element in P, und c in Q sowohl als in Q' enthalten, so ist auch $a^{-1}ca$ in Q und in Q', also auch in R enthalten. Also ist $a^{-1}Ra = R$, und R ist Normalteiler von P und mithin Normalteiler von Q und von Q'. Um Q in die Nebengruppen von R zu zerlegen, wählen wir ein passendes System von Elementen $B = b_1, b_2, b_3, \ldots$ in Q, so daß $Rb_1, Rb_2, Rb_3, \ldots$ alle voneinander verschieden werden, und setzen

$$Q = RB. \tag{18}$$

Nun ist nach (17) $Q'Q = P$, also auch

$$P = Q'RB = Q'B. \tag{19}$$

Die Nebengruppen

$$Q'b_1,\ Q'b_2,\ Q'b_3, \ldots \tag{20}$$

sind aber alle voneinander verschieden. Denn wäre etwa

$$Q'b_2 = Q'b_1,$$

so wäre auch

$$Q'b_2b_1^{-1} = Q'.$$

Es wäre also $b_2b_1^{-1}$ sowohl in Q als in Q' und folglich auch in R enthalten. Dann aber wäre auch $Rb_2 = Rb_1$, was gegen die Voraussetzung ist.

Hiernach können wir die Elemente der zu R komplementären Gruppe in bezug auf Q und der zu Q' komplementären Gruppe in bezug auf P bilden. Wir erhalten diese beiden Gruppen:

$$\begin{aligned} Q/R &= Rb_1,\ Rb_2,\ Rb_3, \ldots \\ P/Q' &= Q'b_1,\ Q'b_2,\ Q'b_3, \ldots \end{aligned} \tag{21}$$

Daraus aber erkennt man leicht, daß diese Gruppen nicht nur von gleichem Grade sind, sondern daß sie isomorph sind. Denn es ist gleichzeitig

$$Rb_1Rb_2 = Rb_1b_2$$

und

$$Q'b_1Q'b_2 = Q'b_1b_2.$$

Ebenso kann man auch schließen, daß die beiden Gruppen

$$Q'/R,\quad P/Q$$

isomorph sind.

Die Gruppen P/Q und P/Q' sind aber, da Q, Q' größte Normalteiler sind, nach 4. einfach, und folglich sind auch die damit isomorphen Gruppen Q'/R und Q/R einfach, und folglich haben wir den Satz:

5. Sind Q und Q' zwei verschiedene größte Normalteiler von P, und ist R der Durchschnitt von Q und Q', so ist R größter Normalteiler sowohl von Q als von Q', und für die Indices gilt das Gesetz:

$$(Q, R) = (P, Q').$$

Damit haben wir die Grundlage gewonnen, um sehr einfach den Satz von der Konstanz der Indexreihe nachzuweisen.

Wir schreiben zur besseren Übersicht die verschiedenen in Vergleich zu ziehenden Kompositionsreihen so, daß wir den Index eines jeden Gliedes in bezug auf das vorangehende unter das Glied setzen.

Danach mögen irgend zwei gegebene Kompositionsreihen von P mit den zugehörigen Indices folgende sein:

$$\begin{matrix} P, & Q, & Q_1, & Q_2, & Q_3, \ldots \\ & \nu, & \nu_1, & \nu_2, & \nu_3, \ldots \end{matrix} \tag{22}$$

$$\begin{matrix} P, & Q', & Q_1', & Q_2', & Q_3', \ldots \\ & \nu', & \nu_1', & \nu_2', & \nu_3, \ldots \end{matrix} \tag{23}$$

Es sei nun R der Durchschnitt von Q und Q', und μ und μ' seien die Indices von R in bezug auf Q und Q'. Wegen des Isomorphismus der Gruppen P/Q und Q'/R ist dann $\mu' = \nu$, und aus dem entsprechenden Grunde $\mu = \nu'$. Wir bilden eine Kompositionsreihe von R mit der zugehörigen Indexreihe

$$\begin{matrix} R, & R_1, & R_2, & R_3, \ldots \\ & \mu_1, & \mu_2, & \mu_3, \ldots, \end{matrix} \tag{24}$$

und da nun R ein größter Normalteiler von Q und von Q' ist, so können wir daraus zwei neue Kompositionsreihen von P bilden, nämlich:

$$\begin{matrix} P, & Q, & R, & R_1, & R_2, & R_3, \ldots \\ & \nu, & \nu', & \mu_1, & \mu_2, & \mu_3, \ldots \end{matrix} \tag{25}$$

$$\begin{matrix} P, & Q', & R, & R_1, & R_2, & R_3, \ldots \\ & \nu', & \nu, & \mu_1, & \mu_2, & \mu_3, \ldots \end{matrix} \tag{26}$$

Die beiden Kompositionsreihen (25) und (26) von P haben also dieselbe Indexreihe.

Nehmen wir jetzt an, der zu beweisende Satz sei bereits als richtig erwiesen für Gruppen, deren Grad bei der Zerlegung in Primfaktoren einen Primfaktor weniger enthält als n (wenn n den Grad von P bedeutet), so folgt, daß alle Indexreihen von Q, dessen Grad ja ein echter Teiler von n und also weniger Primfaktoren als n enthält, miteinander übereinstimmen, daß also die Reihen

$$\nu', \mu_1, \mu_2, \mu_3, \ldots$$
$$\nu_1, \nu_2, \nu_3, \nu_4, \ldots,$$

von der Anordnung abgesehen, übereinstimmen. Ebenso stimmen die Indexreihen von Q'

$$\nu, \mu_1, \mu_2, \mu_3 \ldots$$
$$\nu_1', \nu_2', \nu_3', \nu_4' \ldots$$

überein. Also stimmen auch die beiden Indexreihen von P

$$\nu, \nu_1, \nu_2, \nu_3, \nu_4 \ldots$$
$$\nu', \nu_1', \nu_2, \nu_3', \nu_4' \ldots$$

miteinander überein, da die erste mit $\nu, \nu', \mu_1, \mu_2, \mu_3 \ldots$, die zweite mit $\nu', \nu, \mu_1, \mu_2, \mu_3 \ldots$ übereinstimmt.

Für Gruppen, deren Grad eine Primzahl ist, die nur den einen Normalteiler 1 haben, ist aber der Satz evident, und also ist er durch vollständige Induktion allgemein bewiesen.

Wenn in zwei Kompositionsreihen von P, die wir mit ihren Indexreihen jetzt so darstellen wollen:

$$(27) \qquad \begin{matrix} P, & P_1, & P_2, & P_3, & \ldots & P_{\mu-1}, & 1 \\ & j_1, & j_2, & j_3, & \cdots & j_{\mu-1}, & j_\mu \end{matrix}$$

$$(28) \qquad \begin{matrix} P, & P_1', & P_2', & P_3', & \ldots & P_{\mu-1}', & 1 \\ & j_1', & j_2', & j_3', & \cdots & j_{\mu-1}', & j_\mu', \end{matrix}$$

ein gemeinschaftliches Glied $P_r = P_s'$ vorkommt, so gilt der Satz von der Konstanz der Indexreihen auch für die Gruppe $P_r = P_s'$, und daraus folgt zunächst, daß $r = s$ sein muß, und daß die Indexreihen $j_{r+1}, j_{r+2} \ldots j_\mu$ und $j_{r+1}', j_{r+2}' \ldots j_\mu'$, von der Ordnung abgesehen, übereinstimmen. Folglich müssen auch die vorangehenden Teile der Indexreihen

$$j_1, j_2, \ldots j_r \text{ und } j_1', j_2', \ldots j_r'$$

übereinstimmen.

§ 48.

Permutationsgruppen.

Von besonderer Bedeutung für die Algebra sind die Permutationsgruppen. Aus diesen hat sich der Begriff der Gruppen

zuerst entwickelt, und auf sie lassen sich alle endlichen Gruppen zurückführen.

Es sei

$$P = a_1, a_2, \ldots a_n \tag{1}$$

eine Gruppe vom Grade n mit beliebigen Elementen und einem Kompositionsgesetz. Greifen wir aus P irgend ein Element b heraus, so ist der Komplex Pb mit P völlig identisch. Es können sich also die beiden Reihen

$$\begin{aligned} A &= a_1, \quad a_2, \quad a_3, \ \ldots, \ a_n \\ Ab &= a_1 b, \ a_2 b, \ a_3 b, \ \ldots, \ a_n b \end{aligned} \tag{2}$$

nur durch die Anordnung voneinander unterscheiden und sie müssen sich auch so unterscheiden, wenn b nicht das Einheitselement ist.

Der Übergang von der Anordnung A zu der Anordnung Ab heißt eine Permutation der n Ziffern

$$1, 2, 3, \ldots n, \tag{3}$$

und jedem Element b von P entspricht eine solche Permutation, die wir mit π_b bezeichnen wollen. Zwei verschiedenen Elementen b, c entsprechen zwei verschiedene Permutationen π_b, π_c und dem Einheitselement in P entspricht die identische Permutation π_1, die nichts in der Anordnung A ändert.

Wenn wir unter den π die Kompositionsregel festsetzen, die in der Formel

$$\pi_b \pi_c = \pi_{bc} \tag{4}$$

ausgedrückt ist, so sind die Permutationen π zu einer mit P isomorphen Gruppe Π vereinigt.

Nach (4) ist

$$\pi_b \pi_b^{-1} = \pi_1 \tag{5}$$

und folglich ist

$$\pi_b^{-1} = (\pi_b)^{-1}. \tag{6}$$

Man bezeichnet die Permutation π_b zweckmäßig nach (2) durch

$$\pi_b = (A, Ab). \tag{7}$$

Es ist aber auch, wenn c irgend ein Element in P ist:

$$\pi_b = (Ac, Acb), \tag{8}$$

denn wenn man in (2) die n übereinanderstehenden Paare irgendwie permutiert, also A durch Ac ersetzt, so ist der Übergang von A zu Ab nichts anderes als der Übergang von Ac zu Acb.

Demnach drückt sich die Kompositionsregel (4) auch in der Formel aus:

$$(A,\ Ab)\ (Ab,\ Ac) = (A,\ Ac). \tag{9}$$

Die zu π_b entgegengesetzte Permutation $\pi_b{}^{-1}$ hat dann die Bezeichnung

$$\pi_b{}^{-1} = (Ab,\ A) \tag{10}$$

und $(A, A) = \pi_1$ ist die identische Permutation.

Ausführlicher bezeichnet man die Permutation π_b auch so:

$$\begin{pmatrix} a_1, & a_2, & a_3, & \dots a_n \\ a_1 b, & a_2 b, & a_3 b, & \dots a_n b \end{pmatrix}. \tag{11}$$

Da a, b unter den Elementen a enthalten ist, so kann man es mit a_{b_1} bezeichnen und, indem man in der Bezeichnung (11) nur die Indices setzt, dafür schreiben:

$$\begin{pmatrix} 1, & 2, & 3, & \dots n \\ b_1, & b_2, & b_3, & \dots b_n \end{pmatrix}, \tag{12}$$

was gleichbedeutend ist mit

$$\begin{pmatrix} a_1, & a_2, & a_3, & \dots a_n \\ a_{b_1}, & a_{b_2}, & a_{b_3}, & \dots a_{b_n} \end{pmatrix}. \tag{13}$$

Hier läßt sich die Zusammensetzung nach der Formel (9) leicht ausführen.

So ist z. B.:

$$\begin{pmatrix} 1, & 2, & 3 \\ 2, & 3, & 1 \end{pmatrix} \begin{pmatrix} 1, & 2, & 3 \\ 3, & 1, & 2 \end{pmatrix} = \begin{pmatrix} 1, & 2, & 3 \\ 2, & 3, & 1 \end{pmatrix} \begin{pmatrix} 2, & 3, & 1 \\ 1, & 2, & 3 \end{pmatrix}$$
$$= \begin{pmatrix} 1, & 2, & 3 \\ 1, & 2, & 3 \end{pmatrix} = \pi_1.$$

Wenn durch π_b einige Ziffern ungeändert bleiben, so werden diese in der Bezeichnung (12) bisweilen unterdrückt. Es ist also z. B.

$$\begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 2, & 3, & 1, & 4, & 5 \end{pmatrix} = \begin{pmatrix} 1, & 2, & 3 \\ 2, & 3, & 1 \end{pmatrix}.$$

Die Gesamtzahl aller möglichen Permutationen von n Ziffern beträgt bekanntlich:

$$1.2.3 \dots n = n!$$

und diese bilden, nach unserer Kompositionsregel behandelt, eine Gruppe des Grades $n!$. Wir nennen sie, vorläufig nur um einen Namen dafür zu haben, die **symmetrische Gruppe der n Elemente**. Der Name bezieht sich auf ihre Anwendung in der Algebra, die wir später kennen lernen werden.

Jede andere Permutationsgruppe der n Ziffern ist ein Teiler dieser Gruppe und folglich ihr Grad ein Teiler von $n!$

Es ist eine fundamentale Aufgabe der Algebra, die freilich noch lange nicht gelöst ist, alle möglichen Gruppen der Permutationen, d. h. alle Teiler der symmetrischen Gruppe zu finden.

Nach diesen Ausführungen ist die Theorie der endlichen Gruppen wesentlich identisch mit der Theorie der Permutationsgruppen.

§ 49.

Zerlegung von Permutationen in Transpositionen und in Zyklen.

Die Vertauschung zweier Ziffern in einer Anordnung heißt eine Transposition. Sie wäre nach § 48 zu bezeichnen durch

$$\begin{pmatrix} 1, & 2 \\ 2, & 1 \end{pmatrix},$$

wofür man einfacher (1, 2) schreibt. Wiederholt man dieselbe Transformation, so gelangt man zur Identität. Daher ist eine Transposition sich selbst entgegengesetzt.

Ist nun

$$\pi = \begin{pmatrix} 1, & 2, & \ldots, & n \\ a_1, & a_2, & \ldots, & a_n \end{pmatrix}$$

eine beliebige Permutation von n Ziffern, so ist, wenn man π mit (n, a_n) komponiert,

$$\pi\,(n, a_n) = \pi'$$

eine Permutation, die n ungeändert läßt, die also durch

$$\begin{pmatrix} 1, & 2, & \ldots & n-1 \\ a_1, & a_2, & \ldots & a_{n-1} \end{pmatrix}$$

dargestellt werden kann und also eine Permutation von höchstens $n - 1$ Ziffern ist. Da nun auch

$$\pi = \pi'\,(n, a_n)$$

ist, so folgt hieraus durch vollständige Induktion:

1. Man kann jede Permutation (und zwar auf unendlich viele verschiedene Arten) in Transpositionen zerlegen.

Und daraus folgt weiter:

2. Wenn eine Permutationsgruppe von n Ziffern alle Transpositionen mit einer festen Ziffer, z. B.:

$$(1, 2), (1, 3), \ldots (1, n)$$

enthält, so ist sie mit der symmetrischen Gruppe identisch.

Denn eine solche Gruppe enthält, wie man aus der Zusammensetzung

$$(2, 3) = (1, 2)\,(1, 3)\,(1, 2)$$

erkennt, auch alle anderen Transpositionen, und nach 1. lassen sich daraus alle Permutationen der symmetrischen Gruppe zusammensetzen.

Eine Permutation π heißt zyklisch, wenn sich die Ziffern so in eine Reihe ordnen lassen, daß durch π jede Ziffer in die folgende und die letzte wieder in die erste übergeht, also z. B.

$$\begin{pmatrix} 1, 2, \ldots n-1, n \\ 2, 3, \ldots n, \quad 1 \end{pmatrix};$$

solche zyklische Permutationen bezeichnet man einfacher, indem man die Ziffern des Zyklus nebeneinander setzt, durch

$$(1, 2, 3, \ldots n).$$

Dabei ist es gleichgültig, mit welcher Ziffer man anfängt; man könnte also auch

$$(2, 3, \ldots n, 1)$$

dafür setzen. Die Transpositionen sind nichts weiter als zweigliedrige Zyklen.

Es gilt nun der folgende Satz:

3. Jede Permutation π läßt sich, und zwar nur auf eine Weise, in eine Reihe von zyklischen Permutationen zerlegen, so daß keine zwei dieser Zyklen eine Ziffer gemeinschaftlich haben.

Ist nämlich

$$\pi = \begin{pmatrix} 1, & 2, & \ldots n \\ a_1, & a_2, & \ldots a_n \end{pmatrix}, \tag{1}$$

so fange man mit einer beliebigen Ziffer, etwa mit 1 an, und setze die Reihe

$$1, \quad a_1 = b, \quad a_b = c \ldots \tag{2}$$

so lange fort, bis man auf eine Ziffer zum zweiten Male stößt. Die zuerst wiederkehrende Ziffer muß 1 sein, da zu jeder Ziffer

die in der Reihe (2) vorangehende durch (1) eindeutig bestimmt ist. Dann bilden die Ziffern

(3) $$(1, b, c \ldots)$$

einen ersten Zyklus. Sind dadurch noch nicht alle n Ziffern von (1) erschöpft, so greift man aus den übrigen eine heraus und verfährt ebenso, bis alle n Ziffern von (1) in den Zyklen untergebracht sind. Da in jedem solchen Zyklus zu jeder Ziffer die vorangehende sowohl als die nachfolgende durch (1) völlig bestimmt ist, so sind auch die Zyklen selbst eindeutig bestimmt. Bei der Bezeichnung von π durch die Zyklen können aber nicht nur die verschiedenen Zyklen beliebig angeordnet werden, sondern man kann auch in jedem Zyklus mit einer beliebigen seiner Ziffern anfangen. Eine Ziffer, die nicht geändert wird, bildet für sich einen eingliedrigen Zyklus.

Wir wählen ein ganz beliebiges Beispiel, wodurch das Verfahren sofort klar wird:

$$\pi = \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 3, & 6, & 7, & 1, & 4, & 5, & 8, & 2 \end{pmatrix} = (1, 3, 7, 8, 2, 6, 5, 4)$$

$$\pi' = \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 2, & 1, & 3, & 5, & 7, & 8, & 6, & 4 \end{pmatrix} = (8, 4, 5, 7, 6)\ (1, 2)\ (3)$$

$$\pi'' = \begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8 \\ 2, & 3 & 6, & 5, & 8, & 1, & 4, & 7 \end{pmatrix} = (8, 7, 4, 5)\ (1, 2, 3, 6).$$

Das Nebeneinandersetzen der Zyklen ist mit einer Komposition in dem bisherigen Sinne gleichbedeutend, nur ist zu bemerken, daß Permutationen, die gar keine gemeinschaftliche Ziffer enthalten, bei der Komposition immer vertauscht werden können.

Eingliedrige Zyklen, die nichts ändern, werden in der Bezeichnung auch oft weggelassen.

Da im allgemeinen, wenn π und $\varkappa$ irgend zwei Permutationen von n Ziffern sind, $\varkappa\pi$ von $\pi\varkappa$ verschieden ist, so ist auch

$$\pi^{-1}\varkappa\pi = \varkappa'$$

von $\varkappa$ verschieden. Ist $\varkappa$ in seine Zyklen zerlegt, so kann man $\varkappa'$ nach dem folgenden Satze sehr einfach aus $\varkappa$ ableiten:

4. Man erhält die Permutation $\pi^{-1}\varkappa\pi$ dadurch, daß man in den Zyklen von $\varkappa$ die Permutation π ausführt.

Um diese Regel zu beweisen, sei, in Zyklen zerlegt,

$$\varkappa = (\alpha, \beta, \gamma \ldots)\ (\alpha', \beta', \gamma' \ldots) \ldots$$

und es sei

$$\pi = \begin{pmatrix} \alpha, & \beta, & \gamma, & \ldots & \alpha', & \beta', & \gamma' & \ldots \\ \alpha_\pi, & \beta_\pi, & \gamma_\pi, & \ldots & \alpha'_\pi, & \beta'_\pi, & \gamma'_\pi & \ldots \end{pmatrix}.$$

Durch π^{-1} geht α_π in α über, durch $\varkappa$ geht α in β über und durch π wird β in β_π übergeführt. Durch $\pi^{-1}\varkappa\pi$ geht also α_π in β_π über. Da dieselbe Betrachtung auf β_π, γ_π ... usf. anwendbar ist, so folgt:

$$\pi^{-1}\varkappa\pi = (\alpha_\pi, \beta_\pi, \gamma_\pi \ldots)\ (\alpha'_\pi, \beta'_\pi, \gamma'_\pi \ldots) \ldots,$$

und dies ist der Inhalt des Satzes 4.

Man könnte ihn auch aus der folgenden Darstellung ablesen. Ist

$$\pi = \begin{pmatrix} 1, & 2 & \ldots \\ \alpha_1, & \alpha_2 & \ldots \end{pmatrix} = \begin{pmatrix} \beta_1, & \beta_2 & \ldots \\ \alpha_{\beta_1}, & \alpha_{\beta_2} & \ldots \end{pmatrix}$$

$$\varkappa = \begin{pmatrix} 1, & 2 & \ldots \\ \beta_1, & \beta_2 & \ldots \end{pmatrix},$$

dann ist

$$(4)\quad \pi^{-1}\varkappa\pi = \begin{pmatrix} \alpha_1, & \alpha_2 & \ldots \\ 1, & 2 & \ldots \end{pmatrix} \begin{pmatrix} 1, & 2 & \ldots \\ \beta_1, & \beta_2 & \ldots \end{pmatrix} \begin{pmatrix} \beta_1, & \beta_2 & \ldots \\ \alpha_{\beta_1}, & \alpha_{\beta_2} & \ldots \end{pmatrix} = \begin{pmatrix} \alpha_1, & \alpha_2 & \ldots \\ \alpha_{\beta_1}, & \alpha_{\beta_2} & \ldots \end{pmatrix}.$$

Hieraus können wir einen ersten Teiler der symmetrischen Gruppe vom Index 2 ableiten.

Bedeutet τ eine Transposition und π eine beliebige Permutation, so ist in der zusammengesetzten Permutation $\tau\pi$ die Anzahl der Zyklen um eins größer oder um eins kleiner als in π. Die beiden Ziffern, die durch τ miteinander vertauscht werden, können entweder in demselben Zyklus von π vorkommen oder in zwei verschiedenen Zyklen. Nehmen wir an, es sei ein in π vorkommender Zyklus $\gamma = (1, 2, \ldots a, a + 1, \ldots b)$ und es enthalte τ zwei Ziffern, die in γ vorkommen, etwa $\tau = (1, a)$, dann ist

$$\tau\gamma = (1, a + 1, \ldots b)\ (2, 3, \ldots a),$$

d. h. γ wird durch Zusammensetzung mit τ in zwei Zyklen zerlegt, während die übrigen Zyklen durch τ nicht berührt werden.

Wenn aber die beiden Ziffern von τ in zwei verschiedenen Zyklen von π vorkommen, so mögen diese beiden Zyklen

$$\gamma = (1, 2, 3, \ldots a), \quad \gamma' = (1', 2', 3', \ldots a')$$

sein und $\tau = (1, 1')$. Es ist dann

$$\tau\gamma\gamma' = (1, 2', 3', \ldots a', 1', 2, 3, \ldots a),$$

d. h. die beiden Zyklen γ, γ' werden durch τ zu einem einzigen Zyklus vereinigt, während wieder die übrigen Zyklen ungeändert bleiben.

Wenn wir also eine Permutation π von n Ziffern, die aus ν Zyklen besteht (wobei die eingliedrigen Zyklen mitgezählt werden), durch μ Transpositionen dargestellt haben, so kann sie durch Zusammensetzung mit diesen μ Transpositionen in umgekehrter Reihenfolge in die identische Permutation verwandelt werden, die aus n eingliedrigen Zyklen besteht. Es sind daher im ganzen $n - \nu$ Zyklen gewonnen, und da jeder Transposition ein gewonnener oder ein verlorener Zyklus entspricht, so muß $\mu \equiv n - \nu \pmod 2$ sein, d. h. μ ist eine gerade oder eine ungerade Zahl, je nachdem $n - \nu$ gerade oder ungerade ist. Die letzte Zahl ist aber nur von der Permutation π, nicht von der Art der Zerlegung in Transpositionen abhängig. Wir haben damit den Satz:

5. Die Permutationen von n Ziffern zerfallen in zwei Arten, von denen die erste in eine gerade, die zweite in eine ungerade Anzahl von Transpositionen zerlegbar ist.

Jede dieser beiden Arten umfaßt gleich viele Permutationen, nämlich $\frac{1}{2}n!$; denn durch Hinzufügung einer festen Transposition geht jede Permutation der ersten Art in eine der zweiten Art über und umgekehrt.

Die Zusammensetzung zweier Permutationen von gleicher Art gibt stets eine Permutation der ersten Art, während eine Permutation erster und zweiter Art, zusammengesetzt, eine Permutation zweiter Art ergeben, wie aus der Zerlegung in Transpositionen sofort zu ersehen ist. Daraus folgt:

6. Die Permutationen der ersten Art bilden eine Gruppe, also einen Teiler der symmetrischen Gruppe vom Index 2; sie wird die alternierende Gruppe genannt.

Die Permutationen der zweiten Art bilden die zugehörige Nebengruppe.

Diese Gruppenzerlegung ist die Grundlage der Determinantenbildung.

Eine zyklische Permutation von n Gliedern läßt sich in $n - 1$ Transpositionen zerlegen, wie man aus der Zusammensetzung

$$(1, 2, 3 \ldots n) = (1, 2)\ (1, 3) \ldots (1, n)$$

erkennt, und daraus folgt, daß eine zyklische Permutation zur ersten oder zur zweiten Art gehört, je nachdem die Anzahl der Ziffern eine ungerade oder eine gerade ist.

Wie man jede Permutation aus Transpositionen zusammensetzen kann, ebenso kann man jede Permutation der ersten Art aus dreigliedrigen Zyklen zusammensetzen. Es genügt, wenn dies für jedes Paar von Transpositionen bewiesen ist, da jede Permutation der ersten Art sich aus solchen Paaren komponieren läßt. Es ist aber

$$(1, 2)\ (1, 3) = (1, 2, 3)$$
$$(1, 4)\ (2, 3) = (1, 2, 3)\ (1, 2, 4),$$

woraus die Richtigkeit der Behauptung zu ersehen ist, da die beiden Transpositionen eines Paares entweder eine oder keine Ziffer gemein haben. Also:

7. Alle Permutationen der ersten Art lassen sich (auf unendlich viele Weisen) in zyklische Permutationen von drei Elementen zerlegen.

Aus 7. folgt weiter:

8. Eine Permutationsgruppe von n Ziffern, die alle dreigliedrigen zyklischen Permutationen mit zwei festen Ziffern enthält, muß die ganze alternierende Gruppe enthalten.

Denn aus den Zyklen $(1, 2, 3)$, $(1, 2, 4) \ldots (1, 2, n)$ lassen sich alle dreigliedrigen Zyklen und also nach 7. alle Permutationen der ersten Art komponieren, wie man aus den Zusammensetzungen

$$(2, 1, 3) = (1, 2, 3)\ (1, 2, 3),$$
$$(1, 3, 4) = (1, 2, 3)\ (2, 1, 4)\ (2, 1, 3),$$
$$(2, 4, 5) = (2, 1, 4)\ (1, 2, 5)\ (1, 2, 4),$$
$$(3, 4, 5) = (2, 1, 3)\ (2, 4, 5)\ (1, 2, 3),$$

ersieht.

9. Ist n größer als 4, so ist eine Permutationsgruppe von n Ziffern, die alle aus je zwei Transpositionen ohne gemeinsame Ziffer zusammengesetzten Permutationen $(1, 2)\ (3, 4)$ enthält, durch die alternierende Gruppe teilbar.

Denn es ist

$$(1, 2, 3) = (1, 2)\ (4, 5)\ (4, 5)\ (1, 3)$$

und man kann also, wenn außer den vier Ziffern 1, 2, 3, 4 noch eine fünfte, 5, vorhanden ist, alle dreigliedrigen Zyklen aus Transpositionspaaren von der Form (1, 2) (4, 5) zusammensetzen.

Ist dagegen $n = 4$, so bilden die vier Permutationen

$$1; \quad (1, 2)\,(3, 4); \quad (1, 3)\,(2, 4); \quad (1, 4)\,(2, 3)$$

eine Gruppe, die kleiner ist als die alternierende Gruppe, weil in dieser außerdem noch die acht dreigliedrigen Zyklen vorkommen.

§ 50.

Transitive und primitive Permutationsgruppen.

1 Zwei Elemente a_1, a_2 heißen durch die Permutationsgruppe Π verbunden (transitiv verbunden), wenn es in Π eine Permutation π gibt, durch die a_1 in a_2 übergeht. Wenn zwei Elemente mit einem dritten verbunden sind, so sind sie auch untereinander verbunden.

Ist a_1 durch π mit a_2 verbunden, so ist a_2 durch π^{-1} mit a_1 verbunden, und wenn a_1 durch π_2 mit a_2, durch π_3 mit a_3 verbunden ist, so ist a_2 mit a_3 durch $\pi_2^{-1}\,\pi_3$ verbunden.

Die Gruppe der Permutationen der Elemente

$$P = a_1, a_2, a_3, \ldots a_n \tag{1}$$

heißt transitiv, wenn alle ihre Elemente miteinander verbunden sind, im entgegengesetzten Falle intransitiv.

Durch eine intransitive Gruppe wird das System P in mehrere Teilsysteme A, B, C ... zerlegt, so daß alle Elemente eines dieser Systeme untereinander verbunden sind, aber kein Element des einen Systems mit einem des anderen. Diese Systeme heißen die Systeme der Intransitivität und die Elemente eines jeden Systems transitiv verbunden.

Eine Gruppe ist also transitiv, wenn sie Permutationen enthält, durch die ein beliebiges Element a_1 in ein beliebiges anderes Element b_1 übergeht.

Gibt es Permutationen, durch die zwei beliebige Elemente a_1, a_2 in zwei beliebige andere Elemente b_1, b_2 übergehen, so heißt die Gruppe zweifach transitiv und ebenso werden mehrfach transitive Gruppen definiert. Die symmetrische Gruppe ist hiernach $n$fach transitiv.

2. Eine transitive Gruppe Π, die einen intransitiven Normalteiler N hat, heißt imprimitiv. Die Systeme der Intransitivität von N heißen die Systeme der Imprimitivität von Π.

Sind etwa

$$A = a_1, a_2, \dots a_r \tag{2}$$

die durch N mit a_1 verbundenen Elemente und π eine Permutation, durch die a_1 in ein nicht in A enthaltenes Element b_1 übergeht, so geht A durch π in ein System von Elementen

$$B = b_1, b_2, \dots b_r \tag{3}$$

über, die alle durch $\pi^{-1} N \pi = N$ mit b_1 verbunden sind, und folglich ist B entweder mit A identisch oder ganz davon verschieden und umfaßt alle mit b_1 verbundenen Elemente. Die Systeme der Intransitivität von N sind alle von gleichem Grad r, und r ist ein Teiler von n.

Die identische Permutation bildet für sich eine Gruppe, und ist also ein intransitiver Normalteiler einer jeden anderen Gruppe. Hiernach könnte man jede Gruppe Π imprimitiv nennen, wenn wir nicht noch die Forderung zufügen würden:

$$r > 1. \tag{4}$$

Sind jetzt

$$\begin{aligned} A &= a_1, a_2, \dots a_r \\ B &= b_1, b_2, \dots b_r \\ &\dots\dots\dots\dots \\ S &= s_1, s_2, \dots s_r \end{aligned} \tag{5}$$

die Systeme der Intransitivität von N, so werden durch N die A nur unter sich, die B nur unter sich, die S nur unter sich vertauscht.

Weil Π transitiv ist, so gibt es in Π eine Permutation β, durch die a_1 in b_1 übergeführt wird, und weil durch N die B verbunden sind, so wird a_1 durch βN in jedes b übergeführt und durch $N \beta N$ jedes a in jedes b, also A in B, und es ist

$$\Pi = N + \beta N + \gamma N + \cdots \sigma N. \tag{6}$$

Wir können daher auch so definieren:

3. Eine transitive Gruppe Π heißt imprimitiv, wenn die Elemente von P derart in gleich starke Systeme $A, B, \dots S$ von mehr als einem Element zerlegt werden können, daß alle Permutationen

von Π dieses System nicht auseinanderreißen, sondern nur die Systeme untereinander und die Elemente der Systeme untereinander permutieren.

Die Permutationen von Π, die die Systeme nicht, sondern nur deren Elemente permutieren, bilden die Gruppe N.

Das folgende Beispiel zeigt, daß eine Gruppe auf mehrfache Weise imprimitiv sein kann.

Es sei $n = 6$ und Π die zyklische Gruppe, die aus folgenden in ihre Zyklen zerlegten Permutationen besteht:

$$\pi^0 = (1)$$
$$\pi^1 = (1, 2, 3, 4, 5, 6)$$
$$\pi^2 = (1, 3, 5)\,(2, 4, 6)$$
$$\pi^3 = (1, 4)\,(2, 5)\,(3, 6)$$
$$\pi^4 = (1, 5, 3)\,(2, 6, 4)$$
$$\pi^5 = (1, 6, 5, 4, 3, 2).$$

Hier haben wir die Systeme der Imprimitivität:

$$A = 1, 3, 5$$
$$B = 2, 4, 6$$

mit der Gruppe

$$N = 1, \pi^2, \pi^4,$$

und

$$A = 1, 4$$
$$B = 2, 5$$
$$C = 3, 6$$

mit der Gruppe

$$N = 1, \pi^3.$$

Wir beweisen noch die beiden folgenden Sätze:

4. Wenn eine transitive Permutationsgruppe von n Ziffern eine einzelne Transposition enthält, so ist sie entweder die symmetrische Gruppe, oder sie ist imprimitiv.

Die fragliche Gruppe Π enthalte die Transposition (1, 2) und außerdem

$$(1, 3),\ (1, 4),\ \ldots\ (1, \mu), \tag{7}$$

aber keine andere Transposition mit der Ziffer 1. Ist $\mu = n$, so ist Π nach § 49, 2. die symmetrische Gruppe. Ist $\mu < n$, so bilden die durch Zusammensetzung der Transpositionen (7) ab-

geleiteten Permutationen der Ziffern 1, 2, 3, ... μ einen intransitiven Normalteiler von H und Π ist also nach 2. imprimitiv, wie bewiesen werden sollte.

Ebenso wird auf Grund von § 49, 8. bewiesen:

5. Wenn eine transitive Permutationsgruppe einen dreigliedrigen Zyklus enthält, so enthält sie die ganze alternierende Gruppe oder ist imprimitiv.

§ 51.

Einfachheit der alternierenden Gruppe.

Wir gehen nun zur Betrachtung einzelner Gruppen über und beginnen mit der alternierenden Gruppe A. Diese ist ein Normalteiler der symmetrischen Gruppe S. Denn sind π, $\varkappa$ irgend zwei Permutationen aus S, so gehört $\pi^{-1}\varkappa\pi$ zu derselben Art wie $\varkappa$, also wenn $\varkappa$ zu A gehört, ebenfalls zu A.

Es gilt aber der wichtige Satz, den wir nun zu beweisen haben:

> Wenn $n > 4$ ist, so hat die alternierende Gruppe A außer sich selbst und der Einheit keinen Normalteiler, ist also nach der Bezeichnung in § 47 einfach.

Der Beweis wird so geführt:

Sei A die alternierende Gruppe der Permutationen von n Ziffern 1, 2, 3, ... n und Q ein normaler Teiler von A. Wir haben im § 49, 7. und 4. gesehen, daß sich alle Permutationen von A aus dreigliedrigen Zyklen zusammensetzen lassen, und daß man, wenn $\varkappa$ und π irgend welche Permutationen sind, die transformierte Permutation zu $\varkappa$, $\pi^{-1}\varkappa\pi$ erhält, wenn man in den Zyklen von $\varkappa$ die Vertauschungen π vornimmt. Ist nun $\varkappa$ ein dreigliedriger Zyklus, etwa (1, 2, 3), so kann man π aus A so wählen, daß $\pi^{-1}\varkappa\pi$ jeden beliebigen dreigliedrigen Zyklus der n Ziffern darstellt; denn man kann in

$$\pi = \begin{pmatrix} 1, & 2, & 3, .. & n \\ a_1, & a_2, & a_3, \ldots & a_n \end{pmatrix}$$

die drei ersten Ziffern a_1, a_2, a_3 beliebig wählen, und, wenn es nötig ist, damit π zu A gehöre, noch a_1 und a_2 vertauschen. Dadurch geht aus $\varkappa$ einer der beiden Zyklen (a_1, a_2, a_3), (a_2, a_1, a_3) hervor, von denen jeder die zweite Potenz des anderen ist. Wenn

nun Q ein Normalteiler von A ist, und $\varkappa$ eine Permutation aus Q, so ist $\pi^{-1}\varkappa\pi$ auch in Q enthalten, wenn π in A enthalten ist, und daraus folgt, daß, wenn in Q ein dreigliedriger Zyklus vorkommt, Q mit A identisch ist.

Unser Beweis beruht nun darauf, daß, wenn $\varkappa$ irgend eine Permutation in Q ist, auch $\pi^{-1}\varkappa\pi$ und folglich auch

$$\lambda = \varkappa^{-1}\pi^{-1}\varkappa\pi \tag{1}$$

in Q vorkommen muß, und es ist dann zu zeigen, daß man, wenn $\varkappa$ irgend eine nicht identische Permutation ist, π immer so aus A wählen kann, daß die Permutation λ ein dreigliedriger Zyklus und folglich Q mit A identisch wird.

Wir nehmen zu diesem Zwecke sowohl $\varkappa$ als π in ihre Zyklen zerlegt an und bemerken, daß man bei der Bildung von λ solche Zyklen von $\varkappa$ gar nicht zu berücksichtigen braucht, deren Ziffern durch π ungeändert bleiben, weil sie sich in $\varkappa^{-1}$ und $\varkappa$ gegenseitig aufheben. Wir müssen nun die verschiedenen möglichen Formen von $\varkappa$ einzeln betrachten.

1. Es enthalte $\varkappa$ einen Zyklus von mehr als drei Ziffern, etwa $(1, 2, 3, \ldots m)$, wir nehmen $\pi = (1, 2, 3)$, $\pi^{-1} = (1, 3, 2)$ an und erhalten, indem wir $\varkappa^{-1}\pi^{-1}\varkappa$ zuerst (nach § 49, 4.) bilden,
$$\lambda = \varkappa^{-1}\pi^{-1}\varkappa\pi = (2, 4, 3)\,(1, 2, 3) = (1, 2, 4).$$
In Q kommt also ein dreigliedriger Zyklus vor.
2. Es enthalte $\varkappa$ zwei dreigliedrige Zyklen $(1, 2, 3)$ $(4, 5, 6)$. Wir nehmen $\pi = (1, 3, 4)$ an und erhalten
$$\lambda = (2, 5, 1)\,(1, 3, 4) = (1, 2, 5, 3, 4).$$
Diese Permutation λ, die in Q enthalten ist, fällt aber unter den Fall 1.
3. Es enthalte $\varkappa$ einen dreigliedrigen und einen zweigliedrigen Zyklus $(1, 2, 3)$ $(4, 5)$. (Daß in $\varkappa$, wenn es zu A gehört, noch ein zweiter Zyklus von gerader Gliederzahl vorkommen muß, ist hier gleichgültig.) Für $\pi = (1, 2, 4)$ ergibt sich
$$\lambda = (2, 5, 3)\,(1, 2, 4) = (1, 2, 5, 3, 4),$$
was wieder unter den Fall 1 fällt.
4. Es enthalte $\varkappa$ drei Transpositionen $(1, 2)$ $(3, 4)$ $(5, 6)$. Für $\pi = (1, 3, 5)$ folgt
$$\lambda = (2, 6, 4)\,(1, 3, 5),$$
was unter den Fall 2 fällt.

5. Es enthalte $\varkappa$ zwei Transpositionen und ein unverändertes Element (1, 2) (3, 4) (5). Man setzt $\pi = (1, 2, 5)$ und erhält

$$\lambda = (1, 2, 5)\,(1, 2, 5) = (1, 5, 2).$$

Damit sind alle Fälle erschöpft, wenn $n > 4$ ist. Für $n = 4$ bleibt der eine Fall noch übrig, daß $\varkappa$ aus zwei Transpositionen besteht.

6. In dem Falle $n = 4$ bilden in der Tat die vier Permutationen

$$1;\ (1, 2)\,(3, 4);\ (1, 3)\,(2, 4);\ (1, 4)\,(2, 3)$$

einen Normalteiler von A; denn ist $\varkappa$ eines dieser drei Paare von Transpositionen, so ist $\pi^{-1}\varkappa\pi$ immer wieder eines dieser Paare.

Es folgt aus diesem Satze weiter, daß die symmetrische Gruppe außer im Falle 6. keine anderen normalen Teiler hat als sich selbst, die alternierende Gruppe und die Einheitsgruppe. Denn ist S die symmetrische, A die alternierende Gruppe, und Q ein normaler Teiler von S, so ist der größte gemeinschaftliche Teiler Q' von A und Q ein normaler Teiler von A, ist also gleich A oder gleich 1. Denn ist $\varkappa'$ ein Element in Q', also auch in A, π ein beliebiges Element in S, so ist $\pi^{-1}\varkappa'\pi$ zunächst in Q enthalten, weil Q Normalteiler von S ist. Es ist aber zugleich $\pi^{-1}\varkappa'\pi$ in A, also auch in Q' enthalten.

Ist Q' gleich A, so ist Q entweder auch gleich A oder gleich S. Ist Q' aber $= 1$, so enthält Q außer der Einheit keine Permutation der ersten Art. Sind also $\varkappa$ und λ zwei verschiedene und von der Einheit verschiedene Permutationen von Q, so müssen $\varkappa^2$ und $\varkappa\lambda$ als von der ersten Art $= 1$ sein, d. h. λ muß $= \varkappa$ sein. Es kann also Q höchstens eine von der identischen verschiedene Permutation $\varkappa$ enthalten. Da aber Q ein Normalteiler von S sein soll, so muß für jede Permutation π aus der symmetrischen Gruppe $\pi^{-1}\varkappa\pi = \varkappa$ sein, d. h. $\varkappa$ darf sich nicht ändern, wenn in seinen Zyklen irgend eine Permutation ausgeführt wird. Dies ist aber nur dann möglich, wenn überhaupt nur zwei Ziffern 1, 2 vorhanden sind und $\varkappa = (1, 2)$ ist. Dann aber ist 1, (1, 2) die ganze Gruppe S.

§ 52.

Abelsche Gruppen.

Man nennt solche Gruppen, in denen bei der Komposition das kommutative Gesetz gilt, kommutative oder Abelsche Gruppen. Bei diesen Gruppen herrschen viel einfachere Gesetze, als in den allgemeinen Gruppen, wie wir jetzt sehen werden. Es gelten für die Zusammensetzung der Elemente in diesen Gruppen dieselben Regeln, wie bei der Multiplikation von Zahlen, und wir nennen demnach die Komposita von Elementen einer solchen Gruppe, wie bei der Multiplikation, Produkte und Potenzen.

Wir betrachten hier nur die endlichen Abelschen Gruppen.

Aus der Definition der Abelschen Gruppen folgt, daß das kommutative Gesetz auch bei der Komposition der Teile einer solchen Gruppe gilt. Jeder Teiler einer Abelschen Gruppe ist selbst eine Abelsche Gruppe, und ist zugleich Normalteiler, so daß wir bei diesen Gruppen den Zusatz „Normal“ weglassen können.

Eine Gruppe, die nur aus den Wiederholungen eines Elementes besteht, wie

$$1,\ A,\ A^2 \ldots A^{a-1},$$

heißt eine zyklische Gruppe. Wenn a die kleinste positive Zahl ist, für die $A^a = 1$ ist, also a der Grad der Gruppe, so heißt a auch der Grad des Elementes A. Ist m irgend ein Exponent, für den $A^m = 1$ ist, so ist m ein Vielfaches von a. Denn ist m nicht durch a teilbar, so hat m bei der Division durch a einen Rest $a' < a$ und es ist $A^m = A^{a'} = 1$. Es muß also $a' = 0$, d. h. m durch a teilbar sein. Das Element 1 hat den Grad 1. Alle zyklischen Gruppen sind kommutativ.

Es gilt nun für jede endliche Abelsche Gruppe S der Fundamentalsatz, daß sich immer ein System von Elementen A, B, $C \ldots$ von den Graden a, b, $c \ldots$ so auswählen läßt, daß sich in der Form

$$\Theta = A^\alpha B^\beta C^\gamma \ldots$$

jedes Element Θ von S, und jedes nur einmal, darstellen läßt, wenn α ein volles Restsystem nach dem Modul a, β ein volles Restsystem nach dem Modul b, γ ein volles Restsystem nach dem Modul c usw. durchläuft.

Ein System von Elementen, das zu einer solchen Darstellung geeignet ist, heißt eine Basis der Gruppe, und wir können den zu beweisenden Satz demnach so aussprechen:

I. Jede Abelsche Gruppe läßt sich durch eine Basis darstellen.

Der Beweis des Satzes gründet sich auf folgende Reihe von Schlüssen:

1. Ist n der Grad der Gruppe S, sind $A_1, A_2 \dots A_n$ ihre Elemente und $a_1, a_2, \dots a_n$ deren Grade; durchlaufen ferner $\alpha_1, \alpha_2, \dots \alpha_n$ volle Restsysteme nach den Moduln $a_1, a_2, \dots a_n$, so wird in der Form

$$\Theta = A_1^{\alpha_1} A_2^{\alpha_2} \dots A_n^{\alpha_n} \tag{1}$$

jedes Element von S und jedes gleich oft dargestellt.

Daß man jedes Element in der Form (1) überhaupt erhält, sieht man unmittelbar; denn man erhält z. B. A_1, wenn man $\alpha_1 = 1$, $\alpha_2 = 0, \dots \alpha_n = 0$ setzt.

Es ist noch zu beweisen, daß man jedes Element gleich oft erhält.

Wenn aber

$$1 = A_1^{h_1} A_2^{h_2} \dots A_n^{h_n} \tag{2}$$

eine Darstellung des Elementes 1 ist, so ändert sich Θ nicht, wenn man $\alpha_1, \alpha_2, \dots \alpha_n$ in (1) durch $\alpha_1 + h_1$, $\alpha_2 + h_2, \dots \alpha_n + h_n$ ersetzt. Es wird also jedes Element Θ durch (1) mindestens so oft dargestellt, als das Element 1 durch (2).

Geben andererseits die beiden Exponentenreihen $\alpha_1, \alpha_2, \dots \alpha_n$ und $\alpha_1', \alpha_2' \dots \alpha_n'$ zwei Darstellungen des Elementes Θ, so hat man

$$1 = A_1^{\alpha_1' - \alpha_1} A_2^{\alpha_2' - \alpha_2} \dots A_n^{\alpha_n' - \alpha_n}, \tag{3}$$

also eine Darstellung des Elementes 1. Also können wir nach (2) setzen:

$$\alpha_1' = \alpha_1 + h_1, \quad \alpha_2' = \alpha_2 + h_2 \dots \alpha_n' = \alpha_n + h_n.$$

Es kann folglich auch nicht mehr verschiedene Darstellungen des Elementes Θ geben, als die Anzahl der Darstellungen des Elementes 1 beträgt. Wir bezeichnen die Anzahl dieser Darstellungen mit h. Im ganzen ist aber die Anzahl der verschiedenen möglichen Bestimmungen der α in (1) gleich dem

Produkte $a_1 a_2 \dots a_n$, und da n die Anzahl der Elemente von S ist, so folgt

$$nh = a_1 a_2 \dots a_n. \tag{4}$$

Aus der Formel (4) ergibt sich der Satz:

2. Wenn r eine im Grade n der Gruppe aufgehende Primzahl ist, so gibt es in S ein Element vom Grade r.

Denn aus (4) sieht man zunächst, daß einer der Grade $a_1, a_2 \dots a_n$ durch r teilbar ist. Ist also etwa $a_1 = rk$, so ist A_1^k ein Element vom Grade r.

3. Sind $A, B, \dots$ irgend welche Elemente in S von den Graden $a, b, \dots$, so ist auch $AB \dots$ ein Element in S, und der Grad dieses Produktes ist ein Teiler des kleinsten gemeinschaftlichen Vielfachen m von $a, b, \dots$

Denn es ist

$$(AB \dots)^m = A^m B^m \dots = 1,$$

und folglich m ein Vielfaches des Grades von $AB \dots$

4. Ist der Grad n der Gruppe S in zwei Faktoren $n = ab$ zerlegt, so daß a und b relativ prim sind, so gibt es in S genau a Elemente A, deren Grad ein Teiler von a ist, und b Elemente B, deren Grad ein Teiler von b ist, und in der Form

$$\Theta = AB \tag{5}$$

sind sämtliche Elemente von S und jedes nur einmal enthalten.

Um die Richtigkeit dieses Satzes einzusehen, stellen wir folgende Überlegung an. Der Inbegriff $\mathfrak{A}$ der Elemente A, deren Grad ein Teiler von a ist, ist eine in S enthaltene Gruppe; denn sind A, A' zwei solche Elemente, so ist nach 3. der Grad von AA' ein Teiler von a, und AA' ist also auch in $\mathfrak{A}$ enthalten.

Ebenso ist der Inbegriff $\mathfrak{B}$ der Elemente B, deren Grad ein Teiler von b ist, eine Gruppe. Das Element 1 kommt sowohl in $\mathfrak{A}$ als in $\mathfrak{B}$ vor, sonst enthalten beide kein gemeinschaftliches Element.

Der Grad a' von $\mathfrak{A}$ ist relativ prim zu b. Denn ist r eine in a' aufgehende Primzahl, so gibt es nach 2. in $\mathfrak{A}$ ein Element

vom Grade r. Da aber der Grad jedes Elementes von $\mathfrak{A}$ ein Teiler von a ist, so ist auch r ein Teiler von a und nicht von b.

Ebenso beweist man, daß der Grad b' von $\mathfrak{B}$ relativ prim zu a ist.

Nun bestimme man zwei ganze Zahlen x und y aus der diophantischen Gleichung

$$ax + by = 1 \tag{6}$$

und nehme irgend ein Element Θ von S. Dann ist

$$\Theta = \Theta^{ax}\,\Theta^{by}, \tag{7}$$

und nun ist, wegen $ab = n$

$$(\Theta^{by})^a = 1, \quad (\Theta^{ax})^b = 1,$$

also Θ^{by} in $\mathfrak{A}$ und Θ^{ax} in $\mathfrak{B}$ enthalten.

Also folgt aus (7):

$$\Theta = AB. \tag{8}$$

Demnach ist jedes Element Θ in der Form AB enthalten. Eine solche Darstellung ist aber auch nur auf eine Art möglich. Denn sind A', B' zwei Elemente aus $\mathfrak{A}$ und $\mathfrak{B}$, und ist

$$AB = A'B',$$

so folgt, wenn man in die Potenz $by = 1 - ax$ erhebt, $A = A'$ und folglich auch $B = B'$. Die Anzahl der verschiedenen Produkte der Form AB ist aber $= a'b'$, und daher hat man

$$n = ab = a'b',$$

und da a relativ prim zu b' und b relativ prim zu a' ist:

$$a' = a, \quad b' = b.$$

Damit ist der Satz 4 in allen seinen Teilen bewiesen.

Wenn nun die beiden Gruppen $\mathfrak{A}$, $\mathfrak{B}$ durch Basen dargestellt sind, so folgt aus der Formel (8), daß auch S durch eine Basis dargestellt ist, und die Basis von S enthält die Elemente der Basen von $\mathfrak{A}$ und $\mathfrak{B}$ und keine anderen.

Wenn a und b weiter in Faktoren zerlegbar sind, die zueinander relativ prim sind, so können wir mit den Gruppen $\mathfrak{A}$ und $\mathfrak{B}$ wieder ebenso verfahren, und wir kommen also zu dem Resultate, daß unser Theorem I. allgemein bewiesen ist, wenn wir es noch für Gruppen nachweisen können, deren Grad eine Potenz einer Primzahl ist.

Es sei also jetzt der Grad n der Gruppe S eine Potenz einer Primzahl p

$$n = p^k. \tag{9}$$

Die Grade aller Elemente von S, die ja Divisoren von n sind, müssen dann gleichfalls Potenzen von p sein. Es ist nachzuweisen, daß eine solche Gruppe durch eine Basis darstellbar ist.

Man wähle in S ein Element A von möglichst hohem Grade a. Dann ist a eine Potenz von p und die Grade aller anderen Elemente sind Teiler von a, so daß für jedes Element Θ von S

$$\Theta^a = 1 \tag{10}$$

ist. Die Elemente

$$1, A, A^2, \ldots A^{a-1} \tag{11}$$

sind alle voneinander verschieden, und ihre Gesamtheit ist ein Teiler von S. Ist damit die Gruppe S erschöpft, ist also jedes Element von der Form A^α, so ist S durch eine eingliedrige Basis dargestellt. Wenn aber S mit (11) noch nicht erschöpft ist, so wird es doch für jedes Element Θ von S einen gewissen Exponenten h geben, so daß Θ^h in der Reihe (11) enthalten ist. Gewiß wird das eintreten, wenn h der Grad von Θ, also $\Theta^h = 1$ ist. Unter diesen Zahlen h wird eine die kleinste positive sein, die wir mit b bezeichnen wollen. Es gibt also für jedes Element Θ eine gewisse kleinste positive Zahl b, so daß

$$\Theta^b = A^\lambda \tag{12}$$

in der Reihe (11) enthalten ist.

Diese Zahl b ist ein Teiler von a, also auch eine Potenz von p und zugleich ein Teiler von λ. Denn setzen wir $a = qb + b'$, wo $0 \leqq b' < b$ ist, so folgt aus (10) und (12)

$$\Theta^a = A^{\lambda q}\,\Theta^{b'} = 1,$$

oder $\Theta^{b'} = A^{-\lambda q}$, und wenn also b' nicht Null ist, so gibt es gegen die Voraussetzung eine noch kleinere Zahl als b, nämlich b', die der Forderung (12) genügt.

Aus (12) folgt ferner

$$1 = \Theta^a = A^{\frac{\lambda a}{b}};$$

also muß $\lambda a : b$ ein Vielfaches von a und folglich λ ein Vielfaches von b sein. Wenn wir daher

$$B = \Theta\, A^{-\frac{\lambda}{b}} \tag{13}$$

setzen, so ist $B^b = 1$, und zugleich ist B^b die niedrigste Potenz von B, die einer Potenz von A gleich wird, weil, wenn $B^{b'}$ eine Potenz von A ist, dasselbe nach (13) auch von $\Theta^{b'}$ gilt. Wir

nehmen das Element Θ so gewählt an, daß b so groß als möglich wird. Dann ist auch für jedes andere Element Θ_1 aus S immer Θ_1^b eine Potenz von A, deren Exponent durch b teilbar ist.

Ist nun

$$\alpha = 0, 1, 2, \ldots a - 1$$
$$\beta = 0, 1, 2, \ldots b - 1,$$

so sind die Elemente

$$A^\alpha B^\beta \tag{14}$$

in der Anzahl ab alle voneinander verschieden. Sie bilden einen in S enthaltenen Teiler S', der durch die Basis A, B dargestellt ist. Ist S' mit S identisch, so sind wir am Ziele.

Ist aber S durch (14) noch nicht erschöpft, so fahren wir in derselben Weise fort.

Wir wollen durch Anwendung der vollständigen Induktion gleich allgemein schließen.

Es sei ein Teiler $S_{\nu-1}$ von S ermittelt, der durch eine Basis in der Form dargestellt ist:

$$S_{\nu-1} = A_1^{\alpha_1} A_2^{\alpha_2} \ldots A_{\nu-1}^{\alpha_{\nu-1}}, \tag{15}$$

von dem wir folgende Voraussetzungen machen:

1. Die Grade von $A_1, A_2, \ldots A_{\nu-1}$ seien $a_1, a_2, \ldots a_{\nu-1}$, und es sei

$$a_1 \geqq a_2 \geqq \cdots \geqq a_{\nu-1}.$$

2. Für jedes in S enthaltene Element Θ sei

$$\Theta^{a_{\nu-1}} = A_1^{\lambda_1} A_2^{\lambda_2} \ldots A_{\nu-2}^{\lambda_{\nu-2}}, \tag{16}$$

d. h. in $S_{\nu-2}$ enthalten, und die Exponenten $\lambda_1, \lambda_2 \ldots \lambda_{\nu-2}$ seien durch $a_{\nu-1}$ teilbar.

Die oben abgeleitete Gruppe S' genügt diesen Forderungen, wenn $\nu = 3$ und $a_{\nu-1} = b$ gesetzt wird, und es ist nun zu zeigen, wie man, wenn $S_{\nu-1}$ noch nicht die ganze Gruppe S erschöpft, daraus eine ebensolche umfassendere Gruppe S_ν ableiten kann.

Für jedes Element Θ von S wird es einen gewissen niedrigsten positiven Exponenten a_ν geben, für den Θ^{a_ν} in $S_{\nu-1}$ enthalten ist, und dies a_ν ist wegen (16) gleich oder kleiner als $a_{\nu-1}$, und ist außerdem eine Potenz von p, da es ein Teiler des Grades von Θ sein muß. Wir wählen Θ so, daß der Exponent a_ν so groß als möglich wird. Ist Θ_1 ein beliebiges anderes Element in S, und Θ_1^h die niedrigste Potenz von Θ_1, die in $S_{\nu-1}$ enthalten ist, so ist auch h eine Potenz von p und gleich oder kleiner als a_ν. Es ist daher Θ_1^a in $S_{\nu-1}$ enthalten.

Setzen wir aber

$$\Theta_1^{a_\nu} = A_1^{\lambda_1} A_2^{\lambda_2} \dots A_{\nu-1}^{\lambda_{\nu-1}}, \tag{17}$$

so sind die sämtlichen Exponenten λ durch a_ν teilbar. Denn weil $a_{\nu-1} \geqq a_\nu$ und beides Potenzen von p sind, ist $a_{\nu-1}, a_\nu$ eine ganze Zahl, und wenn man (17) in die Potenz $a_{\nu-1}, a_\nu$ erhebt, so ergibt sich:

$$\Theta_1^{a_{\nu-1}} = A_1^{\frac{\lambda_1 a_{\nu-1}}{a_\nu}} A_2^{\frac{\lambda_2 a_{\nu-1}}{a_\nu}} \dots A_{\nu-1}^{\frac{\lambda_{\nu-1} a_{\nu-1}}{a_\nu}},$$

und nach der Voraussetzung 2 müssen die Exponenten von $A_1, A_2, \dots A_{\nu-1}$ auf der rechten Seite dieser Formel durch $a_{\nu-1}$ teilbare ganze Zahlen sein. Folglich sind $\lambda_1, \lambda_2, \dots \lambda_{\nu-1}$ durch a_ν teilbar. Wir können also, indem wir zu dem oben betrachteten speziellen Θ zurückkehren und unter $h_1, h_2 \dots h_{\nu-1}$ ganze Zahlen verstehen,

$$\Theta^{a_\nu} = A_1^{h_1 a_\nu} A_2^{h_2 a_\nu} \dots A_{\nu-1}^{h_{\nu-1} a_\nu}$$

setzen, und wenn wir dann

$$A_\nu = \Theta A_1^{-h_1} A_2^{-h_2} \dots A_{\nu-1}^{-h_{\nu-1}}$$

annehmen so ist

$$A_\nu^{a_\nu} = 1 \tag{18}$$

die niedrigste Potenz von A_ν, die in $S_{\nu-1}$ enthalten ist (weil sonst auch noch eine niedrigere Potenz von Θ in $S_{\nu-1}$ enthalten wäre).

Dann ist

$$A_1^{\alpha_1} A_2^{\alpha_2} \dots A_\nu^{\alpha_\nu}, \qquad \alpha_i = 0, 1, 2, \dots a_i - 1 \tag{19}$$

nur dann $= 1$, wenn $\alpha_1, \alpha_2, \dots \alpha_\nu$ durch $a_1, a_2, \dots a_\nu$ teilbar und folglich $= 0$ sind, und demnach sind die in (19) dargestellten Elemente alle voneinander verschieden. Diese Elemente bilden aber eine Gruppe S_ν mit der Basis $A_1, A_2, \dots A_\nu$, die der Forderung 1. genügt.

Zugleich ist, wie die Formel (17) zeigt, wenn Θ_1 ein beliebiges Element in S ist, $\Theta_1^{a_\nu}$ in $S_{\nu-1}$ enthalten, und die Exponenten $\lambda_1, \lambda_2, \dots \lambda_{\nu-1}$ sind durch a_ν teilbar. Es ist daher auch die Forderung 2 befriedigt.

Damit ist also unser Satz I. bewiesen. Zugleich sehen wir aus dieser Ableitung, daß man für eine beliebige Abelsche Gruppe S die Elemente der Basis immer so annehmen kann, daß ihre Grade Primzahlpotenzen sind.

Der Beweis, den wir hier für die Möglichkeit der Darstellung einer Abelschen Gruppe durch eine Basis mitgeteilt haben, enthält zugleich einen Weg, eine solche Basis zu finden, und zwar eine, bei der die Grade der Basiselemente Primzahlpotenzen sind. Gleichwohl kann es vorkommen, daß eine und dieselbe Gruppe auf verschiedene Arten durch Basen dargestellt werden kann.

Betrachten wir z. B. zwei Elemente A, B, deren Grade a, b relativ prim sind, so wird durch diese als Elemente einer zweigliedrigen Basis eine Gruppe

$$A^\alpha B^\beta \qquad \begin{matrix} \alpha = 0, 1, 2, \dots a - 1 \\ \beta = 0, 1, 2, \dots b - 1 \end{matrix}$$

dargestellt. Dieselbe Gruppe kann aber auch durch die eingliedrige Basis AB dargestellt werden in der Form

$$(AB)^s \qquad s = 0, 1, 2, \dots ab - 1.$$

Denn sind α und β beliebig gegeben, so kann man s nach dem Modul ab so bestimmen, daß

$$s \equiv \alpha \pmod{a}, \quad s \equiv \beta \pmod{b},$$

wodurch $A^\alpha B^\beta$ und $(AB)^s$ identisch werden. Trotzdem ist in gewissem Sinne die Konstitution der Basis durch die Natur der Gruppen völlig bestimmt, wie folgender Satz besagt:

II. Sind

$$A_1, A_2, \dots A_\nu$$

mit den Graden

$$a_1, a_2, \dots a_\nu$$

und

$$B_1, B_2, \dots B_\mu$$

mit den Graden

$$b_1, b_2, \dots b_\mu$$

zwei Basen einer Abelschen Gruppe S vom Grade n, ist p eine in n aufgehende Primzahl und

$$a_1 = p_1 a'_1, \; a_2 = p_2 a'_2, \dots a_\nu = p_\nu a'_\nu,$$
$$b_1 = p'_1 b'_1, \; b_2 = p'_2 b'_2, \dots b_\mu = p'_\mu b'_\mu,$$

worin $p_1, p_2, \dots p_\nu, p'_1, p'_2, \dots p'_\mu$ die höchsten Potenzen von p sind, die in $a_1, a_2, \dots a_\nu, b_1, b_2, \dots b_\mu$ aufgehen, so kommen alle Primzahlpotenzen $p_1, p_2, \dots p_\nu$, die größer als 1 sind, auch unter den $p'_1, p'_2, \dots p'_\mu$ vor, und umgekehrt.

Um diesen Satz zu beweisen, nehmen wir die Elemente der beiden Basen A und B so geordnet an, daß

$$p_1 \geqq p_2 \geqq \cdots \geqq p_\nu \tag{20}$$
$$p'_1 \geqq p'_2 \geqq \cdots \geqq p'_\mu.$$

Die ganzen Zahlen $a'_1, a'_2, \ldots a'_\nu, b'_1, b'_2, \ldots b'_\mu$ sind nach ihrer Definition durch p nicht teilbar, und wenn wir also mit m das kleinste gemeinschaftliche Vielfache von $a'_1, a'_2, \ldots a'_\nu$ bezeichnen, so ist auch m nicht durch p teilbar. Ist aber

$$\Theta = A_1^{\alpha_1} A_2^{\alpha_2} \ldots A_\nu^{\alpha_\nu} \tag{21}$$

ein beliebiges Element von S, so folgt, daß

$$\Theta^{p_1 m} = 1$$

sein muß.

Setzt man hierin $B_1, B_2, \ldots B_\mu$ für Θ, so folgt, daß $p_1 m$ durch jede der Zahlen $b_1, b_2 \ldots b_\mu$ teilbar ist. Es ist also p_1 teilbar durch p'_1, und m durch das kleinste gemeinschaftliche Vielfache von $b'_1, b'_2, \ldots b'_\mu$. In diesem Schlusse können wir nun durchweg A mit B vertauschen. Es ist also auch p'_1 durch p_1 teilbar, und daher

$$p_1 = p'_1,$$

und außerdem ergibt sich, daß m auch das kleinste gemeinschaftliche Vielfache von $b'_1, b'_2, \ldots b'_\mu$ ist.

Um daraus unseren Satz allgemein zu beweisen, nehmen wir an, es sei für irgend ein s bewiesen:

$$p_1 = p'_1, \; p_2 = p'_2, \ldots p_{s-1} = p'_{s-1}.$$

Es ist unter dieser Voraussetzung das System

$$A_1^{p_s m}, \; A_2^{p_s m}, \ldots A_{s-1}^{p_s m} \tag{22}$$

die Basis eines Teilers S' von S, dessen Grad z sich gleich

$$\frac{p_1}{p_s} \frac{p_2}{p_s} \ldots \frac{p_{s-1}}{p_s} \tag{23}$$

ergibt. Von den beiden Zahlen p_s, p'_s wird, wenn sie nicht gleich sind, eine die größere sein. Wir nehmen an, es sei

$$p'_s \gtrless p_s. \tag{24}$$

Die Gruppe S' ist aber nach dieser Voraussetzung auch durch die Elemente

$$B_1^{p_s m}, \; B_2^{p_s m}, \ldots B_{s-1}^{p_s m} \tag{25}$$

ausdrückbar, aus denen man die Zahl der Elemente

$$z' = \frac{p_1}{p'_s}\,\frac{p_2}{p'_s} \dots \frac{p_{s-1}}{p'_s}$$

erhält, und z' kann jedenfalls nicht kleiner als z sein, also

folglich
$$p'_s \gtreqless p_s,$$
$$p'_s = p_s.$$

Hiermit ist unser Theorem II. bewiesen.

Die in den Gradzahlen einer Basis von S enthaltenen Primzahlpotenzen sind also von der besonderen Wahl der Basis unabhängig und wir nennen sie daher die Invarianten der Gruppe. Das Produkt aller Invarianten ist gleich dem Grade n der Gruppe.

Da wir für S eine Basis bestimmen können, bei der die Grade der Elemente lauter Primzahlpotenzen sind, so sind nach dem Theorem II. diese Grade die Invarianten der Gruppe und sind also durch die Gruppe vollständig bestimmt.

Daß die Invarianten auch die Natur der Gruppe vollständig bestimmen, ergibt sich aus dem Satze:

III. Zwei Abelsche Gruppen mit denselben Invarianten sind isomorph, und isomorphe Abelsche Gruppen haben dieselben Invarianten.

Wenn nämlich zwei Gruppen S und S' der Anzahl und dem Grade nach übereinstimmende Basiselemente haben, wenn etwa

$$\Theta = A_1^{\alpha_1} A_2^{\alpha_2} \dots A_\nu^{\alpha_\nu}$$

die Elemente von S sind, und

$$\Theta' = B_1^{\alpha_1} B_2^{\alpha_2} \dots B_\nu^{\alpha_\nu}$$

die Elemente von S', wo die $\alpha_1, \alpha_2, \dots \alpha_\nu$ in beiden Fällen Restsysteme nach den Moduln $a_1, a_2, \dots a_\nu$ durchlaufen, so brauchen wir nur Θ und Θ' dann einander entsprechen zu lassen, wenn die Exponenten α in beiden dieselben sind; dann sind beide Gruppen isomorph aufeinander bezogen.

Haben also zwei Gruppen dieselben Invarianten, so können wir diese Invarianten in beiden Gruppen zu Graden der Basiselemente machen und erhalten dann diesen Fall.

Wenn umgekehrt zwei Abelsche Gruppen isomorph sind, so bilden die den Basiselementen der einen Gruppe entsprechenden Elemente der anderen eine Basis der letzteren, und also müssen auch die Invarianten in beiden dieselben sein.

Neunter Abschnitt.

Die Galoissche Theorie.

§ 53.

Adjunktion. Algebraische Körper.

Wenn einem Körper irgend welcher Größen (§ 13), den wir mit Ω bezeichnen wollen, irgend eine Größe α hinzugefügt wird die nicht in ihm enthalten ist, so entsteht ein neues Größensystem

$$\Omega, \alpha,$$

das aber kein Körper sein wird; um daraus einen Körper abzuleiten, muß man alle Größen hinzufügen, die sich durch die Verbindung von α mit den Größen von Ω mittels der Grundrechnungsarten ableiten lassen. So erhält man einen erweiterten Körper Ω', der zugleich α und Ω enthält, und der durch α und Ω völlig bestimmt ist. Wir wollen ihn den Körper Ω, α nennen

Dies Hinzufügen einer neuen Größe α zu einem Körper Ω heißt Adjungieren, und man sagt, der Körper Ω' entsteht aus Ω durch Adjunktion von α.

Zu Ω' kann man wieder eine Größe α' adjungieren, und erhält einen dritten Körper Ω'', der dann Ω, α und α' enthält usf. Man kann aber auch gleichzeitig zwei oder mehrere Zahlen zu Ω adjungieren, und erhält so denselben Körper Ω'', wenn man gleichzeitig α und α' dem Ω adjungiert. Man kann auch einem Körper einen zweiten Körper adjungieren und erhält so einen Körper, der zugleich die Zahlen der beiden Körper enthält. In dem obigen Beispiel würde man durch Adjunktion des Körpers Ω, α' zu dem Körper $\Omega' = \Omega, \alpha$ wieder den Körper Ω'' erhalten.

Adjungiert man einem Körper einen seiner Teile, so entsteht kein neuer Körper.

So erhält man z. B., wenn man zum Körper der rationalen Zahlen, der mit $\mathfrak{R}$ bezeichnet sein mag, die Zahl $i = \sqrt{-1}$

adjungiert, den Körper der komplexen Zahlen $x + yi$, mit rationalen x, y, den wir $\mathfrak{J}$ nennen wollen. Durch Adjunktion einer Variablen u zum Körper $\mathfrak{R}$ erhält man den Körper, der aus allen ganzen und gebrochenen rationalen Funktionen von u mit rationalen Koeffizienten besteht, durch Adjunktion von u und i zu $\mathfrak{R}$ den Körper der rationalen Funktionen von u, deren Koeffizienten Zahlen in $\mathfrak{J}$ sind usf.

Wenn Ω ein beliebiger Körper und $f(x)$ eine Funktion in Ω ist, so sind zwei Fälle möglich.

Die Funktion $f(x)$ ist entweder zerlegbar oder nicht zerlegbar in Faktoren niedrigeren Grades, deren Koeffizienten dem Körper Ω angehören. Im ersten Falle heißt die Funktion $f(x)$ reduzibel, im zweiten irreduzibel in Ω.

Eine lineare Funktion ist selbstverständlich in jedem Körper irreduzibel. Multipliziert man mehrere Funktionen in Ω miteinander, so entsteht eine Funktion derselben Form, die aber dann natürlich reduzibel ist, da sie wieder in die Faktoren zerlegt werden kann, aus denen sie entstanden ist.

Eine Funktion $f(x)$, die in Ω irreduzibel ist, kann in einem erweiterten Körper Ω', der aus Ω durch irgend eine Adjunktion entsteht, reduzibel werden; so wird jede Funktion $f(x)$ reduzibel in dem Körper Ω', der aus Ω durch Adjunktion einer Wurzel α von $f(x)$ entsteht; denn es kann von $f(x)$ ein linearer Faktor $x - \alpha$ abgesondert werden. In dem Körper, der aus allen Zahlen besteht, ist jede Funktion mit Zahlenkoeffizienten reduzibel.

I. Eine irreduzible Funktion $f(x)$ kann mit einer anderen Funktion $F(x)$, deren Koeffizienten demselben Körper Ω angehören, keinen gemeinsamen Teiler haben, wenn nicht $F(x)$ durch $f(x)$ teilbar ist.

Dieser Satz, der uns in der Folge noch sehr wichtige Dienste leisten wird, ist fast selbstverständlich; denn der größte gemeinschaftliche Teiler zweier Funktionen $F(x)$ und $f(x)$ läßt sich durch rationale Rechnung finden, und ist daher auch in Ω enthalten. Da nun aber $f(x)$ keinen in Ω enthaltenen Teiler hat als sich selbst oder eine Konstante (d. h. eine in Ω enthaltene Größe), so muß also dieser größte gemeinschaftliche Teiler entweder eine Konstante oder $f(x)$ selbst sein.

Es folgt aus diesem Satze, daß eine irreduzible Funktion niemals mehrfache Faktoren haben kann; denn sie müßte sonst mit ihrer Ableitung $f'(x)$ einen gemeinsamen Teiler haben; also müßte $f'(x)$ durch $f(x)$ teilbar sein, was unmöglich ist, da der Grad von $f'(x)$ niedriger ist als der von $f(x)$.

Wir können diese Sätze auch so ausdrücken:

Ist $f(x)$ irreduzibel, und verschwindet $F(x)$ für eine Wurzel von $f(x)$, so verschwindet $F(x)$ auch für alle anderen Wurzeln von $f(x)$.

Insbesondere können wir daraus schließen:

Ist $F(x)$ von niedrigerem Grade, als die irreduzible Funktion $f(x)$, und verschwindet $F(x)$ für eine Wurzel von $f(x)$, so muß $F(x)$ identisch verschwinden, d. h. alle Koeffizienten von $F(x)$ müssen Null sein.

Ist Ω ein beliebiger Körper, und $F(x)$ eine Funktion in Ω, so heißt die Gleichung

$$F(x) = 0 \tag{1}$$

eine Gleichung in Ω. Diese Gleichung heißt reduzibel oder irreduzibel, je nachdem die Funktion $F(x)$ reduzibel oder irreduzibel ist. Durch Adjunktion einer Wurzel α einer solchen Gleichung zu Ω entsteht (wenn α nicht selbst schon zu Ω gehört) ein neuer Körper Ω', den wir einen algebraischen Körper „über“ Ω oder auch, wenn Zweifel über die Bedeutung ausgeschlossen sind, kurz einen algebraischen Körper nennen wollen. Wir brauchen für einen solchen Körper das Zeichen

$$\Omega(\alpha).$$

Sind $\beta, \gamma \ldots$ Wurzeln von anderen Gleichungen in Ω oder auch andere Wurzeln derselben Gleichung, so erhalten wir durch gleichzeitige Adjunktion von $\alpha, \beta, \gamma \ldots$ gleichfalls algebraische Körper über Ω, die wir mit

$$\Omega(\alpha, \beta, \gamma \ldots)$$

bezeichnen. Wir werden gleich sehen, daß diese Erweiterung des Begriffes algebraischer Körper nur scheinbar ist, und bleiben also zunächst bei der Adjunktion einer Größe α, also bei der Betrachtung von $\Omega(\alpha)$ stehen.

Die Gleichung (1), deren Wurzel α ist, kann reduzibel sein. Unter den irreduzibeln Faktoren von $F(x)$ ist aber wenigstens einer, der für $x = \alpha$ verschwindet, den wir mit $f(x)$ bezeichnen wollen; diese Funktion $f(x)$ ist, wenn wir den Koeffizienten der

höchsten Potenz von x gleich 1 annehmen, völlig bestimmt, weil $x = \alpha$ nicht die Wurzel von zwei verschiedenen irreduzibeln Gleichungen sein kann.

Die Gleichung $f(x) = 0$ hat also die Form:

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0, \tag{2}$$

worin $a_1, a_2, \ldots a_n$ Größen in Ω sind. Ist n der Grad dieser Gleichung, so nennen wir n auch den Grad des Körpers $\Omega(\alpha)$.

Alle Größen Θ des Körpers $\Omega(\alpha)$ lassen sich ableiten durch Addition, Subtraktion, Multiplikation und Division aus α und aus Zahlen in Ω; sie lassen sich also darstellen als rationale Funktionen von α mit Koeffizienten in Ω, oder in der Form

$$\Theta = \frac{\chi(\alpha)}{\psi(\alpha)}, \tag{3}$$

worin $\chi(x)$ und $\psi(x)$ Funktionen in Ω sind. Da aber $\psi(\alpha)$ von Null verschieden sein muß, so ist $\psi(x)$ nicht durch $f(x)$ teilbar, und da $f(x)$ irreduzibel angenommen ist, so ist $\psi(x)$ relativ prim zu $f(x)$.

Danach können wir nach § 15, III. die Funktionen $Q(x)$, $\varphi(x)$ in Ω, und zwar $\varphi(x)$ höchstens vom Grade $n - 1$, so bestimmen, daß

$$Q(x) f(x) + \varphi(x) \psi(x) = \chi(x) \tag{4}$$

wird, und wenn wir also hierin $x = \alpha$ setzen, so daß $f(\alpha) = 0$ wird, so folgt

$$\frac{\chi(\alpha)}{\psi(\alpha)} = \varphi(\alpha). \tag{5}$$

Bezeichnen wir die Koeffizienten von $\varphi(x)$, die, wie wir gesehen haben, Größen in Ω sind, mit $c_0, c_1, \ldots c_{n-1}$, so kann also jede Größe Θ in $\Omega(\alpha)$ in der Form dargestellt werden:

$$\Theta = c_0 + c_1 \alpha + c_2 \alpha^2 + \cdots + c_{n-1} \alpha^{n-1}.$$

Diese Darstellung ist nur auf eine Art möglich; denn ist gleichzeitig

$$\Theta = c_0' + c_1' \alpha + c_2' \alpha^2 + \cdots + c_{n-1}' \alpha^{n-1}, \tag{6}$$

so verschwindet die Funktion $(n-1)$ten Grades

$$c_0 - c_0' + (c_1 - c_1') x + (c_2 - c_2') x^2 + \cdots + (c_{n-1} - c_{n-1}') x^{n-1}$$

für $x = \alpha$. Das ist aber wegen der Irreduzibilität von $f(x)$ nur möglich, wenn

$$c_0 = c_0', \; c_1 = c_1', \ldots c_{n-1} = c_{n-1}'$$

ist.

Wir können also nach Fixierung des Rationalitätsbereichs auch sagen:

Der Körper $\Omega(\alpha)$ besteht aus allen rationalen Funktionen von α in Ω.

§ 54.

Gleichzeitige Adjunktion mehrerer algebraischer Größen.

Es soll jetzt zunächst nachgewiesen werden, daß man die gleichzeitige Adjunktion mehrerer algebraischer Größen durch die Adjunktion einer einzigen ersetzen kann; mit anderen Worten, daß jeder Körper $\Omega(\alpha, \beta, \gamma \ldots)$ angesehen werden kann als ein Körper $\Omega(\alpha)$.

Es seien also $\alpha, \beta, \gamma \ldots$ algebraische Größen in beliebiger Anzahl und

$$A(x),\ B(x),\ C(x) \ldots \tag{1}$$

Funktionen in Ω, deren Wurzeln $\alpha, \beta, \gamma \ldots$ sind. Keine dieser Funktionen soll eine mehrfache Wurzel haben, eine Voraussetzung, durch die die Allgemeinheit nicht beschränkt wird. Dagegen ist nicht ausgeschlossen, daß unter diesen Funktionen dieselbe mehrmals vorkommt, wenn etwa α, β verschiedene Wurzeln einer Funktion sind.

Wir haben im vorigen Paragraphen gesehen, daß sich jede Größe eines Körpers $\Omega(\alpha)$ als ganze rationale Funktion in Ω von α darstellen läßt. Setzen wir in diesem Satze an Stelle des Körpers Ω den Körper $\Omega(\beta, \gamma \ldots)$, so folgt, daß jede Zahl in $\Omega(\alpha, \beta, \gamma \ldots)$ als ganze rationale Funktion von α mit Koeffizienten aus $\Omega(\beta, \gamma \ldots)$ dargestellt werden kann; und wenn wir dieselbe Schlußweise in Beziehung auf die Koeffizienten wiederholen, so ergibt sich:

Jede Größe des Körpers $\Omega(\alpha, \beta, \gamma \ldots)$ kann als ganze rationale Funktion in Ω von $\alpha, \beta, \gamma \ldots$ dargestellt werden.

Wir bilden eine lineare Funktion der $\alpha, \beta, \gamma \ldots$

$$\xi = x\alpha + y\beta + z\gamma + \cdots,$$

und beachten, daß jede der Gleichungen (1) nicht nur eine, sondern mehrere Wurzeln hat, deren Zahl gleich dem Grade der Gleichung ist. Ist also $\alpha', \beta', \gamma', \ldots$ irgend eine von $\alpha, \beta, \gamma \ldots$ verschiedene Kombination von je einer Wurzel von jeder der Gleichungen (1), so setzen wir:

$$\xi' = x\alpha' + y\beta' + z\gamma' + \cdots,$$

und bilden auf die gleiche Weise ξ'', ξ''' ... Die Anzahl der so gebildeten Größen ξ ist, wenn a der Grad von $A(x)$, b der Grad von $B(x)$, c der Grad von $C(x)$ ist usf.,

$$m = abc \ldots$$

Die Differenzen $\xi - \xi'$, $\xi - \xi''$, $\xi' - \xi''$, ... sind lineare Funktionen von x, y, z ... und keine von ihnen ist identisch Null, da wir angenommen haben, daß keine der Funktionen (1) gleiche Wurzeln habe. Wir können also nach dem in § 14 bewiesenen Satze für x, y, z ... solche rationale Zahlen setzen, daß keine von diesen Differenzen verschwindet, daß also die m Werte ξ, ξ', ξ'' ... alle voneinander verschieden sind.

Nun ist jede rationale Funktion, die in bezug auf die Wurzeln jeder der Funktionen (1) symmetrisch ist, nach dem Satz von den symmetrischen Funktionen eine Zahl in Ω. Dazu gehören auch die Koeffizienten der Funktion $m$ten Grades von t:

$$F(t) = (t - \xi)\,(t - \xi')\,(t - \xi'') \ldots, \tag{2}$$

und $F(t) = 0$ ist also eine Gleichung in Ω, die keine gleichen Wurzeln hat, und deren eine Wurzel ξ ist.

Ist ferner Θ eine Größe in Ω $(\alpha, \beta, \gamma \ldots)$, also eine ganze rationale Funktion von $\alpha, \beta, \gamma \ldots$, und bezeichnen wir mit Θ', Θ''... die Größen, die aus dieser rationalen Funktion hervorgehen, wenn die Argumente durch α', β', γ'... oder $\alpha'', \beta'', \gamma''$ usf. ersetzt werden, so ist

$$F(t)\left(\frac{\Theta}{t - \xi} + \frac{\Theta'}{t - \xi'} + \frac{\Theta''}{t - \xi''} + \cdots\right) = \psi(t)$$

als symmetrische Funktion der α und der β ... eine ganze rationale Funktion $(m - 1)$ten Grades von t in Ω, und wenn man $t = \xi$ setzt, so folgt

$$\Theta = \frac{\psi(\xi)}{F'(\xi)}, \tag{3}$$

und Θ ist also in $\Omega(\xi)$ enthalten. Da umgekehrt auch jede Größe in $\Omega(\xi)$ zugleich in $\Omega(\alpha, \beta, \gamma \ldots)$ enthalten ist, so sind beide Körper identisch, d. h. es ist

$$\Omega(\xi) = \Omega(\alpha, \beta, \gamma \ldots).$$

Damit ist unsere Behauptung erwiesen.

§ 55.
Primitive und imprimitive Körper.

Nach dem zuletzt bewiesenen Satze beschränken wir uns jetzt auf die Betrachtung algebraischer Körper $\Omega(\alpha)$, worin α die Wurzel einer Gleichung in Ω ist. Die Gleichung möge, von mehrfachen Faktoren befreit, mit

$$F(x) = 0 \tag{1}$$

bezeichnet und vom Grade m sein. Ihre Wurzeln seien

$$\alpha, \alpha_1, \alpha_2, \dots \alpha_{m-1}. \tag{2}$$

Es kann $F(x)$ in Ω reduzibel sein; dann wird es einen und nur einen irreduzibeln Faktor $f(x)$ von $F(x)$ geben, so daß α eine Wurzel der Gleichung

$$f(x) = 0 \tag{3}$$

ist. Der Grad von $f(x)$ sei n und die Wurzeln von (3), die alle unter den Wurzeln (2) enthalten sind, seien

$$\alpha, \alpha_1, \alpha_2, \dots \alpha_{n-1}. \tag{4}$$

Der Grad des Körpers $\Omega(\alpha)$ ist dann gleichfalls n.

Jede der Größen (4) gibt zu einem algebraischen Körper Anlaß, und so entstehen die Körper

$$\Omega(\alpha), \Omega(\alpha_1), \dots \Omega(\alpha_{n-1}), \tag{5}$$

die wir die mit $\Omega(\alpha)$ konjugierten Körper nennen. Es kann sein, daß diese Körper alle oder teilweise identisch sind, sie können aber auch alle voneinander verschieden sein.

Nach § 53 erhalten wir jede Größe Θ in $\Omega(\alpha)$, wenn wir in einer ganzen rationalen Funktion $\chi(t)$ in Ω, höchstens vom Grade $(n-1)$, für die Variable t die Zahl α setzen.

Setzen wir für t in $\chi(t)$ die verschiedenen Größen (4), so erhalten wir n Größen, eine aus jedem der konjugierten Körper

$$\Theta = \chi(\alpha), \Theta_1 = \chi(\alpha_1), \dots \Theta_{n-1} = \chi(\alpha_{n-1}), \tag{6}$$

und diese heißen die mit Θ konjugierten Größen.

Jede symmetrische Funktion dieser Größen ist zugleich eine symmetrische Funktion der Wurzeln der Gleichung (3), und mithin in Ω enthalten.

Unter diesen symmetrischen Funktionen wollen wir zwei häufig vorkommende durch besondere Namen und Bezeichnungen hervorheben; es ist die Summe

$$S(\Theta) = \Theta + \Theta_1 + \Theta_2 + \dots + \Theta_{n-1}, \tag{7}$$

die wir die Spur von Θ nennen, und das Produkt

$$N(\Theta) = \Theta\,\Theta_1\,\Theta_2 \dots \Theta_{n-1}, \tag{8}$$

das die Norm von Θ heißt. Konjugierte Zahlen haben hiernach dieselben Spuren und Normen.

Das Produkt

$$(t - \Theta)\,(t - \Theta_1) \dots (t - \Theta_{n-1}) = \Phi(t) \tag{9}$$

ist eine ganze rationale Funktion $n$ten Grades von t in Ω und ihre Wurzeln sind Θ und die mit Θ konjugierten Größen. Daraus ergibt sich der Satz:

II. Jede Größe Θ in $\Omega(\alpha)$ ist Wurzel einer Gleichung $n$ten Grades in Ω, deren übrige Wurzeln die mit Θ konjugierten Größen sind.

Die Berechnung der Funktion $\Phi(t)$ aus $f(x)$ heißt eine Tschirnhausen-Transformation von $f(x)$.

Wir haben nun die Funktion $\Phi(t)$ auf ihre Irreduzibilität zu untersuchen.

Wenn die Funktion $\Phi(t)$ reduzibel ist, so hat sie einen irreduzibeln Faktor $\varphi(t)$, in dem wir den Koeffizienten der höchsten Potenz von t gleich 1 annehmen können, und jeder solche Faktor verschwindet wenigstens für einen der Werte Θ, $\Theta_1 \dots \Theta_{n-1}$.

Es sei also

$$\varphi(\Theta) = \varphi[\chi(\alpha)] = 0,$$

d. h. die Gleichungen $\varphi[\chi(x)] = 0$ und $f(x) = 0$ haben eine gemeinsame Wurzel. Da aber $f(x)$ irreduzibel vorausgesetzt ist, so muß $\varphi[\chi(x)]$ nach I., § 53 für alle Wurzeln α, α_1, α_2, $\dots$ α_{n-1} von $f(x) = 0$ verschwinden, d. h. es ist

$$\varphi(\Theta) = 0, \quad \varphi(\Theta_1) = 0, \quad \varphi(\Theta_2) = 0 \dots \varphi(\Theta_{n-1}) = 0.$$

Wenn also die mit Θ konjugierten Werte alle voneinander verschieden sind, so ist $\varphi(t)$ mit $\Phi(t)$ identisch, d. h. $\Phi(t)$ ist irreduzibel. Sind aber unter den mit Θ konjugierten Werten nur n_1 voneinander verschiedene vorhanden, etwa Θ, Θ_1, Θ_2, $\dots$ Θ_{n_1-1}, so ist

$$\varphi(t) = (t - \Theta)\,(t - \Theta_1) \dots (t - \Theta_{n_1-1}), \tag{10}$$

und jeder andere irreduzibele Faktor von $\Phi(t)$ ist, da er wenigstens für einen der konjugierten Werte Θ, und folglich für alle verschwinden muß, mit $\varphi(t)$ identisch. Es ist also $\Phi(t)$ eine Potenz von $\varphi(t)$, etwa

(11) $$\Phi(t) = \varphi(t)^{n_2}$$

und

(12) $$n = n_1 n_2.$$

Daraus ergibt sich der Satz:

III. Die Funktion $\Phi(t)$ ist entweder irreduzibel oder sie ist eine Potenz einer irreduzibeln Funktion. Die n mit einer Zahl in $\Omega(\alpha)$ konjugierten Zahlen sind entweder alle voneinander verschieden oder sie zerfallen in n_1 Systeme von je n_2 untereinander gleichen Zahlen. Im ersten Falle ist $\Phi(t)$ irreduzibel, im zweiten die $n_2$te Potenz einer irreduzibeln Funktion $n_1$ten Grades in Ω.

Eine Größe Θ in $\Omega(\alpha)$, die von allen ihren konjugierten Zahlen verschieden ist, und die also einer irreduzibeln Gleichung $n$ten Grades genügt, heißt eine primitive Größe des Körpers. Nach dem Satze § 14 lassen sich unendlich viele solche primitive Größen bestimmen, sogar so, daß die Koeffizienten von $\Theta = \chi(\alpha)$ rationale Zahlen sind. Man braucht nur über die Koeffizienten von χ so zu verfügen, daß unter den konjugierten Größen $\chi(\alpha)$ keine gleichen vorkommen.

IV. Jede Größe ω des Körpers $\Omega(\alpha)$ kann rational durch eine beliebige primitive Größe Θ des Körpers ausgedrückt werden.

Denn sind $\omega, \omega_1, \omega_2, \ldots \omega_{n-1}$ die zu ω konjugierten Zahlen, ebenso $\Theta, \Theta_1, \Theta_2, \ldots \Theta_{n-1}$ die zu Θ konjugierten und

$$\Phi(t) = (t - \Theta)(t - \Theta_1) \ldots (t - \Theta_{n-1}),$$

so ist

$$\Phi(t)\left(\frac{\omega}{t-\Theta} + \frac{\omega_1}{t-\Theta_1} + \cdots + \frac{\omega_{n-1}}{t-\Theta_{n-1}}\right) = \Psi(t)$$

eine ganze rationale Funktion $n-1$ten Grades von t, deren Koeffizienten Zahlen in Ω sind, und daraus ergibt sich, wenn man $t = \Theta$ setzt,

$$\omega = \frac{\Psi(\Theta)}{\Phi'(\Theta)},$$

worin $\Phi'(\Theta)$ von Null verschieden ist.

Es ist hiernach der Körper $\Omega(\Theta)$ mit dem Körper $\Omega(\alpha)$ identisch.

V. Ist Θ nicht primitiv, so kann nicht jede Größe in $\Omega(\alpha)$ rational durch Θ ausgedrückt werden.

Der Körper $\Omega(\Theta)$ ist ein Teiler des Körpers $\Omega(\alpha)$ und der Grad von $\Omega(\Theta)$ ist ein Teiler des Grades von $\Omega(\alpha)$.

Denn jede Zahl des Körpers $\Omega(\Theta)$ genügt einer Gleichung in $\dot{\Omega}$ vom Grade n_1, wenn n_1 ein Teiler von n und kleiner als n ist. Also kann eine primitive Größe des Körpers $\Omega(\alpha)$, die einer irreduzibeln Gleichung $n$ten Grades genügt, nicht in $\Omega(\Theta)$ enthalten sein.

VI. Der Körper $\Omega(\alpha)$ heißt primitiv, wenn er außer den Größen in Ω keine imprimitiven Größen enthält, imprimitiv, wenn er noch andere imprimitive Größen enthält.

Aus dieser Definition ergibt sich zunächst, daß ein Körper, dessen Grad eine Primzahl ist, notwendig primitiv ist; denn eine imprimitive Größe Θ in $\Omega(\alpha)$ genügt einer Gleichung, deren Grad ein von n verschiedener Teiler des Grades n von $\Omega(\alpha)$ ist. Wenn dieser Teiler gleich 1 ist, so ist Θ in Ω enthalten.

Wir wollen jetzt noch einige der wichtigsten Eigenschaften der imprimitiven Körper kennen lernen.

Hat $\Omega(\alpha)$ einen von Ω verschiedenen Teiler Ω', der seinerseits Ω als Teiler enthält (ist also Ω' ein Körper über Ω), und ist β eine Größe, die dem Körper Ω', aber nicht Ω angehört, so ist der Körper $\Omega(\beta)$ ein algebraischer Körper über Ω und zugleich ein Teiler von Ω' und von $\Omega(\alpha)$, und nach dem, was wir vorhin bewiesen haben, ist der Grad n_1 des Körpers $\Omega(\beta)$ ein Teiler von n. Ist nun durch $\Omega(\beta)$ der Körper Ω' nicht erschöpft, so nehmen wir eine Größe γ in Ω', aber nicht in $\Omega(\beta)$; dann ist der Körper $\Omega(\beta, \gamma)$ ein Teiler von Ω' und von $\Omega(\alpha)$ und hat seinerseits $\Omega(\beta)$ zum Teiler. Der Grad von $\Omega(\beta, \gamma)$ ist also größer als der von $\Omega(\beta)$ und kleiner als der von $\Omega(\alpha)$. Ist damit Ω' noch nicht erschöpft, so fahren wir so fort, müssen aber endlich zum Abschluß kommen, da die Grade der Körper $\Omega(\beta)$, $\Omega(\beta, \gamma)$... immer wachsen und doch kleiner als n bleiben. Daraus folgt:

VII. Jeder Teiler von $\Omega(\alpha)$, der den Körper Ω enthält, ist ein algebraischer Körper $\Omega(\beta)$ über Ω. Der Grad von $\Omega(\beta)$ ist ein Teiler des Grades von $\Omega(\alpha)$, und wenn beide Körper den gleichen Grad haben, so sind sie identisch.

Wir können daher unsere Definition auch so fassen:

VIII. Ein algebraischer Körper über Ω ist primitiv, wenn er außer Ω und sich selbst keinen Körper über Ω zum Teiler hat.

§ 56.

Normalkörper. Galoissche Resolvente.

Ist $\Omega(\alpha)$ ein algebraischer Körper vom $m$ten Grade, so sind die konjugierten Körper

$$\Omega(\alpha_1),\ \Omega(\alpha_2),\ \ldots\ \Omega(\alpha_m)$$

alle von gleichem Grade, und wenn also einer von ihnen im anderen enthalten ist, so sind beide identisch.

Ein Körper, der mit allen seinen konjugierten Körpern identisch ist, heißt ein Normalkörper. In den Normalkörpern herrschen viel einfachere Gesetze, und der große Fortschritt, den die Algebra Galois verdankt, beruht im wesentlichen darauf, daß beliebige Körper auf Normalkörper zurückgeführt werden. Die Normalkörper heißen daher auch Galoissche Körper.

Wenn ein Körper μten Grades $\Omega(\varrho)$ die Eigenschaft hat, daß die zu ϱ konjugierten Zahlen $\varrho_1,\ \varrho_2\ \ldots\ \varrho_{\mu-1}$ alle in $\Omega(\varrho)$ enthalten sind, so ist $\Omega(\varrho)$ ein Normalkörper; denn dann sind die Körper $\Omega(\varrho_1),\ \Omega(\varrho_2)\ \ldots\ \Omega(\varrho_{\mu-1})$ auch alle in $\Omega(\varrho)$ enthalten und also alle mit $\Omega(\varrho)$ identisch.

Wir wollen eine Gleichung eine Normalgleichung oder auch eine Galoissche Gleichung nennen, wenn sie irreduzibel ist und die Eigenschaft hat, daß alle ihre Wurzeln rational (in Ω) durch eine von ihnen ausgedrückt werden können. Dann ist ein primitives Element eines Normalkörpers μten Grades Wurzel einer Normalgleichung μten Grades und umgekehrt erzeugt auch eine Wurzel ϱ einer Normalgleichung einen Normalkörper des gleichen Grades. Es folgt daraus noch, daß bei einer Normalgleichung jede Wurzel nicht nur durch eine bestimmte unter ihnen, sondern durch jede beliebige rational ausdrückbar ist.

Man kann nun auf folgendem Wege aus beliebigen algebraischen Körpern Normalkörper ableiten.

Es sei

$$F(x) = 0 \tag{1}$$

eine beliebige reduzible oder irreduzible Gleichung in Ω vom $m$ten Grade, von der wir nur voraussetzen wollen, daß sie keine mehrfachen Wurzeln habe. Ihre Wurzeln seien

$$\alpha, \alpha_1, \ldots \alpha_{m-1}. \tag{2}$$

Man erhält daraus m algebraische Körper

$$\Omega(\alpha), \Omega(\alpha_1), \ldots \Omega(\alpha_{m-1}), \tag{3}$$

und man kann die Funktion $F(x)$ so wählen, daß unter den Körpern (3) irgend ein gegebener algebraischer Körper über Ω vorkommt. Wir nennen den aus allen Größen (2), d. h. aus allen Wurzeln der Gleichung (1) abgeleiteten Körper über Ω

$$N = \Omega(\alpha, \alpha_1, \ldots \alpha_{m-1}) \tag{4}$$

den Galoisschen Körper der Gleichung $F(x) = 0$.

Ist $F(x)$ irreduzibel, so sind die Körper (3) die mit $\Omega(\alpha)$ konjugierten Körper. In diesem Falle soll der Körper N die Norm eines jeden der Körper $\Omega(\alpha)$ heißen.

Ist $\Omega(\alpha)$ ein Normalkörper, so ist er mit seiner Norm identisch. Im allgemeinen Falle ist nachzuweisen, daß N ein Normalkörper ist. Wir wählen ein primitives Element ϱ des Körpers N und können dann den Körper N auch durch $\Omega(\varrho)$ bezeichnen. Ist μ der Grad von N, so genügt ϱ einer irreduzibeln Gleichung μten Grades

$$g(t) = 0, \tag{5}$$

von der zu zeigen ist, daß es eine Normalgleichung ist. Zu diesem Zweck bemerken wir zunächst, daß die eine Wurzel ϱ dieser Gleichung eine rationale Funktion der $\alpha, \alpha_1, \ldots \alpha_{m-1}$ ist, weil sie in N enthalten ist. Setzen wir, um dies anzudeuten,

$$\varrho = \varrho(\alpha, \alpha_1, \ldots \alpha_{m-1}),$$

und bilden nun alle verschiedenen Anordnungen der Ziffern

$$0, 1, 2, \ldots m-1,$$

deren Anzahl $\Pi m!$ beträgt:

$$(0,1,2 \ldots m-1),\ (0',1',2',\ldots m'-1),\ (0'',1'',2'',\ldots m''-1)\ldots, \tag{6}$$

worin die Ziffern mit einem, zwei usw. Akzenten dieselben sind, wie die ohne Akzent, nur in anderer Reihenfolge, und bilden hieraus die Funktionen

$$\varrho = \varrho(\alpha, \alpha_1, \ldots \alpha_{m-1}),\quad \varrho' = \varrho(\alpha', \alpha_1', \ldots \alpha_{m-1}'), \tag{7}$$
$$\varrho'' = \varrho(\alpha'', \alpha_1'', \ldots \alpha_{m-1}'')\ldots,$$

unbekümmert darum, ob darunter etwa untereinander gleiche vorkommen oder nicht.

Wenn wir in allen den Anordnungen (6) ein und dieselbe Vertauschung vornehmen, z. B. 0 mit 1, so ändert sich die Gesamtheit dieser Anordnungen nicht, sondern nur ihre Reihenfo ge wird eine andere. Denn erstens kann durch eine solche Vertauschung nichts anderes entstehen, als Anordnungen der Ziffern, und zweitens können nicht zwei verschiedene Anordnungen durch eine und dieselbe Vertauschung in dieselbe Anordnung übergehen.

Es werden also auch durch jede solche Vertauschung die Funktionen (7) nur untereinander permutiert werden. Bilden wir also das Produkt

$$G(t) = (t - \varrho)\,(t - \varrho')\,(t - \varrho'')\,\ldots$$

für eine Veränderliche t, so bleiben seine Koeffizienten, die gewiß Funktionen von $\alpha, \alpha_1, \ldots \alpha_{m-1}$ sind, ungeändert, wenn diese Größen irgendwie permutiert werden; d. h. es sind symmetrische Funktionen der Wurzeln der Funktion $F(x)$, und also ist $G(t)$ eine Funktion von t in Ω. Alle Wurzeln von $G(t)$ sind Größen in N, da sie durch die α rational ausgedrückt sind.

Nun haben $G(t)$ und $g(t)$ eine Wurzel gemein, und da $g(t)$ irreduzibel ist, so muß $G(t)$ durch $g(t)$ teilbar sein; also sind auch alle Wurzeln von $g(t)$ in N enthalten, d. h. N ist ein Normalkörper, w. z. b. w.

Jede Gleichung $g(t) = 0$ heißt eine Galoissche Resolvente der Gleichung $F(x) = 0$, und eine Galoissche Resolvente ist also durch folgende Bestimmung definiert:

IX. Eine Gleichung $g(t) = 0$ ist eine Galoissche Resolvente einer gegebenen Gleichung $F(x) = 0$ in Ω, wenn 1. $g(t)$ irreduzibel ist, wenn 2. alle Wurzeln von $F(x)$ rational durch eine Wurzel ϱ von $g(t)$ ausdrückbar sind, und 3. eine Wurzel von $g(t)$ rational durch die Wurzeln von $F(x)$ ausdrückbar ist.

Denn nach 2. ist N in $\Omega(\varrho)$ enthalten, und nach 3. ist einer der mit $\Omega(\varrho)$ konjugierten Körper, $\Omega(\varrho_1)$, in N enthalten. Die Grade von N und $\Omega(\varrho)$ können also nicht verschieden sein und folglich ist $N = \Omega(\varrho)$.

Jede Galoissche Resolvente ist eine Normalgleichung, und eine Normalgleichung ist ihre eigene Galoissche Resolvente.

§ 57.

Die Substitutionen eines Normalkörpers.

Es sei jetzt $\Omega(\varrho)$ ein Normalkörper μten Grades und ϱ eine seiner primitiven Zahlen, ferner

(1) $$g(t) = 0$$

die irreduzible Gleichung μten Grades, deren eine Wurzel $t = \varrho$ ist. Die zu ϱ konjugierten Elemente seien

(2) $$\varrho, \varrho_1, \varrho_2 \cdots \varrho_{\mu-1}.$$

Da nach der Definition des Normalkörpers die Größen (2) alle in $\Omega(\varrho)$ enthalten sind, so können wir $\Theta_1(t)$, $\Theta_2(t)$, $\cdots$ $\Theta_{\mu-1}(t)$ als ganze rationale Funktionen in Ω, höchstens vom Grade $\mu - 1$, so bestimmen daß

(3) $$\varrho_1 = \Theta_1(\varrho),\ \varrho_2 = \Theta_2(\varrho),\ \cdots\ \varrho_{\mu-1} = \Theta_{\mu-1}(\varrho),$$

wird. Ist ω eine beliebige Größe in $\Omega(\varrho)$, so kann man die mit ω konjugierten Größen so darstellen:

(4) $$\omega = \varphi(\varrho),\ \omega_1 = \varphi(\varrho_1),\ \cdots\ \omega_{\mu-1} = \varphi(\varrho_{\mu-1}),$$

worin $\varphi(t)$ eine rationale Funktion in Ω ist.

Da nun $g(t)$ irreduzibel ist, so gilt der Satz § 53, I., den wir jetzt so aussprechen:

1. Wenn eine rationale Funktion $\Phi(t)$ in Ω eine Wurzel mit $g(t)$ gemeinsam hat, so verschwinden alle konjugierten Größen
$$\Phi(\varrho),\ \Phi(\varrho_1),\ \cdots\ \Phi(\varrho_{\mu-1}).$$

Wenn in einer der Funktionen $\Theta_k(\varrho)$, durch die nach (3) die Wurzeln von (1) ausgedrückt sind, ϱ durch eine andere Wurzel ϱ_h ersetzt wird, so entsteht daraus wieder eine der Wurzeln; denn ist

$$g[\Theta_k(\varrho)] = 0,$$

so ist nach 1. auch

$$g[\Theta_k(\varrho_h)] = 0,$$

und die Reihe der Größen

(5) $$\varrho_h,\ \Theta_1(\varrho_h),\ \Theta_2(\varrho_h),\ \cdots\ \Theta_{\mu-1}(\varrho_h)$$

stimmt, von der Anordnung abgesehen, mit der Reihe

(6) $$\varrho,\ \Theta_1(\varrho),\ \Theta_2(\varrho),\ \cdots\ \Theta_{\mu-1}(\varrho)$$

überein. Dies wird erwiesen sein, wenn wir zeigen, daß in (5) keine zwei gleichen Werte vorkommen. Bezeichnen wir der Über-

einstimmung halber mit $\Theta_0(\varrho)$ oder $\Theta(\varrho)$ die Wurzel ϱ selbst, so folgt aus der Gleichheit zweier der Größen (5)

$$\Theta_i(\varrho_h) = \Theta_k(\varrho_h) \tag{7}$$

nach dem Satze 1.

$$\Theta_i(\varrho) = \Theta_k(\varrho), \tag{8}$$

was aber, wenn i von k verschieden ist, nicht möglich ist. Also sind zugleich mit den Größen (6) auch die Großen (5) untereinander verschieden.

Wir können dies als Satz zusammenfassen:

2. Vertauscht man ϱ mit ϱ_h, so geht zugleich jede mit ϱ konjugierte Größe in eine bestimmte andere über, und niemals zwei verschiedene in dieselbe.

Wenn wir in allen Funktionen von ϱ statt ϱ eine andere Wurzel ϱ_a setzen, so führen wir eine Substitution aus. Wir bezeichnen diese Substitution mit

$$\sigma_a = (\varrho, \varrho_a), \qquad a = 0, 1, 2, \ldots \mu - 1,$$

wobei unter σ_0 oder σ die identische Substitution (ϱ, ϱ) verstanden wird, die darin besteht, daß ϱ durch sich selbst ersetzt wird, also alle Zahlen in $\Omega(\varrho)$ ungeändert bleiben.

Wenn wir in einer beliebigen der Wurzeln (3)

$$\varrho_h = \Theta_h(\varrho)$$

die Substitution σ_a ausführen, so geht ϱ_h in eine andere Wurzel ϱ_k über, die bestimmt ist durch

$$\varrho_k = \Theta_h(\varrho_a) = \Theta_h\Theta_a(\varrho) = \Theta_k(\varrho). \tag{9}$$

Es ist also ϱ_k dieselbe Funktion von ϱ_a, wie ϱ_h von ϱ.

Eine beliebige Größe $\omega = \varphi(\varrho)$ des Körpers $\Omega(\varrho)$ geht durch die Substitution σ_a in $\omega_a = \varphi(\varrho_a)$ über. Drücken wir ω durch ϱ_h aus, so ergibt sich eine rationale Funktion ψ, die der Bedingung genügt:

$$\omega = \psi(\varrho_h), \quad \omega_a = \psi(\varrho_k).$$

Es geht also ω in ω_a über durch die Substitution

$$\sigma_a = (\varrho_h, \varrho_k), \tag{10}$$

und dies ist nur ein anderer Ausdruck für die Substitution (ϱ, ϱ_a).

Hierin ist ϱ_h eine beliebige Wurzel von $g(t)$ und das zugehörige ϱ_k ist nach (9) durch σ_a bestimmt. Man kann aber bei gegebenem σ_a auch ϱ_k beliebig annehmen und das zugehörige ϱ_h bestimmen, indem man in $\Theta_h(\varrho_a)$ den Index h die Werte

0, 1, 2, ... $\mu - 1$ durchlaufen läßt, wobei jede Wurzel ϱ_k einmal zum Vorschein kommt. Also:

3. Jede Substitution σ_a kann in der Form (ϱ_h, ϱ_k) dargestellt werden, worin entweder ϱ_h oder ϱ_k eine beliebig gegebene der μ Wurzeln ϱ ist, während die andere durch σ_a völlig bestimmt ist.

Die Anzahl aller voneinander verschiedenen Substitutionen σ_a ist also, die identische Substitution mitgerechnet, gleich dem Grade μ des Körpers $\Omega(\varrho)$. Jede dieser Substitutionen führt die Gesamtheit der Größen des Körpers $\Omega(\varrho)$ in sich selbst über, so daß jede in eine bestimmte andere übergeht, und niemals zwei verschiedene Größen in die gleiche.

Wir nennen daher die σ_a die Substitutionen des Körpers $\Omega(\varrho)$.

Bleibt $\omega = \varphi(\varrho)$ im Körper $\Omega(\varrho)$ ungeändert, wenn ϱ durch ϱ_a ersetzt wird, wenn also

$$\varphi(\varrho) = \varphi(\varrho_a)$$

ist, so sagen wir, ω erlaubt oder gestattet die Substitution (ϱ, ϱ_a) oder σ_a.

Die Größen ω, die außer der identischen Substitution keine Substitution σ_a gestatten, sind die primitiven Elemente des Körpers $\Omega(\varrho)$.

4. Eine Größe, die alle Substitutionen σ_a gestattet, ist notwendig ein Element von Ω.

Denn eine solche Größe ist mit allen ihren konjugierten Größen identisch, und genügt also nach § 55 einer Gleichung ersten Grades in Ω.

§ 58.
Die Galoissche Gruppe.

Wenn wir in irgend einer Funktion von ϱ die Wurzel ϱ zuerst durch ϱ_a und dann ϱ_a durch ϱ_b ersetzen, so ist der Erfolg derselbe, als ob wir gleich von vornherein ϱ mit ϱ_b vertauscht hätten. Setzen wir also

$$(1) \qquad \sigma = (\varrho, \varrho_a), \quad \sigma' = (\varrho_a, \varrho_b), \quad \sigma'' = (\varrho, \varrho_b),$$

so ist es für das Ergebnis gleichgültig, ob wir in den Zahlen des Körpers $\Omega(\varrho)$ zuerst die Substitution σ und dann die Substitution σ' ausführen, oder ob wir für einmal die Substitution σ'' machen.

1. Wir nennen daher σ'' aus σ und σ' komponiert oder zusammengesetzt, und bezeichnen diese Beziehung durch die symbolische Gleichung

$$\sigma\sigma' = \sigma''. \tag{2}$$

Da wir nach dem Satze 3. des vorigen Paragraphen in der Bezeichnung einer Substitution (ϱ_h, ϱ_k) das erste oder das zweite Element beliebig wählen können, so lassen sich ϱ_a, ϱ_b so auswählen, daß σ, σ' zwei beliebig gegebene unter den μ Substitutionen des Körpers $\Omega(\varrho)$ sind; σ'' ist aber dadurch völlig bestimmt. Ebenso ist aber auch, wenn σ und σ'' gegeben sind σ' eindeutig bestimmt. Denn wählen wir ϱ beliebig, so ist ϱ_a durch σ völlig bestimmt und ϱ_b durch σ'', womit auch $\sigma' = (\varrho_a, \varrho_b)$ gegeben ist. Ist endlich σ' und σ'' gegeben, so ist σ eindeutig bestimmt, da zunächst ϱ_b durch σ'' und dann ϱ_a durch σ' bestimmt ist. Wir haben daher:

2. Von den durch die symbolische Gleichung $\sigma\sigma' = \sigma''$ verbundenen drei Substitutionen des Körpers $\Omega(\varrho)$ können irgend zwei beliebig gegeben sein, während die dritte dadurch eindeutig bestimmt ist.

Es ist bei dieser Komposition aber wohl auf die Reihenfolge der Komponenten zu achten, weil $\sigma\sigma'$ von $\sigma'\sigma$ verschieden sein kann, d. h. bei der die Komposition ausdrückenden symbolischen Multiplikation gilt nicht das kommutative Gesetz der gewöhnlichen Multiplikation; wohl aber gilt das assoziative Gesetz, das sich in folgendem Satz ausspricht:

3. Sind σ, σ', σ'' irgend drei der μ Substitutionen des Körpers $\Omega(\varrho)$, so ist

$$(\sigma\sigma')\sigma'' = \sigma(\sigma'\sigma''). \tag{3}$$

Der Beweis ist sehr einfach; denn nach dem Satze 3., § 57, können wir ϱ_a, ϱ_b, ϱ_c so bestimmen, daß

$$\sigma = (\varrho, \varrho_a), \quad \sigma' = (\varrho_a, \varrho_b), \quad \sigma'' = (\varrho_b, \varrho_c), \tag{4}$$

und dann ist

$$\sigma\sigma' = (\varrho, \varrho_b), \qquad (\sigma\sigma')\sigma'' = (\varrho, \varrho_b)(\varrho_b, \varrho_c) = (\varrho, \varrho_c),$$
$$\sigma'\sigma'' = (\varrho_a, \varrho_c), \qquad \sigma(\sigma'\sigma'') = (\varrho, \varrho_a)(\varrho_a, \varrho_c) = (\varrho, \varrho_c).$$

Wir bezeichnen daher kurz die aus den drei Substitutionen σ, σ', σ'' zusammengesetzte Substitution mit

$$\sigma\sigma'\sigma'',$$

und können überhaupt aus beliebig vielen Komponenten eine bestimmte Substitution

$$\sigma\sigma'\sigma''\sigma''' \ldots$$

dadurch zusammensetzen, daß wir nacheinander immer zwei benachbarte unter den Komponenten zu einer Substitution vereinigen, bis das ganze symbolische Produkt sich auf eine einzige Substitution zusammengezogen hat. Denn wir können nach § 57, 3. setzen:

$$\sigma = (\varrho, \varrho_a), \quad \sigma' = (\varrho_a, \varrho_b), \quad \sigma'' = (\varrho_b, \varrho_c), \quad \sigma''' = (\varrho_c, \varrho_e),$$

und erhalten immer

$$(\varrho, \varrho_a)(\varrho_a, \varrho_b)(\varrho_b, \varrho_c)(\varrho_c, \varrho_e) = (\varrho, \varrho_e),$$

was offenbar auf eine beliebige Anzahl von Komponenten ausgedehnt werden kann.

Die Substitutionen eines Normalkörpers haben also die charakteristischen Kennzeichen einer Gruppe. Sie heißt die Gruppe der Substitutionen des Körpers $\Omega(\varrho)$ und mag mit G_ϱ bezeichnet werden.

Wir haben schon im § 48 jede Gruppe mit einer Permutationsgruppe in Beziehung gesetzt. Diese Permutationsgruppen erhalten nun für die Algebra eine besondere Bedeutung. Wir waren nämlich ausgegangen von einer beliebigen Gleichung $m$ten Grades ohne mehrfache Wurzeln, $F(x) = 0$, und erhielten daraus einen Normalkörper $\Omega(\varrho)$, wenn wir eine Funktion ϱ der Wurzeln nahmen, die durch jede Vertauschung einen anderen Wert erhält. Durch diese Funktion ϱ können alle Wurzeln von $F(x)$ rational in Ω ausgedrückt werden:

$$\alpha_1 = \chi_1(\varrho), \quad \alpha_2 = \chi_2(\varrho), \ldots \alpha_m = \chi_m(\varrho).$$

Diese α sind also Zahlen in $\Omega(\varrho)$ und sie erfahren eine Permutation durch jede Substitution (ϱ, ϱ_a). Führt man alle Substitutionen der Gruppe G_ϱ aus, so erhält man eine Permutationsgruppe P, die mit G_ϱ isomorph ist. Wir nennen sie die Galoissche Gruppe der Gleichung $F = 0$. Sie hat ebenso wie G_ϱ den Grad der Galoisschen Resolvente $g(\varrho) = 0$. Wenn man statt von ϱ von einer anderen Funktion derselben Art ausgegangen wäre, so hätte man zwar eine andere Resolvente, aber keinen anderen Körper $\Omega(\varrho)$ und keine andere Permutationsgruppe P bekommen.

Von einer Funktion der Wurzeln α, die ihren Wert nicht ändert, wenn die Wurzeln der Permutation π unterworfen werden, sagen wir (wie bei den Substitutionen), sie gestatte die Permutation π.

Dann kann die Galoissche Gruppe P durch folgende Eigenschaften charakterisiert werden:

a) Jede rationale Gleichung in Ω, die zwischen den m Wurzeln von $F(x)$ besteht, bleibt richtig, wenn die Wurzeln irgend einer Permutation der Galoisschen Gruppe unterworfen werden.

b) Jede rationale Funktion in Ω von den m Wurzeln von $F(x)$, die sämtliche Permutationen der Galoisschen Gruppe gestattet, ist eine Zahl in Ω.

Denn drücken wir die Wurzeln α von $F(x)$ als rationale Funktionen von ϱ aus, so geht eine Funktion der Wurzeln α in eine Funktion $\varphi(\varrho)$ über. Hat man zuvor eine Permutation der Galoisschen Gruppe ausgeführt, so erhält man eine der konjugierten Zahlen $\varphi(\varrho_a)$, die aus $\varphi(\varrho)$ durch eine Substitution $\sigma_a = (\varrho, \varrho_a)$ entsteht. Ist nun $\varphi(\varrho) = 0$, so sind nach dem Satz § 57, 1. auch alle konjugierten $\varphi(\varrho_a) = 0$, womit a) bewiesen ist; und sind die konjugierten Größen $\varphi(\varrho_a)$ alle einander gleich, so ist ihr gemeinsamer Wert nach § 57, 4. in Ω enthalten, wodurch b) bewiesen ist.

Zu a), b) kommt noch als drittes:

c) Wenn irgend eine Permutation π der Wurzeln von $F(x)$ auf alle rationalen Gleichungen in Ω, die zwischen den Wurzeln bestehen, anwendbar ist, so gehört π der Galoisschen Gruppe an, und die Galoissche Gruppe kann daher auch erklärt werden als der Inbegriff aller Permutationen, die auf sämtliche rationale Gleichungen zwischen den Wurzeln anwendbar sind.

Denn nach § 14 kann man ϱ als rationale (z. B. lineare) Funktion der m Größen $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ so annehmen, daß alle $m!$ Werte, die man durch die $m!$ überhaupt möglichen Permutationen der α daraus erhält, voneinander verschieden sind.

Ist dann $g(t) = 0$ die Galoissche Resolvente von $F(x) = 0$, deren Wurzel dieses ϱ ist, so ist

$$g(\varrho) = 0.$$

Hierauf können wir, wenn ϱ durch die Wurzeln α ausgedrückt ist, nach Voraussetzung die Permutation π anwenden und erhalten also, wenn dadurch ϱ in ϱ_a übergeht, $g(\varrho_a) = 0$; d. h. ϱ_a ist auch eine Wurzel der Resolvente, und die Permutation π entspricht also einer Substitution (ϱ, ϱ_a) des Körpers $\Omega(\varrho)$, d. h. π gehört zur Galoisschen Gruppe.

Daraus schließen wir noch auf folgenden Satz:

d) Ist P eine Gruppe von Permutationen der m Wurzeln α, der die Eigenschaften a) und b) zukommen, so ist P die Galoissche Gruppe der Gleichung $F(x) = 0$.

Denn zunächst gehört nach c) jede Permutation von P der Galoisschen Gruppe an, und P ist also gewiß ein Teiler von dieser.

Wenn aber P nur ν Permutationen umfaßt, so mögen diese mit

$$\pi_1, \pi_2, \ldots \pi_\nu \tag{5}$$

bezeichnet sein. Wenden wir diese Permutationen auf ϱ an, so möge sich ergeben:

$$\varrho_1, \varrho_2, \ldots \varrho_\nu. \tag{6}$$

Wenden wir eine der Permutationen (5), etwa π_k, auf eine der Größen (6), etwa auf ϱ_i an, so ist der Erfolg derselbe, als ob $\pi_i \pi_k$ auf ϱ angewandt worden sei; das Ergebnis dieser Permutation soll ϱ_i' sein. Nun liegt aber in der Voraussetzung, daß P eine Gruppe sei, daß auch die komponierte Permutation $\pi_i \pi_k$ zu P gehört, daß also ϱ_i' unter den Größen (6) enthalten sei. Zugleich ist ϱ_i' nach § 57, 2. von ϱ_h' verschieden, sobald ϱ_i von ϱ_h verschieden ist. Daraus folgt, daß die Größen

$$\varrho_1', \varrho_2', \ldots \varrho_\nu' \tag{7}$$

mit den Größen (6), von der Ordnung abgesehen, übereinstimmen.

Das Produkt

$$g'(t) = (t - \varrho_1)(t - \varrho_2) \ldots (t - \varrho_\nu)$$

bleibt also durch die Permutationen der Gruppe P ungeändert, und ist folglich, da wir die Eigenschaft b) von P voraussetzen eine rationale Funktion von t in Ω. Zugleich ist $g'(t)$ ein Teiler von $g(t)$ und muß daher, da $g(t)$ irreduzibel ist, mit $g(t)$ über-

einstimmen, also ist ν nicht kleiner als der Grad der Galoisschen Gruppe und P ist mit der Galoisschen Gruppe identisch

Wählen wir, wie oben, die Größe ϱ als rationale Funktion der α so, daß die $m!$ durch die Permutationen der α sich ergebenden Größen

$$\varrho,\ \varrho',\ \varrho'' \ldots$$

alle voneinander verschieden sind, so ist

$$G(t) = (t - \varrho)\ (t - \varrho')\ (t - \varrho'') \ldots$$

eine Funktion in Ω. Nun ist jede von den Größen ϱ, ϱ', $\varrho'' \ldots$ eine primitive Größe des Normalkörpers $N = \Omega(\alpha_1, \alpha_2, \ldots \alpha_m)$ und jede von ihnen ist also die Wurzel einer Galoisschen Resolvente μten Grades. Je μ von diesen Größen sind die Wurzeln von einer solchen Gleichung. Es muß also $G(t)$, das keine gleichen Wurzeln hat, in lauter irreduzible Faktoren μten Grades zerfallen, und daraus ergibt sich noch, daß μ ein Teiler von $m!$ ist. Zugleich ist μ der Grad der Galoisschen Gruppe von $F(x)$.

Hat der Grad der Galoisschen Resolvente einer Gleichung $m$ten Grades den größten Wert $m!$, so sagen wir, mit einem von Kronecker herrührenden Ausdruck, die Gleichung hat keinen Affekt. Sie hat einen um so höheren Affekt, je niedriger der Grad μ der Galoisschen Resolvente ist. Den Quotienten $m!:\mu$, der immer eine ganze Zahl, höchstens gleich $m!$ und mindestens gleich 1 sein muß, wollen wir den Grad des Affektes nennen, der also bei einer Gleichung ohne Affekt den Wert 1 hat. Wenn der Affekt den möglichst hohen Grad $m!$ hat, dann sind die Wurzeln der Gleichung selbst im Rationalitätsbereich Ω enthalten, die Gleichung ist also gelöst.

Wenn durch Adjunktion einer algebraischen Größe zu Ω die Galoissche Resolvente reduzibel wird, so entsteht eine neue Resolvente niedrigeren Grades, und der Grad des Affektes der Gleichung $F(x) = 0$ vergrößert sich. Man nähert sich also dadurch der Lösung der Gleichung.

Die Galoissche Auffassung der Aufgabe, eine Gleichung $F(x) = 0$ zu lösen, besteht darin, daß durch aufeinander folgende Adjunktion von algebraischen Größen möglichst einfacher Art die Gruppe allmählich verkleinert, oder der Affekt erhöht werden soll, bis er seinen höchsten Grad erreicht hat.

4. Die allgemeine Gleichung $m$ten Grades hat in dem Körper Ω, der aus den rationalen Funktionen der Koeffizienten $a_1, a_2, \ldots a_m$ als unabhängige Variable besteht, keinen Affekt.

Man kann nämlich in diesem Körper Ω die Galoissche Resolvente von $F(x) = 0$ in folgender Weise bilden.

Es bezeichnen $\alpha_1, \alpha_2, \ldots \alpha_m$ unabhängige Variable und

$$\varrho = u_1\alpha_1 + u_2\alpha_2 + \cdots + u_m\alpha_m$$

eine lineare Funktion, deren Koeffizient $u_1, u_2, \ldots u_m$ beliebige rationale Zahlen, unter denen keine zwei einander gleich sind. Führt man in dieser Funktion ϱ alle Permutationen der $\alpha_1, \alpha_2, \ldots \alpha_m$ aus, so erhält man $m!$ verschiedene Funktionen $\varrho, \varrho', \varrho'' \ldots$ und die Funktion

$$G(t) = (t - \varrho)\,(t - \varrho')\,(t - \varrho'')\ldots$$

ist eine symmetrische Funktion der $\alpha_1, \alpha_2, \ldots \alpha_m$. Sie kann also rational durch die symmetrischen Grundfunktionen ausgedrückt und diese dann durch unabhängige Variable $a_1, a_2, \ldots a_m$ ersetzt werden (§ 25, IV). Es entsteht so eine rationale Funktion $G(t, a)$ von t und a, die in bezug auf t vom Grade $m!$ ist. Diese Funktion $G(t, a)$ ist im Körper Ω, der aus den rationalen Funktionen der a besteht, irreduzibel. Denn zerlegt man $G(t, a)$ in irreduzible Faktoren

$$G(t, a) = g(t, a)\; g_1(t, a)\ldots,$$

so kann man die a auf beiden Seiten wieder umgekehrt durch die symmetrischen Grundfunktionen der α ersetzen und erhält dadurch wieder die Funktion $G(t)$, die als Funktion der α mit $G(t, \alpha)$ bezeichnet sein soll. Es muß also einer der Faktoren, etwa $g(\varrho, a)$ verschwinden, und darin kann man sämtliche Permutationen der α ausführen, wodurch die a nicht geändert werden. Es ist also zugleich $g(\varrho', a) = 0$, $g(\varrho'', a) = 0 \ldots$, d. h. $g(t, a)$ ist irreduzibel und folglich mit $G(t, a)$ identisch.

Wenn Ω der absolute Rationalitätsbereich ist, so sind die Koeffizienten a von $F(x)$ rationale Zahlen und der Affekt ist eine zahlentheoretische Eigenschaft dieser Zahlen. Die α sind in diesem Falle algebraische Zahlen.

Eine Funktion dieser α, die durch die Permutationen einer Gruppe ungeändert bleibt, braucht nicht formal ungeändert zu bleiben, sondern nur mit Benutzung der Gleichung $F = 0$. Man kann aber einer solchen Funktion eine Form geben, die auch

formal ungeändert bleibt, indem man das arithmetische Mittel aller der Ausdrücke nimmt, in die sie durch Anwendung der fraglichen Gruppe übergeht.

§ 59.

Reduzible und imprimitive Gleichungen.

Die erste Anwendung, die wir von der Galoisschen Gruppe machen, gibt uns ein Kriterium für die Reduzibilität und Irreduzibilität.

Wenn die Funktion $F(x)$ reduzibel ist, so zerfällt sie in zwei Faktoren in Ω:

$$F(x) = F_1(x) F_2(x),$$

und $F_1(x)$, $F_2(x)$ haben keine gemeinschaftliche Wurzel, weil wir angenommen haben, daß $F(x)$ keine mehrfache Wurzel habe. Ist also α eine Wurzel von $F_1(x)$ und β eine Wurzel von $F_2(x)$, so ist $F_1(\alpha) = 0$ und $F_1(\beta)$ von Null verschieden. Eine Permutation, die α in β überführt, läßt also die rationale Gleichung $F_1(\alpha) = 0$ nicht bestehen und kann nach a) nicht zur Galoisschen Gruppe von $F(x)$ gehören.

Nach § 50 heißt eine Permutationsgruppe beliebiger Symbole intransitiv, wenn durch Permutationen dieser Gruppe nicht jedes dieser Symbole in jedes andere übergeführt werden kann. Bedeuten also diese Symbole die Wurzeln von $F(x)$ und die Gruppe die Galoissche, so gilt der Satz:

1. Die Galoissche Gruppe einer reduzibeln Gleichung ist intransitiv.

Es gilt auch das Umgekehrte. Denn sind A, B, C ... die Systeme der Intransitivität, und sind $\alpha_1, \alpha_2, \ldots \alpha_\mu$ die Wurzeln des Systems A, $\beta_1, \beta_2, \ldots \beta_\nu$ die des Systems B, $\gamma_1, \gamma_2, \ldots \gamma_\sigma$ die des Systems C, so gestattet jede der Funktionen

$$\begin{aligned} A(x) &= (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_\mu) \\ B(x) &= (x - \beta_1)(x - \beta_2) \ldots (x - \beta_\nu) \\ C(x) &= (x - \gamma_1)(x - \gamma_2) \ldots (x - \gamma_\sigma) \end{aligned}$$

(für ein unbestimmtes x) die Permutationen der Gruppe und ist also in Ω enthalten. Es ist folglich

$$F(x) = A(x) B(x) C(x) \ldots$$

zerlegbar. Wir haben also:

2. Ist die Galoissche Gruppe einer Gleichung $F(x)$ intransitiv, so entspricht jedem System der Intransitivität ein rationaler Faktor von $F(x)$.

Es sei jetzt $f(x) = 0$ eine irreduzible Gleichung $n$ten Grades, also ihre Galoissche Gruppe P transitiv. Die Wurzeln von $f(x)$ seien

$$\alpha,\ \alpha_1,\ \alpha_2,\ \dots\ \alpha_{n-1}. \tag{1}$$

Wenn der Körper $\Omega(\alpha)$ imprimitiv ist (§ 55), so gibt es ein Element $\Theta = \chi(\alpha)$, unter dessen konjugierten Werten mehrere, sagen wir r untereinander gleiche vorkommen. Es ist also

$$\Theta = \chi(\alpha) = \chi(\alpha_1) = \cdots = \chi(\alpha_{r-1}),$$

während alle übrigen unter den konjugierten Werten Θ von $\chi(\alpha)$ verschieden sind.

Wir erhalten also eine erste Reihe von Wurzeln von $f(x)$:

$$A = \alpha,\ \alpha_1\ \dots\ \alpha_{r-1}.$$

Ist β eine nicht in A enthaltene Wurzel, so gibt es in P eine Permutation π, durch die α in β übergeht, und diese Permutation läßt sich auf die Gleichung $\chi(\alpha) = \chi(\alpha_i)$ anwenden. Geht α_i durch π in β_i über, so sind die r Größen

$$B = \beta,\ \beta_1\ \dots\ \beta_{r-1}$$

alle voneinander verschieden, und keine davon ist in A enthalten. Denn wäre etwa $\beta_1 = \alpha_1$, so würde aus $\chi(\alpha) = \chi(\alpha_1)$ durch π folgen $\chi(\beta) = \chi(\beta_1) = \chi(\alpha_1)$, also gegen die Voraussetzung $\chi(\alpha_1) = \chi(\beta)$. Daraus schließt man, indem man so fortfährt, daß r ein Teiler von n ist:

$$n = rs,$$

und daß sich die Werte (1) in s Reihen von je r Wurzeln zerlegen lassen, die wir mit

$$\begin{aligned} A &= \alpha,\ \alpha_1\ \dots\ \alpha_{r-1} \\ B &= \beta,\ \beta_1\ \dots\ \beta_{r-1} \\ &\dots\dots\dots\dots \\ S &= \sigma,\ \sigma_1\ \dots\ \sigma_{r-1} \end{aligned} \tag{2}$$

bezeichnen wollen, so daß

$$\begin{aligned} \Theta &= \chi(\alpha) = \chi(\alpha_1) \cdots = \chi(\alpha_{r-1}) \\ \Theta_1 &= \chi(\beta) = \chi(\beta_1) \cdots = \chi(\beta_{r-1}) \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ \Theta_{s-1} &= \chi(\sigma) = \chi(\sigma_1) \cdots = \chi(\sigma_{r-1}) \end{aligned} \tag{3}$$

die konjugierten Werte von Θ sind. Da diese Größen durch die (transitive) Gruppe P nur unter einander permutiert werden, so ist

$$(t - \Theta)(t - \Theta_1) \dots (t - \Theta_{s-1}) = \varphi(t)$$

eine Funktion in Ω vom Grade s in bezug auf t, deren Wurzeln die Werte (3) sind. Da diese Größen durch P transitiv verbunden sind, so ist $\varphi(t)$ irreduzibel.

Es ergibt sich nun aus (3), daß die Gruppe P so beschaffen sein muß, daß alle ihre Permutationen die Elemente der einzelnen Reihen $A, B, \dots S$ nur untereinander vertauschen und außerdem die ganzen Reihen $A, B, \dots S$ miteinander vertauschen, so daß niemals an Stelle von zwei Elementen derselben Reihe zwei Elemente verschiedener Reihen treten. Denn wenn etwa durch eine Permutation π von P, α und α_1 in β und σ übergeführt würden, so würde folgen, da man die Permutation π [nach a), § 58] in der Gleichung $\chi(\alpha) = \chi(\alpha_1)$ ausführen darf, daß auch $\chi(\beta) = \chi(\sigma)$ sein müßte, was der Annahme widerspricht, daß die Werte (3) voneinander verschieden sind. Die Gruppe P ist also nach § 50 imprimitiv und die $A, B, C, \dots$ sind die Systeme der Imprimitivität. Danach haben wir den Satz.

3. Ein imprimitiver Körper hat eine imprimitive Gruppe.

Es ergibt sich aus der Imprimitivität der Gruppe für die imprimitiven Körper ein wichtiges Resultat.

Wir wollen mit

$$\psi(x, x_1, \dots x_{r-1})$$

eine rationale symmetrische Funktion der r Veränderlichen $x, x_1 \dots x_{r-1}$ bezeichnen und setzen:

$$(4) \qquad \begin{aligned} \omega &= \psi(\alpha, \alpha_1 \dots \alpha_{r-1}) \\ \omega_1 &= \psi(\beta, \beta_1 \dots \beta_{r-1}) \\ &\dots\dots\dots\dots\dots\dots\dots \\ \omega_{s-1} &= \psi(\sigma, \sigma_1 \dots \sigma_{r-1}); \end{aligned}$$

dann ist zu beweisen, daß ω rational durch Θ ausgedrückt werden kann, d. h. im Körper $\Omega(\Theta)$ enthalten ist. Es ist nämlich

$$(5) \qquad \varphi(t)\left(\frac{\omega}{t - \Theta} + \frac{\omega_1}{t - \Theta_1} + \dots + \frac{\omega_{s-1}}{t - \Theta_{s-1}}\right) = \Phi(t)$$

eine ganze rationale Funktion von t, deren Koeffizienten rationale Funktionen der $\alpha, \alpha_1, \dots \alpha_{r-1}$ sind, die ungeändert bleiben, wenn

irgend eine Permutation der Gruppe P vorgenommen wird, weil durch diese Permutationen die ω entweder ungeändert bleiben oder in derselben Weise wie die Θ miteinander permutiert werden. Nach § 58, b) sind diese Koeffizienten also in Ω enthalten. Setzen wir dann in (5) für das unbestimmte t den Wert Θ ein, so folgt, da $\varphi'(\Theta)$ von Null verschieden ist,

$$\omega = \frac{\Phi(\Theta)}{\varphi'(\Theta)}. \tag{6}$$

Wenden wir dies auf die Koeffizienten des Produktes

$$(u - \alpha)(u - \alpha_1) \dots (u - \alpha_{r-1}) = \varphi(u, \Theta) \tag{7}$$

an, wo u eine Variable bedeutet, so ergibt sich, daß diese Funktion $r$ten Grades, deren Wurzeln die $\alpha, \alpha_1, \dots \alpha_{r-1}$ sind, rational durch Θ ausgedrückt werden kann.

Man kann auch in P immer eine Permutation finden, durch die Θ ungeändert bleibt und α in irgend ein beliebiges der $\alpha_1, \alpha_2, \dots \alpha_{r-1}$ übergeht, und daraus folgt, daß die Funktion $\varphi(u, \Theta)$ in $\Omega(\Theta)$ irreduzibel ist.

Es ist also die Größe α, die ursprünglich Wurzel einer Gleichung $n$ten Grades in Ω ist, zugleich Wurzel einer Gleichung $r$ten Grades, deren Koeffizienten rational von der Wurzel einer irreduzibeln Gleichung $s$ten Grades abhängen. Eine Gleichung $f(x) = 0$, deren Wurzel α diese Eigenschaft hat, nennt man eine imprimitive Gleichung.

Wir können das Bewiesene auch so ausdrücken:

4. Der imprimitive Körper $\Omega(\alpha)$ vom $n$ten Grade wird durch die Adjunktion des Körpers $s$ten Grades $\Omega' = \Omega(\Theta)$ zu einem Körper $\Omega'(\alpha)$ vom $r$ten Grade über $\Omega(\Theta)$.

Wir wollen noch untersuchen, ob die Imprimitivität der Gruppe ein ausreichendes Kennzeichen für die Imprimitivität des Körpers ist, ob man also in einem Körper $\Omega(\alpha)$ mit imprimitiver Gruppe immer imprimitive Elemente finden kann.

Sei also jetzt $f(x) = 0$ eine irreduzible Gleichung $n$ten Grades mit imprimitiver Gruppe, und seien (1) die Wurzeln dieser Gleichung und (2) die Systeme der Imprimitivität. Es ist nun zu zeigen, daß eine Funktion $\chi(\alpha)$ existiert, der die durch (3) ausgedrückte Eigenschaft zukommt.

Wir wählen irgend eine symmetrische Funktion $\psi(x, x_1, \ldots x_{r-1})$ so, daß die Werte

$$(8)\qquad \begin{aligned} y &= \psi(\alpha, \alpha_1 \ldots \alpha_{r-1}) \\ y_1 &= \psi(\beta, \beta_1 \ldots \beta_{r-1}) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{s-1} &= \psi(\sigma, \sigma_1 \ldots \sigma_{r-1}) \end{aligned}$$

alle voneinander verschieden sind.

Um die Möglichkeit hiervon einzusehen, können wir z. B.

$$(9)\qquad \begin{aligned} \psi(t, A) &= (t-\alpha)(t-\alpha_1)\ldots(t-\alpha_{r-1}) \\ \psi(t, B) &= (t-\beta)(t-\beta_1)\ldots(t-\beta_{r-1}) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \psi(t, S) &= (t-\sigma)(t-\sigma_1)\ldots(t-\sigma_{r-1}) \end{aligned}$$

setzen, und dann läßt sich für t ein solcher rationaler Wert finden, daß diese Größen alle voneinander verschieden ausfallen. Diese Werte können also für die y in (8) genommen werden.

Wenn wir nun unter u eine unabhängige Variable verstehen und

$$(10)\qquad \varphi(u) = (u-y)(u-y_1)\ldots(u-y_{s-1})$$

setzen, so ist $\varphi(u)$, da seine Koeffizienten durch die Permutationen der Gruppe ungeändert bleiben, eine Funktion in Ω, deren Wurzeln die Größen y sind. Sie ist überdies irreduzibel, denn aus der vorausgesetzten Transitivität der Gruppe folgt, daß, wenn eine rationale Funktion von y für einen der Werte (8) verschwindet, sie auch für alle anderen verschwinden muß.

Ist ω irgend eine symmetrische Funktion von $\alpha, \alpha_1, \ldots \alpha_{r-1}$, so schließen wir aus der Betrachtung des Ausdruckes

$$\varphi(u)\left(\frac{\omega}{u-y} + \frac{\omega_1}{u-y_1} + \cdots + \frac{\omega_{s-1}}{u-y_{s-1}}\right),$$

der eine ganze rationale Funktion von u in Ω ist, ganz wie oben, indem wir $u = y$ setzen, daß ω rational durch y dargestellt werden kann, und demnach kann auch die Funktion

$$(11)\qquad \psi(t, A) = \psi(t, y),$$

deren Wurzeln die Größen $\alpha, \alpha_1, \ldots \alpha_{r-1}$ sind, rational durch y ausgedrückt werden.

Auch die Gleichung $\psi(t, A = 0$ ist irreduzibel in dem Körper $\Omega(y)$; denn ist $\psi_1(t, y)$ ein rationaler Teiler von $\psi(t, y)$, und ist

$$\psi_1(\alpha, y) = 0,$$

so können wir wegen der Transitivität der Gruppe auf diese Gleichung eine Permutation anwenden, durch die α in eine beliebige Größe der Reihe A übergeht, wodurch y ungeändert bleibt; es ist also jede Wurzel von $\psi(t, y)$ zugleich Wurzel von $\psi_1(t, y)$, und also sind beide Funktionen identisch.

Nun läßt sich y als rationale Funktion von α allein darstellen, und wenn wir $y = \chi(\alpha)$ setzen, so hat diese Funktion die in (3) geforderte Eigenschaft.

Denn nach (9) und (10) ist $\varphi(y) = 0$ und $\psi(\alpha, y) = 0$, dagegen $\psi(\alpha y_i)$, wenn $i = 1, 2, \ldots s - 1$ ist, nicht gleich Null. Die beiden Funktionen

$$\psi(\alpha, u), \quad \varphi(u)$$

haben also nur den einen linearen Faktor $u - y$ gemein, und y läßt sich also nach dem Algorithmus des größten gemeinschaftlichen Teilers rational durch α ausdrücken. Damit ist bewiesen:

5. Ein primitiver Körper hat eine primitive Gruppe.

Um aus einer Gleichung $F(x) = 0$ vom Grade m einen Normalkörper abzuleiten, haben wir eine Funktion der Wurzeln ϱ benutzt, die bei allen Vertauschungen der Wurzeln $m!$ verschiedene Werte erhält, woraus wir den Körper $\Omega(\varrho)$ ableiten. Wir konnten für ϱ irgend eine solche $m!$-wertige Funktion nehmen, und die Wurzeln von $F(x)$ sind alle in $\Omega(\varrho)$ enthalten. Es genügt aber auch, daß ϱ irgend eine primitive Größe dieses Körpers ist, also eine Funktion der Wurzeln, die nur so viele verschiedene Werte erhält, als der Grad dieses Körpers, d. h. der Grad der Gruppe beträgt.

Wenn nun $F(x) = 0$ selbst eine irreduzible Normalgleichung ist, so ist sie ihre eigene Galoissche Resolvente, und ihre Galoissche Gruppe hat denselben Grad wie die Gleichung.

Wenn umgekehrt die Gruppe von $F(x) = 0$ transitiv und vom $m$ten Grade ist, so ist x selbst eine Funktion, die durch die sämtlichen Permutationen der Gruppe P m verschiedene Werte erhält, und die also eine den Körper $\Omega(\varrho) = \Omega(x)$ erzeugende Größe ist. Daraus ergibt sich:

6. Die notwendige und hinreichende Bedingung, daß eine Gleichung $F(x) = 0$ eine Normalgleichung ist, besteht darin, daß ihre Gruppe transitiv und von demselben Grade wie $F(x)$ ist.

§ 60.

Reduktion der Galoisschen Resolvente durch Adjunktion.

Es sei jetzt $F(x) = 0$ eine Gleichung $m$ten Grades (ohne mehrfache Wurzeln), und $\alpha_1, \alpha_2, \ldots \alpha_m$ seien die Wurzeln von F; Ω sei der Rationalitätsbereich und P die Galoissche Gruppe vom Grade p. Diese habe einen Teiler Q vom Grade q und vom Index j, also

$$p = qj.$$

Es sei ϱ eine rationale Funktion der Wurzeln $\alpha_1, \alpha_2, \ldots \alpha_m$, die bei allen $m!$ Permutationen lauter voneinander verschiedene Werte annimmt, so daß der Normalkörper $N = \Omega(\alpha_1, \alpha_2, \ldots \alpha_m)$ auch durch $\Omega(\varrho)$ bezeichnet werden kann.

1. Wir nennen eine Funktion

$$\psi = \psi(\alpha_1, \alpha_1, \ldots \alpha_m)$$

zu der Gruppe Q gehörig, wenn sie 1. die Permutationen von Q gestattet und 2. durch jede nicht zu Q gehörige Permutation aus P geändert wird.

Eine solche Funktion ψ erhält man, wenn ϱ durch die Permutationen von Q in $\varrho, \varrho_1, \varrho_2 \ldots \varrho_{q-1}$ übergeht, indem man

$$\psi = \psi(t) = (t - \varrho)(t - \varrho_1) \ldots (t - \varrho_{q-1}) \tag{1}$$

setzt und unter t eine unbestimmte Größe versteht, die man so bestimmt, daß $\psi(t)$ bei jeder nicht zu Q gehörigen Permutation seinen Wert ändert. Wir zerlegen P in die Nebengruppen zu Q:

$$\begin{aligned} P &= Q + Q\pi_1 + Q\pi_2 + \cdots + Q\pi_{j-1} \\ &= Q + \pi_1^{-1}Q + \pi_2^{-1}Q + \cdots + \pi_{j-1}^{-1}Q, \end{aligned} \tag{2}$$

worin $\pi_1, \pi_2, \ldots \pi_{j-1}$ Permutationen aus P sind. Bezeichnen wir mit $\psi|\pi$ den Wert, in den ψ durch die Permutation π übergeht, so ist also

$$\psi|Q = \psi,$$

dagegen $\psi|\pi$ nicht gleich ψ, wenn π eine nicht zu Q gehörige Permutation aus P ist.

Wir wollen kürzer

$$\psi|\pi_n = \psi_n \tag{3}$$

setzen, und erhalten aus (4):

$$\psi|Q\pi_n = \psi_n.$$

Aus der Voraussetzung, daß die ψ_n alle von ψ verschieden sind, folgt, daß sie auch untereinander verschieden sind. Denn wäre $\psi_r = \psi_s$, also:

$$\psi|\pi_r = \psi|\pi_s,$$

so würde folgen:

$$\psi = \psi|\pi_r\pi_s^{-1},$$

$\pi_r\pi_s^{-1} = \varkappa$ würde zu Q gehören und es wäre

$$\pi_r = \varkappa\pi_s,$$

was nicht der Fall sein sollte.

Wir erhalten demnach aus ψ die j konjugierten Werte:

$$\psi, \psi_1, \psi_2, \dots \psi_{j-1}, \tag{4}$$

und die Funktion ψ_r gehört zu der mit Q konjugierten Gruppe $\pi_r^{-1}Q\pi_r$, wie aus (3) hervorgeht.

Wir erhalten den Satz:

2. Die konjugierten Größen (4) sind die Wurzeln einer irreduzibeln Gleichung vom Grade j in Ω.

Wenden wir nämlich auf die Größen (4) irgend eine Permutation π aus P an, so können wir, weil $\pi_r\pi$ nach (2) in einer der Nebengruppen $Q\pi_s$ enthalten ist, setzen:

$$\pi_r\pi = \varkappa\pi_s, \quad \pi_s = \varkappa^{-1}\pi_r\pi, \quad \psi|\pi_r\pi = \psi|\varkappa\pi_s = \psi_s,$$

wo $\varkappa$ in Q enthalten ist.

Es geht also ψ_r durch π in ein bestimmtes ψ_s über, und es ist nur zu zeigen, daß nicht zwei verschiedene ψ_r, ψ_{r_1} in das gleiche ψ_s übergehen können. Aus

$$\varkappa^{-1}\pi_r\pi = \varkappa_1^{-1}\pi_{r_1}\pi$$

würde aber folgen:

$$\pi_{r_1} = \varkappa_1\varkappa^{-1}\pi_r,$$

und folglich wäre auch $\psi_r = \psi_{r_1}$.

Jede symmetrische Funktion der ψ_r, z. B.

$$\varphi(t) = (t - \psi)(t - \psi_1) \dots (t - \psi_{j-1}),$$

gestattet alle Permutationen von P und ist folglich in Ω enthalten, und die Wurzeln von $\varphi(t)$ sind eben die Größen (4), wie bewiesen werden sollte.

Um die Irreduzibilität von $\varphi(t)$ nachzuweisen, nehmen wir an, es sei $\Phi(t)$ irgend eine Funktion in Ω, die für $t = \psi$ verschwindet, also $\Phi(\psi) = 0$. Da auf diese Gleichung alle Permutationen von P angewandt werden dürfen, so folgt, daß auch $\Phi(\psi_1)$, $\Phi(\psi_2)$, $\dots$ $\Phi(\psi_{j-1})$ Null sein müssen, daß also $\Phi(t)$ durch

$\varphi(t)$ teilbar sein muß. Darin aber ist die Irreduzibilität enthalten.

Hieran schließt sich der Satz von Lagrange[1]):

3. Jede Größe des Körpers $N = \Omega(\alpha_1, \alpha_2, \ldots \alpha_m)$, die die Permutationen der Gruppe Q gestattet, ist in dem Körper $\Omega(\psi)$ enthalten, wenn ψ eine zu Q gehörige Funktion ist.

Eine die Permutationen von Q gestattende Funktion ω geht durch die Permutationen einer Nebengruppe $Q\pi_1$ in ein und dieselbe Funktion ω_1 über. Es entsprechen also den konjugierten Werten

$$\psi, \psi_1, \psi_2, \ldots \psi_{j-1} \tag{5}$$

die Werte

$$\omega, \omega_1, \omega_2, \ldots \omega_{j-1}, \tag{6}$$

die jedoch nicht notwendig alle voneinander verschieden sind.

Wendet man auf die Größenreihen (5), (6) eine der Permutationen von P an, so tritt eine gewisse Permutation ein, und zwar in beiden Reihen die gleiche, da, wenn z. B. ψ_1 in ψ_2 übergeht, auch ω_1 in ω_2 übergehen muß.

Betrachten wir nun die Summe

$$\varphi(t)\left(\frac{\omega}{t-\psi} + \frac{\omega_1}{t-\psi_1} + \cdots + \frac{\omega_{j-1}}{t-\psi_{j-1}}\right) = \chi(t),$$

die eine ganze Funktion $(j-1)$ten Grades von t ist, so finden wir, daß sie durch alle Permutationen P ungeändert bleibt und folglich in Ω enthalten ist. Setzt man dann $t = \psi$ und beachtet, daß $\varphi(t)$ keine gleichen Wurzeln hat, so folgt

$$\omega = \frac{\chi(\psi)}{\varphi'(\psi)}. \tag{7}$$

Es ist also ω rational durch ψ ausgedrückt, und dies ist der Inhalt des Satzes 3.

4. Wenn wir eine zu Q gehörige Funktion ψ dem Körper Ω adjungieren, so reduziert sich die Gruppe des Körpers N auf Q.

Denn bezeichnen wir den Körper $\Omega(\psi)$ mit Ω', so gestattet erstens jede Gleichung in Ω' zwischen den Wurzeln der Grund-

[1]) Lagrange, Réflexions sur la résolution algébrique des équations. Mémoires de l'Académie de Berlin, années 1770, 1771. Oeuvres de Lagrange. Tome III. Der Satz ist von Lagrange allerdings nur in einer spezielleren Fassung gegeben. Die allgemeine Formulierung rührt von Galois her.

gleichung $\alpha_1, \alpha_2, \ldots \alpha_m$ die Permutationen von Q, weil Q in P enthalten ist, und weil ψ durch Q ungeändert bleibt; und zweitens ist jede Funktion, die durch Q ungeändert bleibt, nach dem Lagrangeschen Satze in Ω' enthalten. Dies sind aber [nach § 58, a), b)] die charakteristischen Merkmale der Galoisschen Gruppe im Körper Ω'. Um also die Galoissche Gruppe vom Grade p auf den Grad q zu reduzieren, muß eine Wurzel einer Hilfsgleichung $j$ten Grades in Ω adjungiert werden.

Den Satz 4. können wir auch so ausdrücken:

Der Normalkörper $N = \Omega(\alpha_1, \alpha_2, \ldots \alpha_m)$ ist ein Körper $p$ten Grades über Ω und $q$ten Grades über $\Omega' = \Omega(\psi)$.

Der Erniedrigung der Gruppe durch Adjunktion von ψ entspricht, wie schon aus den allgemeinen Grundsätzen hervorgeht, eine Zerfällung der Galoisschen Resolvente. Nehmen wir an, es sei $g(t) = 0$ die Galoissche Resolvente und ϱ eine ihrer Wurzeln, und durch die Permutationen von Q gehe ϱ über in

$$\varrho, \varrho_1, \varrho_2, \cdots \varrho_{q-1}, \tag{8}$$

worin, wenn $\varkappa_h$ irgend eine Permutation aus Q ist, $\varrho_h = \varrho \,|\, \varkappa_h$ ist.

Wendet man auf diese Größen eine Permutation der Gruppe Q an, so werden sie nur untereinander permutiert, und ihre symmetrischen Funktionen, und folglich auch die Funktion der Variablen t

$$g(t, \psi) = (t - \varrho)(t - \varrho_1) \ldots (t - \varrho_{q-1}) \tag{9}$$

sind in $\Omega(\psi)$ enthalten. Zugleich ist $g(t, \psi)$ ein Teiler von $g(t)$.

Wenden wir auf ψ irgend eine Permutation π_i an und setzen

$$\psi \,|\, \pi_i = \psi_i,$$

so ist $\psi_i = \psi$ oder verschieden von ψ, je nachdem π_i zu Q gehört oder nicht. Entsprechend der Zerlegung

$$P = Q + Q\pi_1 + Q\pi_2 + \cdots + Q\pi_{j-1}$$

erhalten wir j verschiedene Funktionen

$$\psi, \psi_1, \psi_2, \ldots \psi_{j-1}. \tag{10}$$

Die Reihe der Größen (8) möge durch π_i übergehen in

$$\varrho_{0,i}, \varrho_{1,i}, \varrho_{2,i}, \cdots \varrho_{q-1,i} \tag{11}$$

und zwei dieser Reihen sind entweder ganz identisch (von der Reihenfolge abgesehen), wenn π_i in Q enthalten ist, oder ganz voneinander verschieden, wenn π_i nicht in Q enthalten ist. Da nun ψ_i zu der konjugierten Gruppe $\pi_i^{-1} Q \pi_i$ gehört, so ist

$$g(t, \psi_i) = (t - \varrho_{0,i})(t - \varrho_{1,i}) \ldots (t - \varrho_{q-1,i})$$

ein in $\Omega(\psi_i)$ enthaltener Teiler von $g(t)$, und es ergibt sich die Zerlegung

$$g(t) = g(t, \psi)\, g(t, \psi_1) \ldots g(t, \psi_{j-1}). \tag{12}$$

Sind Q und Q' zwei verschiedene Teiler der Gruppe P, so werden Q und Q' gewisse Permutationen gemein haben, unter denen sich immer die identische Permutation findet. Es ist möglich, daß dies die einzige gemeinsame Permutation von Q und Q' ist, und dann heißen diese beiden Gruppen teilerfremd. Es können aber noch mehr gemeinsame Elemente vorhanden sein. Den Inbegriff R aller gemeinsamen Permutationen von Q und Q' nennen wir den größten gemeinschaftlichen Teiler, oder den Durchschnitt von Q und Q'. Dies R ist immer eine Gruppe, denn wenn π_1 und π_2 beide sowohl in Q als in Q' vorkommen, so muß auch $\pi_1 \pi_2$ in Q und in Q', also auch in R vorkommen. Der Begriff ist sofort übertragbar auf mehrere Gruppen $Q, Q', Q'' \ldots$

Ist ψ eine zu Q und ψ' eine zu Q' gehörige Funktion, so können wir die rationalen Zahlen x, x' so bestimmen, daß $\omega = x\psi + x'\psi'$ eine zu R gehörige Funktion wird, denn jedenfalls gestattet ω die Permutationen von R, da ψ und ψ' sie gestatten. Ist dann π eine nicht in R enthaltene Permutation aus P, so ist sicher nicht zugleich $\psi = \psi \mid \pi$ und $\psi' = \psi' \mid \pi$; also können wir x, x' so bestimmen, daß auch nicht $\omega = \omega \mid \pi$ wird. Wenn wir also gleichzeitig ψ und ψ', und folglich ω adjungieren, so reduziert sich die Gruppe des Körpers N nach 4. auf R. Ebenso können wir bei mehr als zwei Gruppen Q schließen, und erhalten den Satz:

5. Sind $Q, Q', Q'' \ldots$ Teiler von P und R ihr Durchschnitt, sind ferner $\psi, \psi', \psi'' \ldots$ Funktionen, die zu $Q, Q', Q'' \ldots$ gehören, so reduziert sich die Gruppe des Körpers N durch gleichzeitige Adjunktion von $\psi, \psi', \psi'' \ldots$ auf R.

Wenn wir nicht bloß die Funktion ψ, sondern alle mit ψ konjugierten Größen $\psi, \psi_1, \psi_2 \ldots \psi_{j-1}$ adjungieren, wenn wir also aus Ω den Körper

$$\Omega'' = \Omega(\psi, \psi_1, \psi_2, \ldots \psi_{j-1})$$

ableiten, so ist nach diesem Satze der Erfolg der, daß die Gruppe von N in Ω'' der größte gemeinschaftliche Teiler R aller mit Q konjugierten Gruppen wird, so daß R der Durchschnitt aller

Gruppen $\pi^{-1} Q \pi$ ist, wenn π die Permutationen von P durchläuft.

Von Interesse ist nun der Fall, daß alle konjugierten Gruppen $\pi^{-1} Q \pi$ miteinander identisch sind, also Q ein Normalteiler von P ist. In diesem Falle ist nach dem Theorem 3. jede der konjugierten Größen

$$\psi, \psi_1, \psi_2 \dots \psi_{j-1}$$

in $\Omega(\psi)$ enthalten, die Körper $\Omega(\psi)$, $\Omega(\psi_1)$, $\dots$ $\Omega(\psi_{j-1})$ und $\Omega(\psi, \psi_1, \dots \psi_{j-1})$ sind identisch, und $\Omega(\psi)$ ist ein Normalkörper über Ω.

Die Hilfsgleichung $\varphi(t) = 0$, von der die Bestimmung der Funktion ψ abhängt, geht in die Galoissche Resolvente über, wenn der Teiler Q, zu dem ψ gehört, die identische Gruppe ist. Denn dann kann nach dem Satze von Lagrange jede Funktion des Körpers N, also auch jede Wurzel α, rational durch ψ ausgedrückt werden, und N ist mit $\Omega(\psi)$ identisch.

Wir wollen diese Gleichungen $\varphi(t) = 0$ daher in einem allgemeineren Sinne Resolventen nennen. Es sind aber hier zwei Fälle zu unterscheiden.

1. Wenn die konjugierten Gruppen $\pi^{-1} Q \pi$ teilerfremd sind, so ist die gleichzeitige Adjunktion sämtlicher Wurzeln der Resolvente $\varphi(t) = 0$ gleichbedeutend mit der Adjunktion einer zur Einheitsgruppe gehörigen Funktion, und die Lösung der gegebenen Gleichung ist auf die vollständige Lösung der Resolvente zurückgeführt. Es ist

$$N = \Omega(\psi, \psi_1 \dots \psi_{j-1}),$$

und eine Galoissche Resolvente der Gleichung $\varphi(t) = 0$ ist zugleich eine Galoissche Resolvente der ursprünglichen Gleichung. In diesem Falle nennen wir $\varphi(t) = 0$ eine Totalresolvente der gegebenen Gleichung.

2. Haben die konjugierten Gruppen $\pi^{-1} Q \pi$ einen von der Einheitsgruppe verschiedenen Teiler R vom Grade r, der dann ein Normalteiler von P ist, so ist die gleichzeitige Adjunktion von sämtlichen zu ψ konjugierten Funktionen gleichbedeutend mit der Adjunktion einer zu R gehörigen Größe. Die Galoissche Resolvente der gegebenen Gleichung ist durch diese Adjunktion noch nicht vollständig gelöst, sondern sie ist nur in Faktoren vom Grade r zerlegt. In diesem Falle heißt $\varphi(t) = 0$ eine Partialresolvente.

Ist P eine einfache Gruppe, so existieren nur Totalresolventen, während, wenn P Normalteiler hat, zu jedem solchen Normalteiler eine Partialresolvente gefunden werden kann.

Derselbe Unterschied tritt auch hervor, wenn wir die Galoissche Gruppe der Resolvente $\varphi(t)$ untersuchen. Diese Gruppe besteht aus allen Vertauschungen, die in der Reihe der Größen

$$\psi, \psi_1, \psi_2, \dots \psi_{j-1} \tag{13}$$

durch Anwendung der sämtlichen Operationen π von P hervorgerufen werden; denn jede Gleichung in Ω zwischen den Größen (13) bleibt richtig, wenn eine dieser Permutationen vorgenommen wird, da man die Operationen π auf jede Gleichung zwischen den α, also auch zwischen den ψ, anwenden kann; und wenn eine Funktion in Ω der Größen (13) alle diese Permutationen gestattet, so gestattet sie auch alle Permutationen π und ist also gleich einer Größe in Ω.

Es ist nun der Grad dieser Gruppe zu bestimmen. Unter den Operationen von P werden die und nur die unter den Größen (13) keine Veränderung hervorrufen, die gleichzeitig in den Gruppen $\pi^{-1} Q \pi$ aller dieser Funktionen, also auch in ihrem größten gemeinsamen Teiler R vorkommen. Die Operationen von R mögen mit σ bezeichnet sein. Sind dann π und π' zwei Operationen, die unter den Größen (13) dieselbe Permutation hervorrufen, so wird $\pi' \pi^{-1}$ die ursprüngliche Anordnung der ψ wieder herstellen, also gleich einer der Permutationen σ sein, oder

$$\pi' = \sigma \pi.$$

Wir haben daher das Ergebnis:

Die Permutationen der Nebengruppe $R\pi$ und nur diese rufen unter den Größen (13) eine und dieselbe Permutation hervor.

Der Grad der Galoisschen Gruppe der Resolvente $\varphi(t) = 0$ ist also gleich der Anzahl dieser Nebengruppen, d. h. gleich dem Quotienten $p : r$ oder dem Index des Teilers R von P, und die Gruppe selbst ist isomorph mit der zu R komplementären Gruppe P/R (§ 45).

Ist R die identische Gruppe, also $r = 1$, und folglich $\varphi(t) = 0$ eine Totalresolvente, so ist der Grad ihrer Gruppe ebenso hoch, wie der Grad der Gruppe der ursprünglichen Gleichung, und beide Gruppen sind überdies isomorph. In bezug auf die Gruppe ist

also nichts gewonnen. Die gegebene Gleichung ist mit der Resolvente, so verschieden auch ihre Grade sein mögen, äquivalent.

Ist auf der anderen Seite Q ein Normalteiler, also K mit Q identisch, so ist $\varphi(t) = 0$ eine Partialresolvente und der Grad ihrer Gruppe ist gleich dem Index j des Teilers Q von P. Nach der Adjunktion einer Wurzel dieser Resolvente reduziert sich die Gruppe der ursprünglichen Gleichung auf Q, also auf den Grad q. Es ist also eine Spaltung der Gruppe erfolgt.

Wenn man Resolventen bilden will von möglichst niedrigem Grade, so hat man Teiler der Gruppe P aufzusuchen von möglichst kleinem Index, also von möglichst hohem Grade. Will man aber eine Reduktion der Gruppe herbeiführen, so hat man einen Normalteiler von P aufzusuchen.

Die Aufgabe der Lösung einer algebraischen Gleichung wird nach der Galoisschen Auffassung durch eine andere ersetzt, nämlich durch Adjunktion von algebraischen Größen möglichst einfacher Natur eine Zerfällung der Galoisschen Resolvente, also eine Erniedrigung des Grades der Gruppe herbeizuführen.

Wir stellen jetzt die Frage so. In dem ursprünglichen Körper Ω ist die Galoissche Resolvente $g(t)$ irreduzibel. Der Körper Ω soll zu einem anderen Körper Ω' so erweitert werden, daß $g(t)$ reduzibel wird. Ω' soll dabei ein algebraischer Körper über Ω sein, und es muß also eine algebraische Größe ε geben, so daß $\Omega' = \Omega(\varepsilon)$ wird. Die Größe ε wird Wurzel einer gewissen irreduzibeln Gleichung in Ω sein, die wir mit

$$\chi(\varepsilon) = 0 \tag{14}$$

bezeichnen wollen.

Nehmen wir an, in dem Körper $\Omega(\varepsilon)$ sondere sich von $g(t)$ der irreduzible Faktor

$$g_1(t) = g_1(t, \varepsilon) \tag{15}$$

ab, der die Wurzel $t = \varrho$ hat. Den Grad von $g_1(t, \varepsilon)$ wollen wir mit q bezeichnen und die Wurzeln mit

$$\varrho, \varrho_1, \varrho_2, \cdots \varrho_{q-1}, \tag{16}$$

so daß

$$g_1(t, \varepsilon) = (t - \varrho)(t - \varrho_1) \ldots (t - \varrho_{q-1}). \tag{17}$$

Zunächst ist nun nachzuweisen, daß die in der Gruppe P enthaltenen Substitutionen

$$(\varrho, \varrho), (\varrho, \varrho_1), (\varrho, \varrho_2), \ldots (\varrho, \varrho_{q-1})$$

eine Gruppe Q bilden. Um dies zu zeigen, setzen wir

$$\varrho_i = \Theta_i(\varrho),$$

worin Θ eine Funktion in Ω bedeutet, und bedenken, daß dann die Gleichung

$$g_1[\Theta_i(t), \varepsilon] = 0$$

eine Wurzel, nämlich $t = \varrho$, mit $g_1(t)$, gemein hat, und mithin, da $g_1(t)$ in $\Omega(\varepsilon)$ irreduzibel ist, durch $g_1(t)$ teilbar ist. Daraus folgt, daß, wenn ϱ_1 und ϱ_2 irgend zwei der Wurzeln (16) sind, auch $\Theta_1(\varrho_2) = \varrho_3$ unter diesen Wurzeln enthalten ist. Es ist aber die Substitution (ϱ, ϱ_3) aus (ϱ, ϱ_1) und (ϱ, ϱ_2) zusammengesetzt, und damit ist die Gruppeneigenschaft von Q nachgewiesen. Wir bezeichnen den Index von Q mit j und setzen

$$p = jq. \tag{18}$$

Das Produkt (17) gestattet nun die Substitutionen der Gruppe Q, und wenn also ψ wie oben eine zu dieser Gruppe gehörige Funktion in N bedeutet, so läßt sich nach dem Satze von Lagrange die Funktion $g_1(t, \varepsilon)$ rational durch ψ ausdrücken, d. h. $g_1(t, \varepsilon)$ ist eine Funktion von t im Körper $j$ten Grades $\Omega(\psi)$, und soll demnach durch $g(t, \psi)$ bezeichnet werden.

Die Größe ψ ist eine der Wurzeln einer irreduzibeln Gleichung vom Grade j

$$\varphi(u) = 0, \tag{19}$$

deren übrige Wurzeln $\psi_1, \psi_2, \ldots \psi_{j-1}$ sind, und nach (12) ist $g(t)$ in die Faktoren zerlegbar:

$$g(t) = g(t, \psi)\, g(t, \psi_1) \ldots g(t, \psi_{j-1}). \tag{20}$$

Nun ist

$$g(t, \psi) = g_1(t, \varepsilon),$$

und aus (20) folgt, daß diese Gleichung nicht bestehen bleibt, wenn auf der linken Seite für ψ eine der anderen Wurzeln von (19) gesetzt wird. Denn sonst müßte $g(t, \psi)$ mit einem der übrigen Faktoren auf der rechten Seite von (20) identisch sein. Man kann also für t einen solchen rationalen Wert setzen, daß die Gleichung

$$g(t, u) - g_1(t, \varepsilon) = 0 \tag{21}$$

nur die eine Wurzel $u = \psi$ mit der Gleichung (19) gemein hat. Der größte gemeinschaftliche Teiler von (20) und (21) ist dann in bezug auf u linear, und wenn man darin $u = \psi$ setzt, so ergibt sich, daß ψ rational durch ε ausdrückbar ist, und

daß die Adjunktion von ψ zu Ω genügt, um den Faktor $g_1(t)$ von der Galoisschen Resolvente abzusondern.

Der Körper $\Omega(\psi)$ ist ein Teiler des Körpers $\Omega(\varepsilon)$. Der Grad des Körpers $\Omega(\varepsilon)$, d. h. der Grad der Hilfsgleichung $\chi(\varepsilon) = 0$, ist daher ein Vielfaches von j und niemals niedriger als j. Wenn der Grad der Hilfsgleichung gleich j ist, was z. B. dann immer eintritt, wenn der Grad von χ eine Primzahl ist, so ist $\Omega(\varepsilon)$ mit $\Omega(\psi)$ identisch, d. h. ε ist auch rational durch ψ ausdrückbar. In diesem Falle sind die konjugierten Werte $\varepsilon, \varepsilon_1, \ldots \varepsilon_{j-1}$ die sämtlichen Wurzeln von $\chi = 0$, und man erhält sie, wenn man in dem rationalen Ausdruck von ε durch ψ die Funktion ψ durch jede der konjugierten Funktionen $\psi, \psi_1, \ldots \psi_{j-1}$ ersetzt. Es kann also ε selbst für ψ genommen werden, und man findet eine Zerlegung

$$g(t) = g(t, \varepsilon)\, g(t, \varepsilon_1) \ldots g(t, \varepsilon_{j-1}).$$

Die Größen $\varepsilon, \varepsilon_1, \varepsilon_2, \ldots \varepsilon_{j-1}$ gehören sämtlich dem Körper $\Omega(\varrho)$ an.

Diese Resultate sind von großer Wichtigkeit und verdienen besonders hervorgehoben zu werden.

Nach Kronecker heißen, wenn $F(x) = 0$ die gegebene Gleichung ist, und N der zugehörige Galoissche Körper, die Größen von N die der Gleichung $F(x) = 0$ natürlichen Irrationalitäten.

Wir haben dann den Satz:

6. Jede mögliche Reduktion der Galoisschen Gruppe wird herbeigeführt durch Adjunktion einer natürlichen Irrationalität. Ist j der Index der reduzierten Gruppe in bezug auf die ursprüngliche, so kann die Reduktion nicht eintreten durch Adjunktion eines Körpers von niedrigerem als dem $j$ten Grade. Der Grad kann aber gleich j sein, und dann ist die adjungierte Irrationalität eine natürliche. Ist der Grad des adjungierten Körpers größer als j, so ist er ein Vielfaches von j.

§ 61.

Gegenseitige Reduktion zweier Körper.

Es sei in den folgenden Betrachtungen Ω der ein für allemal zugrunde gelegte Rationalitätsbereich und $N = \Omega(\vartheta)$ ein algebrai-

scher Körper $n$ten Grades über Ω. Die konjugierten Werte $\vartheta, \vartheta_1, \vartheta_2, \ldots \vartheta_{n-1}$ seien alle in N enthalten, also N ein Normalkörper und $\sigma_i = (\vartheta, \vartheta_i)$ seien die n Substitutionen der Gruppe S dieses Körpers. Zu dieser Gruppe gehört Ω in N, denn nur die Zahlen des Körpers N, die zugleich in Ω enthalten sind, bleiben durch S ungeändert.

Wir betrachten nun zwei in N enthaltene Körper über Ω:

$$A = \Omega(\overset{a}{\alpha}), \qquad B = \Omega(\overset{b}{\beta}) \tag{1}$$

der Grade a und b. Hier können A und B beliebige Körper über Ω sein. Denn man kann immer einen Normalkörper N angeben, der die beiden Körper A und B enthält. Man kann z. B. für N den Körper nehmen, der aus allen zu α und zu β konjugierten Werten abgeleitet ist.

Es sei

$$M = \Omega(\alpha, \overset{m}{\beta}) \tag{2}$$

das gemeinschaftliche Vielfache der Körper A und B, dessen Grad wir mit m bezeichnen.

Der Körper M ist imprimitiv, da er den Körper B enthält, und ist ein Körper über $\Omega(\beta)$, dessen Grad mit a' bezeichnet werde. Es ist dann nach § 55:

$$m = b\,a' \quad \text{und ebenso} \quad m = a\,b'. \tag{3}$$

Die primitive Zahl α des Körpers A genügt einer Gleichung vom Grade a in Ω:

$$\varphi(\overset{a}{u}) = 0, \tag{4}$$

und in $\Omega(\beta)$ einer Gleichung vom Grade a':

$$\eta(\overset{a'}{u}, \beta) = 0, \tag{5}$$

worin u eine Variable ist, und $\varphi(u, \beta)$ ist ein Teiler von $\varphi(u)$, der in B irreduzibel ist.

Es ist folglich $\varphi(u, \beta)$ ein Produkt aus a' Linearfaktoren

$$\varphi(u, \beta) = (u - \alpha)\,(u - \alpha')\,(u - \alpha'')\ldots, \tag{6}$$

wenn die $\alpha, \alpha', \alpha'' \ldots$ einige unter den konjugierten Werten zu α sind, und $\varphi(u)$ scheidet nach Adjunktion von β einen irreduzibeln Faktor vom Grade a' aus.

Es sei Δ der größte gemeinschaftliche Teiler von A und B, d. h. der Körper der Zahlen δ, die zugleich in A und B enthalten sind, und d dessen Grad.

Wenn nun A ein Normalkörper ist, so sind alle $\alpha, \alpha', \alpha'' \ldots$ und mithin nach (6) auch die Koeffizienten von $\varphi(u, \beta)$ in A enthalten, und da $\varphi(u, \beta)$ zugleich in B enthalten ist, so ist es auch in Δ enthalten. Wenn $\varphi(u, \beta)$ in B irreduzibel ist, so ist es auch in jedem Teiler von B, also auch in Δ irreduzibel, und folglich ist $\varphi(u, \beta) = \varphi(u, \delta)$ die Funktion niedrigsten Grades in Δ, die für $u = \alpha$ verschwindet, d. h. der Grad a' von M über B ist zugleich der Grad von A über Δ. Also ist

$$m : b = a : d$$

oder:

$$md = ab. \tag{7}$$

Es ist hier vorausgesetzt, daß A ein Normalkörper sei, und da man durchweg A und B vertauschen darf, so genügt es für das Bestehen von (7), daß einer der beiden Körper ein Normalkörper sei.

Es ergibt sich aus (3) und (7):

$$a' = \frac{m}{b} = \frac{a}{d}, \quad b' = \frac{m}{a} = \frac{b}{d},$$

und wenn wir beide Körper A und B als Normalkörper annehmen, so können wir hieraus das Theorem folgern:

7. Wenn der Normalkörper A vom Grade a durch Adjunktion des Normalkörpers B vom Grade b auf den Grad a/d reduziert wird, so wird B durch Adjunktion von A auf den Grad b/d reduziert; d ist der Grad des größten gemeinschaftlichen Teilers und $m = ab/d$ der Grad des gemeinschaftlichen Vielfachen der beiden Körper[1]).

[1]) C. Jordan, Traité des substitutions, art. 378. Frobenius, Mathematische Annalen, Bd. 70.

Zehnter Abschnitt.

Zyklische Gleichungen.

§ 62.

Kubische und biquadratische Gleichungen.

Wir wollen jetzt die Gruppentheorie auf die Auflösung der Gleichungen dritten und vierten Grades anwenden.

Die Gruppe der Permutationen von drei Ziffern besteht aus sechs Elementen, nämlich aus der identischen Permutation 1, aus zwei dreigliedrigen Zyklen und drei Transpositionen:

$$(1) \qquad \begin{matrix} 1 & (0,1,2) & (0,2,1) \\ (1,2) & (2,0) & (0,1). \end{matrix}$$

Die drei ersten

$$1 \qquad (0,1,2) \qquad (0,2,1)$$

bilden die alternierende Gruppe, sie besteht aus den Potenzen der zyklischen Permutation $\pi = (0, 1, 2)$:

$$(2) \qquad 1, \pi, \pi^2,$$

und ist also eine zyklische Gruppe.

Es seien nun α, α_1, α_2 die drei Wurzeln der kubischen Gleichung:

$$(3) \qquad f(x) = x^3 - ax^2 + bx - c = 0.$$

Legen wir den Körper Ω zugrunde, der aus allen rationalen Funktionen der unabhängigen Veränderlichen a, b, c besteht, so ist (1) die Gruppe dieser Gleichung; wenn wir aber

$$(4) \qquad \sqrt{D} = (\alpha - \alpha_1)\,(\alpha - \alpha_2)\,(\alpha_1 - \alpha_2)$$

adjungieren, worin

$$(5) \qquad D = a^2b^2 + 18abc - 4a^3c - 4b^3 - 27c^2$$

die Diskriminante der Gleichung (3) ist [§ 15, (14)], so reduziert sich die Gruppe auf (2).

In dem Körper Ω', der durch diese Adjunktion aus Ω entsteht, kann jede Wurzel rational durch jede andere ausgedrückt werden, denn es ist

$$\alpha_1 + \alpha_2 = a - \alpha$$
$$\alpha_1 - \alpha_2 = \frac{\sqrt{D}}{f'(\alpha)},$$

also

$$2\alpha_1 = a - \alpha + \frac{\sqrt{D}}{f'(\alpha)}$$
$$2\alpha_2 = a - \alpha - \frac{\sqrt{D}}{f'(\alpha)}, \tag{6}$$

und hierin können die Vertauschungen π, π^2 ausgeführt werden. Die kubische Gleichung ist also nach Adjunktion von $\sqrt{D}$ ihre eigene Galoissche Resolvente.

Alles dies bleibt gültig, wenn der Rationalitätsbereich irgend ein spezieller Körper Ω ist, in dem a, b, c und $\sqrt{D}$ enthalten sind, wenn nicht $f(x)$ selbst in Ω reduzibel ist, also eine rationale Wurzel hat. Denn außer sich selbst und der zyklischen Gruppe (2) hat die Gruppe (1) nur noch intransitive Teiler, nämlich die Einheitsgruppe und drei Gruppen vom Typus 1, (1, 2).

Wollen wir die kubische Gleichung nun auflösen, d. h. auf eine reine kubische Gleichung zurückführen, so müssen wir eine Funktion v der Wurzeln suchen, die zwar nicht selbst, deren Kubus aber die zyklische Permutation π gestattet. Eine solche Funktion kann aber nur existieren, wenn die dritten Einheitswurzeln, oder, was dasselbe ist, $\sqrt{-3}$, dem Körper Ω adjungiert wird; denn v muß durch Anwendung von π eine dritte Einheitswurzel als Faktor erhalten.

Durch diese Adjunktion kann eine Reduktion der Gruppe nicht eintreten, weil die Gruppe (2) außer der Einheitsgruppe keinen Teiler hat und die Reduktion auf die Einheitsgruppe nicht durch eine quadratische, sondern nur durch eine kubische Gleichung geschehen kann (§ 60). Wir setzen also

$$\varepsilon = \frac{-1 + \sqrt{-3}}{2}, \quad \varepsilon^2 = \frac{-1 - \sqrt{-3}}{2}$$

und

$$v = \alpha + \varepsilon\alpha_1 + \varepsilon^2\alpha_2$$
$$v' = \alpha + \varepsilon^2\alpha_1 + \varepsilon\alpha_2; \tag{7}$$

dann ist der Erfolg von π der, daß v in $\varepsilon^2 v$, v' in $\varepsilon v'$ übergeht, während v^3, v'^3 und vv' ungeändert bleiben, also in dem Körper

Ω' enthalten sind. Es hat keine Schwierigkeit, diese Größen nach den Sätzen über die symmetrischen Funktionen zu berechnen. Setzen wir zur Abkürzung

$$A = \alpha^2\alpha_1 + \alpha_1^2\alpha_2 + \alpha_2^2\alpha,$$
$$A' = \alpha_1^2\alpha + \alpha_2^2\alpha_1 + \alpha^2\alpha_2,$$

so ist [§ 22, (3)]

$$A + A' = ab - 3c$$
$$A - A' = \sqrt{D},$$

und für v^3 erhält man

$$v^3 = \alpha^3 + \alpha_1^3 + \alpha_2^3 + 6\alpha\alpha_1\alpha_2 + 3\varepsilon A + 3\varepsilon^2 A',$$

also

$$v^3 = a^3 - \frac{9}{2}ab + \frac{27}{2}c + \frac{3}{2}\sqrt{-3D}, \tag{8}$$

und ebenso

$$v'^3 = a^3 - \frac{9}{2}ab + \frac{27}{2}c - \frac{3}{2}\sqrt{-3D}. \tag{9}$$

Es ist aber auch

$$vv' = \alpha^2 + \alpha_1^2 + \alpha_2^2 - \alpha\alpha_1 - \alpha\alpha_2 - \alpha_1\alpha_2 = a^2 - 3b, \tag{10}$$

und aus (8), (9), (10) erhält man

$$-(v^3 - v'^3)^2 = 27D = 4(a^2 - 3b)^3 - (2a^3 - 9ab + 27c)^2.$$

Fügt man zu (7) noch die Gleichung

$$a = \alpha + \alpha_1 + \alpha_2,$$

so findet man in Übereinstimmung mit der Cardanischen Formel:

$$3\alpha = a + v + v'.$$
$$3\alpha_1 = a + \varepsilon^2 v + \varepsilon v'$$
$$3\alpha_2 = a + \varepsilon v + \varepsilon^2 v'.$$

Bei vier Ziffern 0, 1, 2, 3 hat die symmetrische Gruppe 24 Permutationen. Es sind, durch ihre Zyklen dargestellt, die folgenden:

$$\begin{aligned} P = {} & 1, (0,1), (0,2), (0,3), (1,2), (1,3), (2,3) \\ & (0,1)(2,3), (0,2)(1,3), (0,3)(1,2) \\ & (0,1,2), (0,1,3), (0,2,3), (1,2,3) \\ & (0,2,1), (0,3,1), (0,3,2), (1,3,2) \\ & (0,1,2,3), (0,1,3,2), (0,2,3,1) \\ & (0,2,1,3), (0,3,1,2), (0,3,2,1). \end{aligned} \tag{11}$$

Wir haben in P außer der identischen Permutation sechs Transpositionen, acht dreigliedrige und sechs viergliedrige Zyklen und drei Transpositionspaare, die mit der identischen Permu-

tation zusammen eine Gruppe vom Grade 4 bilden. Diese heißt die Vierergruppe:

$$(12) \qquad R = 1,\ (0,1)(2,3),\ (0,2)(1,3),\ (0,3)(1,2).$$

Daß R eine Gruppe ist, ersieht man sofort, weil durch die Komposition je zweier ihrer von 1 verschiedenen Permutationen die dritte entsteht. Bildet man die beiden Nebengruppen $(0,1,2)\,R$, $(0,2,1)\,R$, so erhält man mit R zusammen die ganze alternierende Gruppe Q vom Grade 12:

$$(13) \quad \begin{aligned} Q &= R + (0,1,2)\,R + (0,2,1)\,R \\ &= 1,\ (0,1)\,(2,3),\ (0,2)\,(1,3),\ (0,3)\,(1,2) \\ &\quad (0,1,2),\ (1,3,2),\ (0,3,1),\ (0,2,3) \\ &\quad (0,2,1),\ (0,3,2),\ (1,2,3),\ (0,1,3). \end{aligned}$$

Da es nur drei Transpositionspaare gibt und die konjugierten Gruppen alle vom gleichen Typus sein müssen, so folgt, daß R ein Normalteiler der alternierenden Gruppe Q vom Index 3 und der symmetrischen Gruppe G vom Index 6 ist. Wir haben also hier eine Ausnahme des Satzes von der Einfachheit der alternierenden Gruppe (§ 51).

Benutzt man die Komposition:

$$S = R + (0,1)\,R = R + R\,(0,1),$$

so erhält man einen Teiler von P vom Grade 8 und vom Index 3:

$$(14) \quad \begin{aligned} S = 1,\ &(0,1)\,(2,3),\ (0,2)\,(1,3),\ (0,3)\,(1,2), \\ &(0,1),\ (2,3),\ (0,3,1,2),\ (0,2,1,3). \end{aligned}$$

Daß dies eine Gruppe ist, folgt sofort aus $(0,1)\,R = R\,(0,1)$. Sie ist nicht in Q, sondern nur in P enthalten. Sie ist auch kein Normalteiler von P, sondern es gibt drei konjugierte Teiler, die R zum Durchschnitt haben.

In P ist noch eine zyklische Gruppe von vier Elementen enthalten:

$$(15) \qquad C = 1,\ (0,3,1,2),\ (0,1)(2,3),\ (0,2,1,3),$$

zu der es wieder drei konjugierte gibt. Diese sind Teiler von P vom Index 6.

Unter den intransitiven Gruppen verdient noch eine Gruppe hervorgehoben zu werden:

$$(16) \qquad T = 1,\ (0,1),\ (2,3),\ (0,1)(2,3),$$

die den Index 6 hat und zu einem System von drei konjugierten Teilern von P führt. Außerdem gibt es von intransitiven Teilern

von P nur noch Permutationsgruppen von zwei oder drei Elementen.

Diese Zerlegungen sollen jetzt auf die biquadratische Gleichung angewandt werden. Sei also

$$f(x) = x^4 - a_1 x^3 + a_2 x^2 - a_3 x + a_4 \tag{17}$$

eine Funktion vierten Grades, in der zunächst die a_1, a_2, a_3, a_4 Veränderliche sind, und $\alpha, \alpha_1, \alpha_2, \alpha_3$ seien die Wurzeln von $f(x)$. Adjungieren wir die Quadratwurzel aus der Diskriminante

$$\sqrt{D} = (\alpha - \alpha_1)(\alpha - \alpha_2)(\alpha - \alpha_3)(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3), \tag{18}$$

so reduziert sich die Galoissche Gruppe der Gleichung $f(x) = 0$ von P auf Q (weil $\sqrt{D}$ zu Q gehört).

Man findet leicht eine Funktion, die zu der Gruppe R gehört, nämlich:

$$y = (\alpha - \alpha_1)(\alpha_2 - \alpha_3).$$

Diese Funktion ist innerhalb Q dreiwertig. Sie erhält die Werte:

$$\begin{aligned} y &= (\alpha - \alpha_1)(\alpha_2 - \alpha_3) \\ y_1 &= (\alpha - \alpha_2)(\alpha_3 - \alpha_1) \\ y_2 &= (\alpha - \alpha_3)(\alpha_1 - \alpha_2), \end{aligned} \tag{19}$$

und die symmetrischen Funktionen dieser drei Werte sind also rationale Funktionen der a und $\sqrt{D}$. Das Produkt $y y_1 y_2$ ist $\sqrt{D}$ selbst. Die Summe $y + y_1 + y_2$ verschwindet, und die Größe

$$-A = y y_1 + y y_2 + y_1 y_2 \tag{20}$$

ist eine symmetrische Funktion der α, also eine rationale Funktion der a. Die Größen (19) sind daher die Wurzeln der kubischen Gleichung

$$y^3 - Ay + \sqrt{D} = 0. \tag{21}$$

Dies ist die Resolvente (16) in § 35. Es ist eine Normalgleichung, weil R ein Normalteiler in Q ist. Es ist also dasselbe, ob wir einen oder alle drei Werte y adjungieren.

Durch die Transpositionen (α, α_1) geht y, y_1, y_2 in $-y, -y_2, -y_1$ über und folglich ändert sich

$$B = (y - y_1)(y - y_2)(y_1 - y_2) \tag{22}$$

nicht; und da dasselbe für alle Transpositionen gilt, so ist B eine symmetrische Funktion der α, also eine rationale Funktion der a. Die Diskriminante der kubischen Gleichung (21)

$$B^2 = -27D + 4A^3 \tag{23}$$

ist also ein Quadrat in Ω, woraus zu erkennen, daß (21) eine Normalgleichung ist.

Adjungiert man y, so reduziert sich die Gruppe der biquadratischen Gleichung auf

$$R = 1, \quad (0, 2)\,(1, 3), \quad (0, 1)\,(2, 3), \quad (0, 3)\,(1, 2).$$

Diese Gruppe ist imprimitiv (und zwar nach drei Arten), und die Gleichung kann jetzt durch zwei Quadratwurzeln gelöst werden. Am besten gelangt man dazu auf folgendem Wege. Die drei Größen

$$\begin{aligned} v_1 &= (\alpha + \alpha_1 - \alpha_2 - \alpha_3)^2 \\ v_2 &= (\alpha - \alpha_1 + \alpha_2 - \alpha_3)^2 \\ v_3 &= (\alpha - \alpha_1 - \alpha_2 + \alpha_3)^2 \end{aligned} \tag{24}$$

gestatten alle die Permutationen der Gruppe R und sind daher rational durch y darstellbar.

Die Quadratwurzeln aus diesen Größen:

$$\begin{aligned} \sqrt{v_1} &= \alpha + \alpha_1 - \alpha_2 - \alpha_3 \\ \sqrt{v_2} &= \alpha - \alpha_1 + \alpha_2 - \alpha_3 \\ \sqrt{v_3} &= \alpha - \alpha_1 - \alpha_2 + \alpha_3 \end{aligned} \tag{25}$$

ändern sich durch die Permutationen der Gruppe R so, daß eine von ihnen das Zeichen beibehält, die beiden anderen das Zeichen wechseln, so daß das Produkt $\sqrt{v_1}\cdot\sqrt{v_2}\cdot\sqrt{v_3}$ ungeändert bleibt und folglich dem Körper $\Omega(y)$ angehört. Fügt man noch die Gleichung

$$a_1 = \alpha + \alpha_1 + \alpha_2 + \alpha_3 \tag{26}$$

hinzu, so ergibt sich die vollständige Auflösung der biquadratischen Gleichung:

$$\begin{aligned} 4\alpha &= a_1 + \sqrt{v_1} + \sqrt{v_2} + \sqrt{v_3} \\ 4\alpha_1 &= a_1 + \sqrt{v_1} - \sqrt{v_2} - \sqrt{v_3} \\ 4\alpha_2 &= a_1 - \sqrt{v_1} + \sqrt{v_2} - \sqrt{v_3} \\ 4\alpha_3 &= a_1 - \sqrt{v_1} - \sqrt{v_2} + \sqrt{v_3}. \end{aligned} \tag{27}$$

Ändert man die Vorzeichen von zweien der Quadratwurzeln, so vertauschen sich die vier Wurzeln $\alpha, \alpha_1, \alpha_2, \alpha_3$ nach der Gruppe R Eine Transposition, wie etwa (α, α_1), oder, was auf dasselbe hinauskommt, die Vorzeichenänderung von $\sqrt{D}$, bewirkt keine Änderung von $\sqrt{v_1}$, dagegen eine gleichzeitige Vorzeichenänderung und Vertauschung von $\sqrt{v_2}, \sqrt{v_3}$.

Eine zyklische Permutation der drei Wurzeln $\alpha_1, \alpha_2, \alpha_3$ bewirkt eine zyklische Permutation von y, y_1, y_2 und daher werden die Wurzeln α in derselben Weise permutiert, wenn man eine Wurzel y der Gleichung (21) durch eine andere ersetzt.

Die Funktion v_1 selbst gehört zu der Gruppe S und ist infolgedessen Wurzel einer kubischen Resolvente in Ω. Sie führt auf die kubische Resolvente, die sich aus der Ferrarischen Methode ergibt, und die für den Fall $a_1 = 0$ in § 35, (4) aufgestellt ist.

Welche Werte man den mehrwertigen algebraischen Größen, die bei der Auflösung auftreten, auch beilegen mag, die Ausdrücke (27) stellen immer die Wurzeln unserer biquadratischen Gleichung in irgend einer Reihenfolge dar.

In bezug auf die übrigen Wege zur Auflösung der biquadratischen Gleichung können wir uns kürzer fassen.

Wenn wir zunächst nicht $\sqrt{D}$ adjungieren, sondern eine zu der Gruppe S gehörige Funktion, so erhalten wir eine kubische Resolvente, die nicht Normalgleichung ist. Wir können für diese Funktion etwa y^2 wählen, das der kubischen Gleichung

$$y^6 - 2Ay^4 + A^2y^2 - D = 0 \tag{28}$$

genügt. Besser noch nimmt man als Wurzeln der kubischen Resolvente

$$z = y_1 - y_2, \quad z_1 = y_2 - y, \quad z_2 = y - y_1. \tag{29}$$

Es gehört dann z zur Gruppe S. Aus (20), (22) erhält man für z die kubische Gleichung

$$z^3 - 3Az + B = 0. \tag{30}$$

Man kann ferner zur Lösung der biquadratischen Gleichung dadurch gelangen, daß man zuerst eine zu der zyklischen Gruppe C gehörige (eine zyklische) Funktion der Wurzeln adjungiert, wie z. B.

$$w = \alpha\alpha_1^2 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2 + \alpha_3\alpha^2. \tag{31}$$

Diese Funktion ist sechswertig und ist also die Wurzel einer Gleichung sechsten Grades. Diese Gleichung sechsten Grades ist aber, wie man leicht sieht, imprimitiv, und kann auf eine Gleichung dritten Grades und auf zwei Quadratwurzeln zurückgeführt werden. Diese Gleichung sechsten Grades ist eine Totalresolvente, und zwei ihrer Wurzeln genügen zur rationalen Darstellung der Wurzeln α. Die Form dieser Resolvente wird nicht einfach.

Man kann endlich noch darauf ausgehen, durch Adjunktion einer zur Gruppe T gehörigen Funktion ξ die Funktion $f(x)$ direkt reduzibel zu machen und in zwei quadratische Faktoren

zu zerlegen. Eine solche Funktion ξ ist gleichfalls die Wurzel einer Gleichung sechsten Grades, die sich aber auch durch Imprimitivität auf den dritten und zweiten Grad reduziert. Nimmt man z. B.

$$\xi = \alpha + \alpha_1 - \tfrac{1}{2}a_1 = \tfrac{1}{2}\sqrt{v_1}, \tag{32}$$

so sind von den sechs Wurzeln je zwei entgegengesetzt gleich, so daß eine kubische Gleichung für ξ^2 resultiert. Durch eine dieser Größen ξ lassen sich dann die quadratischen Faktoren von $f(x)$ rational darstellen.

Nehmen wir zur Vereinfachung der Formeln $a_1 = 0$ an und

$$f(x) = x^4 + ax^2 + bx + c, \tag{33}$$

so wird $\xi = \alpha + \alpha_1 = -\alpha_2 - \alpha_3$ und

$$\begin{aligned} (\alpha\alpha_1 - \alpha_2\alpha_3)\xi &= b, \\ \alpha\alpha_1 + \alpha_2\alpha_3 &= \xi^2 + a, \end{aligned}$$

also

$$2\alpha\alpha_1 = \frac{b}{\xi} + \xi^2 + a, \quad 2\alpha_2\alpha_3 = -\frac{b}{\xi} + \xi^2 + a, \tag{34}$$

so daß die beiden quadratischen Faktoren von $f(x)$ folgende werden:

$$\begin{aligned} x^2 - \xi x + \tfrac{1}{2}\left(\frac{b}{\xi} + \xi^2 + a\right) &= 0 \\ x^2 + \xi x + \tfrac{1}{2}\left(-\frac{b}{\xi} + \xi^2 + a\right) &= 0, \end{aligned} \tag{35}$$

und für ξ^2 ergibt sich aus (34) durch Multiplikation die kubische Gleichung

$$4c = (\xi^2 + a)^2 - \frac{b^2}{\xi^2},$$

oder

$$\xi^6 + 2a\xi^4 + (a^2 - 4c)\xi^2 - b^2 = 0. \tag{36}$$

Nimmt man für ξ irgend eine Wurzel dieser Gleichung sechsten Grades, so gibt jede der Gleichungen (35) ein Paar Wurzeln von $f(x) = 0$.

§ 63.

Abelsche Gleichungen.

Der Grad einer transitiven Permutationsgruppe von m Ziffern ist immer durch m teilbar und also niemals kleiner als m. Denn ist P eine solche Gruppe, und Q_0 der Inbegriff der Permutationen von P, die die Ziffer 0 ungeändert lassen, so ist auch

Q_0 eine Gruppe und also ein Teiler von P. Nun kann man wegen der vorausgesetzten Transitivität in P ein System von Permutationen $\pi_1, \pi_2, \ldots \pi_{m-1}$ finden, die 0 in 1, 0 in 2, ... 0 in $m-1$ überführen, und dann ist $Q_0\pi_1$ das System aller der Permutationen von P, die 0 in 1 verwandeln. Danach ist

$$(1) \qquad P = Q_0 + Q_0\pi_1 + Q_0\pi_2 + \cdots + Q_0\pi_{m-1},$$

und der Grad von P gleich dem Produkte aus m und dem Grade von Q_0, also:

I. Die Galoissche Gruppe einer irreduzibeln Gleichung ist niemals von niedrigerem Grade, als die Gleichung selbst.

Eine Normalgleichung haben wir früher durch die Bestimmung erklärt, daß sie irreduzibel sei, und daß jede ihrer Wurzeln rational durch jede andere ausdrückbar sein sollte. Daraus ergibt sich, daß die Gruppe einer Normalgleichung sich auf die identische Gruppe reduzieren muß, wenn man eine Wurzel adjungiert; und umgekehrt ist eine Gleichung, deren Gruppe so beschaffen ist, wenn sie zugleich irreduzibel ist, immer eine Normalgleichung.

Nun reduziert sich P durch Adjunktion der Wurzel α_0 auf Q_0, durch α_1 auf $\pi_1^{-1} Q_0 \pi_1$ usw., und wenn also P die Gruppe einer Normalgleichung sein soll, so ist notwendig und hinreichend, daß Q_0 die Einheitsgruppe ist, daß also durch $\pi, \pi_1, \pi_2, \ldots \pi_{m-1}$ die Gruppe erschöpft sei. Wir haben also:

II. Damit eine irreduzible Gleichung eine Normalgleichung sei, ist notwendig und hinreichend, daß der Grad der Gruppe mit dem Grade der Gleichung übereinstimme.

Wir betrachten hier zunächst die spezielle Art von Gleichungen, zu denen die von Gauß zuerst aufgelösten Kreisteilungsgleichungen gehören, die Abel allgemein auflösen gelehrt hat, und die wir also nach ihm Abelsche Gleichungen nennen wollen[1]).

III. Eine Gleichung $m$ten Grades $F(x) = 0$ mit den Wurzeln $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ heißt eine Abelsche Gleichung, wenn jede Wurzel rational durch eine von ihnen, α, ausdrückbar ist, und wenn, falls

[1]) Abel, Mémoire sur une classe d'équations résolubles algébriquement. Crelles Journal f. Mathematik, Bd. 4, 1829. Oeuvres complètes, nouvelle édition, 1881, Bd. 1, S. 418.

(2) $$\alpha_1 = \Theta_1(\alpha), \quad \alpha_2 = \Theta_2(\alpha). \ldots \alpha_{m-1} = \Theta_{m-1}(\alpha)$$

diese rationalen Ausdrücke sind, die Bedingung

(3) $$\Theta_h \Theta_k(\alpha) = \Theta_k \Theta_h(\alpha)$$

für je zwei dieser Funktionen besteht.

Es bedeutet hierin das Zeichen $\Theta_h \Theta_k(\alpha)$, daß die Funktion $\Theta_h(x)$ für das Argument $x = \Theta_k(\alpha)$ gebildet werden soll. Selbstverständlich bezieht sich diese ganze Definition auf einen bestimmten als rational angenommenen Körper Ω.

Wenn die Funktion $F(x)$ nicht irreduzibel ist, so hat sie einen bestimmten irreduzibeln Faktor $\varphi(x)$, der die Wurzel α hat, und unter den Wurzeln von $\varphi(x)$ bestehen gleichfalls die durch (2) und (3) ausgesprochenen Relationen. Es ist daher $\varphi(x) = 0$ eine Galoissche Resolvente von $F(x) = 0$, und wenn $\varphi(x) = 0$ gelöst ist, so sind damit alle Wurzeln von $F(x)$ bekannt.

Die Galoissche Gruppe einer Abelschen Gleichung, sei sie irreduzibel oder nicht (wenn nur ihre Wurzeln voneinander verschieden sind), hat die Eigenschaft, daß bei der Zusammensetzung ihrer Permutationen das kommutative Gesetz gilt; sie ist also eine kommutative Gruppe.

Es seien nämlich

(4) $$\alpha, \quad \alpha_1 = \Theta_1(\alpha), \quad \alpha_2 = \Theta_2(\alpha) \ldots \alpha_{m-1} = \Theta_{m-1}(\alpha)$$

die Wurzeln von $F(x)$ und

(5) $$\alpha, \quad \alpha' = \Theta'(\alpha), \quad \alpha'' = \Theta''(\alpha) \ldots$$

die darunter enthaltenen Wurzeln des irreduzibeln Faktors $\varphi(x)$ von $F(x)$.

Da, wie schon bemerkt, $\varphi(x) = 0$ eine Galoissche Resolvente ist, so besteht die Gruppe der Gleichung aus den Substitutionen des Körpers $\Omega(\alpha)$, also aus den Substitutionen

(6) $$\sigma = (\alpha, \alpha), \quad \sigma' = (\alpha, \alpha'), \quad \sigma'' = (\alpha, \alpha'') \ldots$$

Diese Gruppe befolgt aber das kommutative Gesetz; denn es sei

$$\sigma' = (\alpha, \alpha') = [\alpha, \Theta'(\alpha)]$$
$$\sigma'' = (\alpha, \alpha'') = [\alpha, \Theta''(\alpha)];$$

dann ist

$$\sigma'\sigma'' = [\alpha, \Theta'(\alpha)]\,[\Theta'(\alpha), \quad \Theta''\Theta'(\alpha)] = [\alpha, \Theta''\Theta'(\alpha)]$$
$$\sigma''\sigma' = [\alpha, \Theta''(\alpha)]\,[\Theta''(\alpha), \quad \Theta'\Theta''(\alpha)] = [\alpha, \Theta'\Theta''(\alpha)].$$

Wegen (3) ist daher $\sigma'\sigma'' = \sigma''\sigma'$, worin σ', σ'' zwei beliebige der Substitutionen σ sein können. Folglich ist die Gruppe der Sub-

stitutionen des Körpers $\Omega(\alpha)$ und damit auch die isomorphe Permutationsgruppe der Gleichung $F(x) = 0$ kommutativ.

Es gilt nun auch das Umgekehrte, wobei aber die Irreduzibilität vorausgesetzt sein muß.

IV. Eine irreduzible Gleichung $\varphi(x) = 0$ mit kommutativer Gruppe ist eine Abelsche Gleichung.

Es seien α, α_1, α_2, ... α_{m-1} die Wurzeln von $\varphi(x)$, und P die Gruppe dieser Gleichung, die wegen der Irreduzibilität von $\varphi(x)$ transitiv ist, und die wir jetzt außerdem als kommutativ annehmen. Es sei Q der Teiler von P, der das Element 0 in Ruhe läßt. Ist dann π_i eine Permutation, die die Ziffer 0 in i überführt, so ist $\pi_i^{-1} Q \pi_i$ die Gruppe der Permutationen in P, die die Ziffer i nicht ändern. Da nun aber in jeder Zusammensetzung von Permutationen aus P die Komponenten vertauscht werden können, so ist

$$\pi_i^{-1} Q \pi_i = \pi_i^{-1} \pi_i Q = Q,$$

d. h. die Gruppe Q läßt auch die Ziffer i ungeändert. Wegen der Transitivität von P kann aber i jede der Ziffern 1, 2, ... $m-1$ bedeuten, folglich besteht Q aus der einzigen identischen Permutation, und es gibt in P außer der identischen keine Permutation, die eine Ziffer ungeändert läßt. Adjungiert man aber eine Wurzel α, so reduziert sich die Gruppe P auf eine Gruppe, die α ungeändert läßt, also auf die Einheitsgruppe, und die Gleichung ist gelöst, d. h. jede Wurzel von $\varphi(x)$ kann durch eine beliebige unter ihnen, α, rational ausgedrückt werden. Demnach ist $\varphi(x) = 0$ eine Normalgleichung und somit ihre eigene Galoissche Resolvente. Ist $\alpha_k = \Theta_k(\alpha)$, so besteht die Galoissche Gruppe dieser Gleichung aus den Substitutionen

$$\sigma_k = [\alpha, \Theta_k(\alpha)],$$

und es ist

$$\sigma_h \sigma_k = [\alpha, \Theta_h \Theta_k(\alpha)], \quad \sigma_k \sigma_h = [\alpha, \Theta_k \Theta_h(\alpha)].$$

Da die Gruppe kommutativ sein soll, so muß

$$\Theta_h \Theta_k(\alpha) = \Theta_k \Theta_h(\alpha)$$

sein, d. h. $\varphi(x) = 0$ ist eine Abelsche Gleichung. Dies ist der Grund, weshalb die kommutativen Gruppen auch Abelsche Gruppen genannt werden. Es folgt also noch aus dem oben gegebenen Satze über die Gruppe von Normalgleichungen, daß bei einer transitiven Abelschen Permutationsgruppe die Zahl

der Permutationen mit der Zahl der vertauschten Ziffern übereinstimmt.

Hier haben wir die Irreduzibilität der gegebenen Gleichung vorausgesetzt. Wollen wir auch reduzible Gleichungen $F(x) = 0$ mit kommutativer Gruppe P berücksichtigen, so betrachten wir einen irreduzibeln Teiler $\varphi(x)$ von $F(x)$. Die Gruppe P' der Gleichung $\varphi(x) = 0$ erhält man, wenn man die Permutationen von P auf die Wurzeln von $\varphi(x)$ anwendet, und wenn daher P kommutativ ist, so ist es auch P'. Daraus ergibt sich:

V. Hat eine reduzible Gleichung $F(x) = 0$ eine kommutative Gruppe, so gibt jeder irreduzible Teiler φ von F eine Abelsche Gleichung $\varphi = 0$.

Es ist aber nicht notwendig, daß die Wurzeln des einen Teilers φ rational durch die eines anderen ausdrückbar sind, und daher können wir, im Hinblick auf unsere Definition, die Gleichung $F(x) = 0$ nicht immer als eine Abelsche bezeichnen.

Es ist endlich noch zu bemerken, daß bei einer kommutativen Gruppe jeder Teiler normal ist, da ja immer $\pi^{-1}\varkappa\pi = \varkappa$ ist, wenn π und $\varkappa$ beliebige Elemente einer kommutativen Gruppe sind.

Wir haben vorhin gesehen, daß in einer transitiven Abelschen Gruppe außer der identischen keine Permutation vorkommt, die eine Ziffer ungeändert läßt. Nehmen wir an, eine Permutation π einer solchen Gruppe sei in ihre Zyklen zerlegt, und der Zyklus, der die wenigsten Glieder enthält, sei ein r-gliedriger. Dann wird π^r die Glieder dieses Zyklus ungeändert lassen und muß also die identische Permutation sein. Daraus folgt aber, daß auch alle übrigen Zyklen von π, wenn noch andere vorhanden sind, aus r Gliedern bestehen müssen, also:

VI. Eine Permutation einer transitiven Abelschen Gruppe enthält nur Zyklen von gleicher Gliederzahl.

Die Anzahl r der Glieder eines Zyklus muß ein Teiler von m sein, wenn m der Grad der Gruppe ist, und wenn $m = rs$ ist, so ist s die Anzahl der r-gliedrigen Zyklen, aus denen π besteht.

Ist nun P die Gruppe einer Abelschen Gleichung $F(x) = 0$, so nehmen wir irgend eine nicht identische Permutation π aus P heraus und zerlegen sie in ihre Zyklen

$$\pi = \gamma\gamma_1\gamma_2 \cdots \gamma_{s-1}, \tag{7}$$

worin jeder der Zyklen γ sich auf r Wurzeln der Gleichung $F(x) = 0$ bezieht. Wir wollen diese Wurzeln so anordnen, daß

$$(8)\qquad \begin{aligned} \gamma &= (\alpha, \alpha_1, \alpha_2 \dots \alpha_{r-1}) \\ \gamma_1 &= (\beta, \beta_1, \beta_2 \dots \beta_{r-1}) \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{s-1} &= (\sigma, \sigma_1, \sigma_2 \dots \sigma_{r-1}) \end{aligned}$$

wird.

Ist π_1 irgend eine Permutation von P, so ist wegen der Vertauschbarkeit

$$(9)\qquad \pi_1^{-1}\pi\pi_1 = \pi.$$

Nach dem Satze über die Bildung der Permutationen $\pi_1^{-1}\pi\pi_1$ (§ 49, 4.) darf also π nicht geändert werden, wenn die Permutationen π_1 in den Zyklen von π ausgeführt werden. Da aber die Zyklen vollständig bestimmt sind, abgesehen von ihrem Anfangselement, so ergibt sich, daß durch Anwendung irgend einer Permutation π_1 aus P die Elemente der einzelnen Zyklen γ nicht voneinander getrennt, sondern nur untereinander (zyklisch) vertauscht und außerdem die Zyklen miteinander vertauscht werden können. Die Gruppe ist also, wenn $s > 1$ ist, imprimitiv.

Eine rationale Funktion der Argumente $\alpha, \alpha_1, \dots \alpha_{r-1}$, die ihren Wert nicht ändert, wenn die α der zyklischen Permutation γ und ihren Wiederholungen unterworfen werden, heißt eine zyklische Funktion der α. Bezeichnet man mit

$$(10)\qquad \omega = \psi(\alpha, \alpha_1, \dots \alpha_{r-1})$$

eine solche zyklische Funktion und mit

$$(11)\qquad \omega, \omega_1, \omega_2, \dots \omega_{s-1}$$

die konjugierten Werte von ω, setzt also z. B.

$$(12)\qquad \omega_1 = \psi(\beta, \beta_1, \dots \beta_{r-1}),$$

so sind diese Größen nach § 59 die Wurzeln einer irreduzibeln Gleichung $s$ten Grades:

$$\Phi(t) = 0,$$

deren Gruppe man erhält, wenn man die durch P hervorgerufenen Permutationen der Größen (11) aufsucht.

Man erhält aber die durch $\pi_1\pi_2$ unter den Größen (11) bewirkte Permutation, wenn man die beiden durch π_1 und π_2 einzeln hervorgerufenen Permutationen zusammensetzt, und die Gruppe der Permutationen von (11) ist daher auch kommutativ. Daher ist $\Phi(t) = 0$ eine Abelsche Gleichung $s$ten Grades.

Adjungiert man ω, so zerfällt $F(x)$ in s Faktoren $r$ten Grades:

$$F(x) = F(x, \omega)\, F(x, \omega_1) \ldots F(x, \omega_{s-1}),$$

von denen der erste, $F(x, \omega)$, die Wurzeln $\alpha, \alpha_1, \ldots \alpha_{r-1}$, hat, und die Gruppe der Gleichung $F(x, \omega) = 0$ besteht allein aus der Periode der zyklischen Permutation γ.

Wir wollen eine Gleichung, deren Gruppe aus einem einzigen Zyklus und seinen Wiederholungen besteht, eine zyklische Gleichung nennen, so daß die zyklischen Gleichungen der einfachste Spezialfall der Abelschen Gleichungen sind. Wir haben dann also bewiesen, daß die Lösung jeder Abelschen Gleichung zurückgeführt wird auf die Lösung einer Abelschen Gleichung niedrigeren Grades und auf die Lösung einer Reihe zyklischer Gleichungen. Diesen Satz kann man wieder auf die Hilfsgleichung $s$ten Grades anwenden, und damit so lange fortfahren, bis diese Hilfsgleichung sich auf den ersten Grad reduziert. Damit erhält man also das Resultat:

VII. Die Lösung einer Abelschen Gleichung läßt sich immer auf die Lösung einer Reihe von zyklischen Gleichungen zurückführen, deren Grade Teiler des Grades der gegebenen Gleichung sind.

Es ist nicht notwendig, in die Definition der zyklischen Gleichungen die Irreduzibilität mit aufzunehmen. Wir können daher allgemein die Definition so fassen:

VIII. Eine Gleichung $m$ten Grades $F(x) = 0$ mit m verschiedenen Wurzeln heißt eine zyklische Gleichung im Körper Ω, wenn ihre Wurzeln $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ nicht rational sind, aber sich so anordnen lassen, daß die zyklischen Funktionen der Wurzeln in Ω rational sind.

Wenn also

$$\pi = (\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}) \tag{13}$$

ist, so muß jede Funktion in Ω enthalten sein, die die Permutationen der zyklischen Gruppe

$$C = 1, \pi, \pi^2, \pi^3 \ldots \pi^{m-1} \tag{14}$$

gestattet. Die Galoissche Gruppe einer zyklischen Gleichung ist entweder die Periode C selbst und dann ist die Gleichung irreduzibel, oder sie ist ein Teiler von C; sie besteht dann, wenn e und f zwei ganzzahlige Faktoren von m sind und

$$m = ef$$

ist, aus den Permutationen

$$C_e = 1, \pi^e, \pi^{2e}, \ldots \pi^{(f-1)e}, \tag{15}$$

und die zyklische Gleichung zerfällt in e Faktoren $f$ten Grades. Denn nehmen wir eine zu der Gruppe C gehörige Funktion $\psi(\alpha, \alpha_1, \ldots \alpha_{m-1})$, so ist diese nach Voraussetzung gleich einer Größe in Ω, die wir mit a bezeichnen. Auf die rationale Gleichung $\psi = a$ ist dann keine nicht in C enthaltene Permutation anwendbar, und folglich kann die Gruppe der Gleichung keine anderen Substitutionen enthalten als solche, die in C vorkommen. Wenn nun π in der Gruppe der Gleichung vorkommt, so ist sie mit C identisch, und da C transitiv ist, ist die Gleichung irreduzibel. Ist aber π^e die niedrigste Potenz von π, die in der Gruppe der Gleichung vorkommt, so ist C_e diese Gruppe. C_e ist aber, wenn $e > 1$ ist, intransitiv, und die Gleichung ist reduzibel.

Die zyklischen Gleichungen haben, wie alle Abelschen Gleichungen, die Eigenschaft, daß jede Wurzel rational durch jede andere ausdrückbar ist. Hier lassen sich diese Ausdrücke folgendermaßen zyklisch anordnen. Sind $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ die Wurzeln der zyklischen Gleichung $F(x) = 0$, so ist die ganze Funktion $(m-1)$ten Grades von x

$$F(x)\left(\frac{\alpha_1}{x-\alpha} + \frac{\alpha_2}{x-\alpha_1} + \cdots + \frac{\alpha}{x-\alpha_{m-1}}\right) = \varphi(x)$$

ungeändert durch die Permutation π und also in Ω enthalten. Wenn wir darin $x = \alpha, \alpha_1, \ldots \alpha_{m-1}$ setzen und das Zeichen

$$\frac{\varphi(x)}{F'(x)} = \Theta(x)$$

einführen, so folgt

$$\alpha_1 = \Theta(\alpha),\ \alpha_2 = \Theta(\alpha_1),\ \ldots \alpha_{m-1} = \Theta(\alpha_{m-2}),\ \alpha = \Theta(\alpha_{m-1}), \tag{16}$$

und dies gilt, mag $F(x)$ reduzibel oder irreduzibel sein, wenn nur $F(x)$ und $F'(x)$ keinen gemeinsamen Teiler haben.

Wenn der Grad der zyklischen Gleichung keine Primzahl ist so läßt sie sich durch das oben auf Abelsche Gleichungen im allgemeinen angewandte Verfahren noch weiter reduzieren.

Wenn nämlich $m = ef$ irgend eine Zerlegung von m in zwei Faktoren ist, so zerfällt die Permutation π^e in e Zyklen γ von je f Gliedern:

$$(17)\qquad \begin{aligned} \gamma &= (\alpha, \alpha_e, \alpha_{2e}, \dots \alpha_{(f-1)e}) \\ \gamma_1 &= (\alpha_1, \alpha_{e+1}, \alpha_{2e+1}, \dots \alpha_{(f-1)e+1}) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{e-1} &= (\alpha_{e-1}, \alpha_{2e-1}, \alpha_{3e-1}, \dots \alpha_{m-1}). \end{aligned}$$

Nun bestimmen wir eine zu dem ersten dieser Zyklen gehörige Funktion ψ und setzen

$$(18)\qquad \begin{aligned} \eta &= \psi(\alpha, \alpha_e, \alpha_{2e}, \dots \alpha_{(f-1)e}) \\ \eta_1 &= \psi(\alpha_1, \alpha_{e+1}, \alpha_{2e+1}, \dots \alpha_{(f-1)e+1}) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \eta_{e-1} &= \psi(\alpha_{e-1}, \alpha_{2e-1}, \alpha_{3e-1}, \dots \alpha_{m-1}). \end{aligned}$$

Diese Größen sind alle voneinander verschieden, aber bei einer zyklischen Permutation ihrer Argumente bleiben sie ungeändert. Durch Anwendung der Permutation π gehen die Größen

$$\eta, \eta_1, \eta_2, \dots \eta_{e-1}$$

zyklisch ineinander über, und nach der Voraussetzung sind also ihre zyklischen Funktionen und folglich auch ihre symmetrischen Funktionen rational. Sie sind also die Wurzeln einer zyklischen Gleichung $e$ten Grades, während $F(x)$ in e Faktoren $f$ten Grades

$$F(x) = F(x, \eta)\, F(x, \eta_1) \dots F(x, \eta_{e-1})$$

zerfällt, deren jeder eine zyklische Gleichung $f$ten Grades für die Wurzeln eines der Zyklen γ ergibt.

Denn setzen wir etwa

$$F_1 = (x - \alpha)\,(x - \alpha_e) \dots (x - \alpha_{(f-1)e}),$$

so gestattet diese Funktion die Permutationen der Periode von γ. Also gestattet sie auch die Permutationen der Gruppe, die nach Adjunktion von η zur Galoisschen Gruppe unserer Gleichung wird, die ja nur aus Potenzen von $\gamma, \gamma_1, \dots \gamma_{e-1}$ bestehen kann. Es ist daher nach dem Satze von Lagrange (§ 60, 3.) F_1 rational durch η darstellbar, also $F_1 = F(x, \eta)$. Die Gleichung $F(x, \eta) = 0$ ist aber wieder zyklisch in $\Omega(\eta)$, da die zyklischen Funktionen ihrer Wurzeln diesem Körper angehören.

Die Auflösung der zyklischen Gleichungen $m$ten Grades ist hierdurch abhängig gemacht von der Lösung zyklischer Gleichungen, deren Grade die Primfaktoren von m sind. Ist also z. B. m eine Potenz von 2, so wird die Lösung durch eine Reihe von Quadratwurzeln bewerkstelligt. Auf diesem Wege hat Gauß zuerst die Kreisteilungsgleichungen behandelt[1]).

[1]) Gauß, Disquisitiones arithmeticae, Sectio VII.

Wir knüpfen noch die für die Folge wichtige Bemerkung hier an, daß für einen Primzahlgrad die Begriffe der Normalgleichung und der zyklischen Gleichung zusammenfallen. Denn jedenfalls ist eine zyklische Gleichung von Primzahlgrad, da sie irreduzibel ist, eine Normalgleichung. Und wenn umgekehrt der Grad n einer Normalgleichung, der zugleich der Grad der Gruppe dieser Gleichung ist, eine Primzahl ist, und π eine nicht identische Permutation dieser Gruppe, so ist der Grad von π, der ja ein Teiler von n sein muß, gleich n, und die Gruppe der Gleichung ist $1, \pi, \pi^2, \ldots \pi^{n-1}$, also zyklisch.

Wenn die Galoissche Gruppe $\mathfrak{G}$ einer Abelschen Gleichung nach der Komposition der Teile (§ 45) in zwei Gruppen zerlegt ist, so daß in der Form

$$\mathfrak{G} = \mathfrak{A}\mathfrak{B}$$

jedes Element von $\mathfrak{G}$ ein- und nur einmal erscheint, so tritt noch eine weitere Reduktion der Abelschen Gleichung ein. Sind m, a, b die Grade von $\mathfrak{G}, \mathfrak{A}, \mathfrak{B}$, so ist $m = ab$. Man nehme zwei Zahlen ξ und η, die zu $\mathfrak{A}$ und zu $\mathfrak{B}$ gehören (§ 60, 1.) und bilde mit rationalen Koeffizienten α, β eine m-wertige Zahl

$$x = \alpha\xi + \beta\eta. \tag{19}$$

Die Gruppen $\mathfrak{G}\,|\,\mathfrak{A}$, $\mathfrak{G}\,|\,\mathfrak{B}$ der Grade b und a der Gleichungen, deren Wurzeln ξ und η sind, sind gleichfalls Abelsche und x ist eine primitive Zahl des Körpers $\Omega(x)$, dessen Gruppe G ist. Nach § 52 läßt sich eine Abelsche Gruppe durch eine Basis darstellen, deren Elemente Primzahlpotenzen zu Graden haben, und so kommen wir durch Wiederholung dieses Verfahrens nach VII. zu dem Satze:

IX. Die Wurzel einer Abelschen Gleichung läßt sich rational (in Ω) darstellen durch die Wurzeln zyklischer Gleichungen, deren Grade Primzahlpotenzen sind.

§ 64.

Resolventen von Lagrange.

Die Methode der Auflösung zyklischer Gleichungen, die wir jetzt kennen lernen wollen, ist gleichmäßig auf Primzahlgrade und zusammengesetzte Grade anwendbar. Man bedient sich dazu gewisser Ausdrücke, die unter dem Namen der Resolventen von Lagrange bekannt und bei allen Untersuchungen über

die algebraische Auflösung von Gleichungen von großem Nutzen sind[1]).

Wenn wir

$$r = \cos\frac{2\pi}{m} + i\sin\frac{2\pi}{m} = e^{i\frac{2\pi}{m}}$$

setzen, so ist nach dem Moivreschen Satze, wenn k eine positive oder negative ganze Zahl ist:

$$r^k = \cos\frac{2\pi k}{m} + i\sin\frac{2\pi k}{m},$$

und mithin:

$$r^m = 1, \qquad r^{km} = 1.$$

Die Größen r^k heißen darum $m$te Einheitswurzeln oder Einheitswurzeln vom Grade m. Es gibt deren m verschiedene für $k = 0, 1, 2 \ldots m-1$, und andere Werte von k ergeben dieselben Werte wieder. Wir bezeichnen die voneinander verschiedenen dieser Einheitswurzeln mit ε und erhalten aus der Summenformel der geometrischen Reihe:

$$(1)\qquad \begin{aligned} \overset{\varepsilon}{\Sigma}\varepsilon^k &= 0, \text{ wenn } k \text{ nicht durch } m \text{ teilbar ist,} \\ &= m, \text{ wenn } k \text{ durch } m \text{ teilbar, also } \varepsilon^k = 1 \text{ ist.} \end{aligned}$$

Die Einheitswurzel ε heißt primitiv, wenn sie nicht zugleich eine Einheitswurzel niedrigeren Grades ist, wenn also k in $\varepsilon = r^k$ teilerfremd zu m ist, sonst imprimitiv.

Es sei $F(x) = 0$ eine Gleichung mit den voneinander verschiedenen Wurzeln $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ und

$$(2)\qquad (\varepsilon, \alpha) = \alpha + \varepsilon\alpha_1 + \varepsilon^2\alpha_2 + \cdots + \varepsilon^{m-1}\alpha_{m-1}.$$

Die so definierten Summen sind es, die man die Lagrangeschen Resolventen nennt. Wenn diese Funktionen für alle $m$ten Einheitswurzeln ε bekannt sind, so ist auch die Gleichung selbst gelöst, denn aus (1) erhält man:

$$(3)\qquad m\alpha = \overset{\varepsilon}{\Sigma}(\varepsilon, \alpha),$$

worin sich die Summe über alle $m$ten Einheitswurzeln ε erstreckt. Man kann auch die anderen Wurzeln in gleicher Weise ausdrücken:

$$(4)\qquad m\alpha_k = \overset{\varepsilon}{\Sigma}\varepsilon^{-k}(\varepsilon, \alpha),$$

[1]) Lagrange, Réflexions etc., s. S. 508. Früher haben wir unter Resolventen auflösende Gleichungen verstanden, hier sind es auflösende Funktionen.

so daß in der Tat alles auf die Kenntnis von ε und der Funktionen (ε, α) zurückgeführt ist.

Die Summen (3) und (4) lassen sich noch in etwas anderer Weise darstellen, da die sämtlichen $m$ten Einheitswurzeln Potenzen von r sind. Setzt man also $\varepsilon = r^\lambda$ und läßt λ die Reihe der Zahlen $0, 1 \ldots m-1$ durchlaufen, so können wir für die Gleichungen (3), (4) setzen:

$$m\alpha = \sum^{\lambda} (r^\lambda, \alpha), \qquad m\alpha_k = \sum^{\lambda} r^{-\lambda k} (r^\lambda, \alpha).$$

Wir untersuchen diese Resolventen (ε, α) zunächst als Funktionen der m unabhängigen Veränderlichen α, und wollen wegen der einfacheren Darstellungsweise der Formeln übereinkommen, daß $\alpha_m = \alpha_0 = \alpha$ und überhaupt $\alpha_h = \alpha_k$ sein soll, wenn sich h von k durch ein Vielfaches von m unterscheidet $[h \equiv k \pmod m]$.

Für diese Resolventen gelten nun die folgenden Sätze.

X. Wenn man auf die Indices der α die zyklische Permutation $\pi = (0, 1, \ldots m-1)$ anwendet, so geht (ε, α) in $\varepsilon^{-1}(\varepsilon, \alpha)$ über, und durch die Permutation π^k geht (ε, α) in $\varepsilon^{-k}(\varepsilon, \alpha)$ über.

Dies zeigt die Definition (2) der Resolventen unmittelbar.

Wir verstehen ferner unter ν einen beliebigen positiven Exponenten und bilden nach dem polynomischen Lehrsatze $(\varepsilon, \alpha)^\nu$. In der entwickelten Potenz setzen wir $\varepsilon^m = 1$ und ordnen nach Potenzen von ε. Es ergibt sich dann ein Ausdruck von der Form

$$(5) \qquad (\varepsilon, \alpha)^\nu = A_0^{(\nu)} + \varepsilon A_1^{(\nu)} + \varepsilon^2 A_2^{(\nu)} + \cdots + \varepsilon^{m-1} A_{m-1}^{(\nu)} = \sum_{0, m-1}^{h} \varepsilon^h A_h^{(\nu)},$$

worin die $A_0^{(\nu)}, A_1^{(\nu)} \ldots A_{m-1}^{(\nu)}$ Formen νten Grades mit ganzzahligen Koeffizienten und den Variablen α sind, aber von ε unabhängig. Auch hier möge $A_h = A_k$ sein, so oft $h \equiv k \pmod m$. Danach beweisen wir den Satz:

XI. Wenn man auf die Indices der α die zyklische Permutation π anwendet, so erleiden die Indices der Koeffizienten von $(\varepsilon, \alpha)^\nu$ die zyklische Permutation π^ν, d. h. $A_h^{(\nu)}$ geht in $A_{h+\nu}^{(\nu)}$ über.

Um dies nachzuweisen, bemerken wir, daß die Formel (5) für jede beliebige $m$te Einheitswurzel ε, einschließlich 1, richtig

bleibt, und hiernach folgt aus (1) und (5):

$$m A_k^{(\nu)} = \sum^{\varepsilon} \varepsilon^{-k} (\varepsilon, \alpha)^\nu. \tag{6}$$

Macht man auf der rechten Seite dieser Formel in den Indices der α die Permutation π, so ergibt sich nach X.:

$$\sum^{\varepsilon} \varepsilon^{-k-\nu} (\varepsilon, \alpha)^\nu,$$

d. h. $A_k^{(\nu)}$ geht in $A_{k+\nu}^{(\nu)}$ über, wie im Satz XI. behauptet ist.

Der Satz XI. ist ein spezieller Fall eines allgemeinen Theorems. Entwickeln wir ein Produkt von beliebig vielen Faktoren

$$(\varepsilon, \alpha)^\nu (\varepsilon^{\lambda_1}, \alpha)^{\nu_1} (\varepsilon^{\lambda_2}, \alpha)^{\nu_2} \ldots,$$

worin ν, ν_1, ν_2 ... positive, λ_1, λ_2 ... beliebige ganze Zahlen sind, und ordnen es, wie vorher $(\varepsilon, \alpha)^\nu$, nach Potenzen von ε, so mag sich ergeben:

$$(\varepsilon, \alpha)^\nu (\varepsilon^{\lambda_1}, \alpha)^{\nu_1} (\varepsilon^{\lambda_2}, \alpha)^{\nu_2} \cdots = \sum_{0, m-1}^{h} \varepsilon^h B_h, \tag{7}$$

worin die B_h von ε unabhängige Formen von der Variablen α sind. Da auch diese Entwickelung für alle $m$ten Einheitswurzeln ε gilt, so kann man die Formel (1) anwenden und erhält

$$m B_k = \sum^{\varepsilon} \varepsilon^{-k} (\varepsilon, \alpha)^\nu (\varepsilon^{\lambda_1}, \alpha)^{\nu_1} (\varepsilon^{\lambda_2}, \alpha)^{\nu_2} \ldots,$$

woraus man nach X. schließen kann, daß B_k durch die Substitutionen π in

$$B_{k+\nu+\lambda_1\nu_1+\lambda_2\nu_2+\cdots}$$

übergeht. Wir haben also:

XII. Die Permutation π, auf die Indices der α angewandt, ruft unter den Indices der Koeffizienten des nach ε geordneten Produktes

$$(\varepsilon, \alpha)^\nu (\varepsilon^{\lambda_1}, \alpha)^{\nu_1} (\varepsilon^{\lambda_2}, \alpha)^{\nu_2} \ldots \tag{8}$$

die Permutation $\pi^{\nu+\lambda_1\nu_1+\lambda_2\nu_2\cdots}$ hervor.

In diesen Theoremen kann die Permutation π wiederholt werden; sie bleiben richtig, wenn π durch irgend eine Potenz von π ersetzt wird, und sie gelten, was auch ε für eine $m$te Einheitswurzel sein mag. Ist aber e ein Teiler von m,

$$m = ef,$$

und ε irgend eine $e$te Einheitswurzel, so ist ε zugleich $m$te Einheitswurzel und der Ausdruck (2) wird

$$\begin{aligned}(\varepsilon, \alpha) = \alpha &+ \varepsilon\alpha_1 &&+ \cdots + \varepsilon^{e-1}\alpha_{e-1}\\ &+ \alpha_e &+ \varepsilon\alpha_{e+1} &+ \cdots + \varepsilon^{e-1}\alpha_{2e-1}\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &+ \alpha_{e(f-1)} &+ \varepsilon\alpha_{e(f-1)+1} &+ \cdots + \varepsilon^{e-1}\alpha_{m-1}.\end{aligned}$$

Wir führen nun die Bezeichnung ein:

$$(9)\qquad \begin{aligned}\eta &= \alpha &+ \alpha_e &+ \alpha_{2e} &+ \cdots + \alpha_{(f-1)e}\\ \eta_1 &= \alpha_1 &+ \alpha_{e+1} &+ \alpha_{2e+1} &+ \cdots + \alpha_{(f-1)e+1}\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \eta_{e-1} &= \alpha_{e-1} &+ \alpha_{2e-1} &+ \alpha_{3e-1} &+ \cdots + \alpha_{m-1},\end{aligned}$$

und nennen diese Größen, wie es Gauß in dem speziellen Falle der Kreisteilung getan hat, die f-gliedrigen Perioden der Größen α. Es wird dann

$$(10)\qquad (\varepsilon, \alpha) = \eta + \varepsilon\eta_1 + \varepsilon^2\eta_2 + \cdots + \varepsilon^{e-1}\eta_{e-1},$$

wofür wir auch (ε, η) schreiben können.

Die Anwendung der Permutation π auf die Indices von α bringt unter den nach dem Modul e zu nehmenden Indices der Größen η die zyklische Permutation

$$(11)\qquad \gamma = (0, 1, 2, \ldots e-1)$$

hervor.

Entsprechend der Formel (10) können wir die Formeln (5) und (7) nun auch so schreiben:

$$\begin{aligned}(\varepsilon, \eta)^\nu &= E_0^{(\nu)} + \varepsilon E_1^{(\nu)} + \varepsilon^2 E_2^{(\nu)} + \cdots + \varepsilon^{e-1} E_{e-1}^{(\nu)}\\ (\varepsilon, \eta)^\nu\,(\varepsilon^{\lambda_1}, \eta)^{\nu_1}\,(\varepsilon^{\lambda_2}, \eta)^{\nu_2}\ldots &\\ &= G_0 + \varepsilon G_1 + \varepsilon^2 G_2 + \cdots + \varepsilon^{e-1} G_{e-1},\end{aligned}$$

worin dann

$$\begin{aligned}E_k^{(\nu)} &= A_k^{(\nu)} + A_{e+k}^{(\nu)} + \cdots + A_{(f-1)e+k}^{(\nu)}\\ G_k &= B_k + B_{e+k} + \cdots + B_{(f-1)e+k}\end{aligned}$$

ist, und $E_k^{(\nu)}$ und G_k ganze homogene Funktionen der α sind, die sich auch als Funktionen der η darstellen lassen.

Die Größen η_k, $E_k^{(\nu)}$, G_k bleiben ungeändert, wenn die Permutation π^e auf ihre Indices angewandt wird, d. h. wenn der Index um ein Vielfaches von e verändert wird, und wir können also jetzt die Sätze XI und XII so vervollständigen:

XIII. Ist ε eine beliebige $e$te Einheitswurzel, und e ein Teiler von m, so erleiden, wenn auf die Indices von α die Permutation π ausgeübt wird, die Indices k der Koeffizienten $E_k^{(\nu)}$ von $(\varepsilon, \eta)^\nu$ die Permutation γ^ν, und die Indices der Koeffizienten G des Produktes $(\varepsilon, \eta)^\nu\,(\varepsilon^{\lambda_1}, \eta)^{\nu_1}\,(\varepsilon^{\lambda_2}, \eta)^{\nu_2}\ldots$ die Permutation $\gamma^{\nu+\lambda_1\nu_1+\lambda_2\nu_2\cdots}$.

§ 65.

Auflösung der zyklischen Gleichungen.

Die Lagrangeschen Resolventen führen durch Anwendung der jetzt bewiesenen Sätze zu der Auflösung der zyklischen Gleichungen, genauer gesagt, zur Reduktion auf reine Gleichungen.

Wir verstehen jetzt unter den α nicht mehr beliebige Variable, sondern die Wurzeln einer zyklischen Gleichung, so daß die zyklischen Funktionen der α als Größen des Rationalitätsbereichs Ω zu betrachten sind.

Nach dem Theorem § 64, X. sind die Koeffizienten von $(\varepsilon, \alpha)^m$ zyklische Funktionen der α.

Verstehen wir unter $a_0, a_1, \ldots a_{m-1}$ Größen in Ω und setzen

$$\psi_\lambda = a_0 + a_1\varepsilon^\lambda + a_2\varepsilon^{2\lambda} + \cdots + a_{m-1}\varepsilon^{(m-1)\lambda}, \tag{1}$$

so folgt aus diesem Theorem:

$$(\varepsilon^\lambda, \alpha) = \sqrt[m]{\psi_\lambda}. \tag{2}$$

Bezeichnet darin ε eine $m$te aber keine niedrigere Einheitswurzel, so sind in der Form $(\varepsilon^\lambda, \alpha)$ alle Resolventen enthalten.

Bemerken wir noch, daß $(1, \alpha) = a$ als die Summe der Wurzeln zu den bekannten Größen gehört, so haben wir nach § 64, (3)

$$m\alpha = a + \sqrt[m]{\psi_1} + \sqrt[m]{\psi_2} + \cdots + \sqrt[m]{\psi_{m-1}}, \tag{3}$$

und damit also α durch Radikale $m$ten Grades ausgedrückt, die unter den Wurzelzeichen außer den Größen, die von Hause aus in Ω vorkommen, noch $m$te Einheitswurzeln enthalten.

Jedes dieser Radikale hat, für sich betrachtet, m verschiedene Werte, die sich um $m$te Einheitswurzeln als Faktoren voneinander unterscheiden. Geben wir jeder $m$ten Wurzel alle ihre Werte, so erhalten wir aus (3) viele verschiedene Werte von α, unter denen nach § 64, (4) die sämtlichen Wurzeln $\alpha, \alpha_1, \alpha_2, \ldots \alpha_{m-1}$ vorkommen. Aber die Zahl der so aus (3) abgeleiteten Ausdrücke ist viel größer, und es handelt sich noch darum, die beizubehaltenden von den abzusondernden zu unterscheiden. Am einfachsten führt dazu folgender Weg.

Wenden wir das Theorem § 64, XII. auf zwei Faktoren an, so ergibt sich, daß

$$(\varepsilon, \alpha)^\lambda\ (\varepsilon^\lambda, \alpha)^\mu$$

eine Funktion in $\Omega(\varepsilon)$ ist, wenn $\nu + \lambda\mu \equiv 0 \pmod m$. Setzen wir also $\mu = 1$, $\nu = m - \lambda$, so folgt, wenn

$$\chi_\lambda = b_0^{(\lambda)} + b_1^{(\lambda)}\varepsilon + \cdots + b_{m-1}^{(\lambda)}\varepsilon^{m-1}$$

eine Größe in $\Omega(\varepsilon)$ bedeutet,

$$(\varepsilon, \alpha)^{m-\lambda}(\varepsilon^\lambda, \alpha) = \chi_\lambda,$$

also nach (2):

$$\sqrt[m]{\psi_\lambda} = \frac{\chi_\lambda}{\left(\sqrt[m]{\psi_1}\right)^{m-\lambda}} = \frac{\left(\sqrt[m]{\psi_1}\right)^\lambda \chi_\lambda}{\psi_1}; \tag{4}$$

und dadurch sind, wenn ε eine festgehaltene primitive $m$te Einheitswurzel bedeutet, die sämtlichen in (3) vorkommenden Radikale rational durch eines von ihnen, $\sqrt[m]{\psi_1}$, ausgedrückt.

Gibt man diesem einen seiner m verschiedenen Werte, so erhält man aus (3) gerade die m verschiedenen Werte α.

Es ist nur ein Ausnahmefall, in dem dieses Verfahren nicht anwendbar ist, das ist der, wenn $\psi_1 = 0$ ist. Wir können aber durch eine kleine Modifikation des Verfahrens uns von einem solchen Ausnahmefall frei machen. Dem schicken wir folgendes voraus.

Es sei p eine in m aufgehende Primzahl und $m = pn$; wie oben sei ε irgend eine festgehaltene primitive $m$te Einheitswurzel. Dann gibt es immer ein durch p nicht teilbares λ, so daß $(\varepsilon^\lambda, \alpha)$ von Null verschieden ist. Denn bilden wir nach der Formel § 64, (4) die Differenz $\alpha_n - \alpha_0$, so erhalten wir:

$$m(\alpha_n - \alpha_0) = \sum_{0, m-1}^{\lambda} (\varepsilon^{-n\lambda} - 1)(\varepsilon^\lambda, \alpha).$$

Nun ist aber $\varepsilon^{-n\lambda} - 1$ immer $= 0$, so oft λ durch p teilbar ist, und wenn $(\varepsilon^\lambda, \alpha)$ in allen anderen Fällen, wo also λ nicht durch p teilbar ist, verschwindet, so ist $\alpha_n = \alpha_0$, gegen die Voraussetzung, daß die α alle verschieden sein sollen. Es gibt also wenigstens ein durch p nicht teilbares λ, so daß $(\varepsilon^\lambda, \alpha)$ von Null verschieden ist.

Nun zerlegen wir m in seine Primfaktoren und setzen

$$m = p_1 p_2 \ldots,$$

worin p_1, p_2 ... Potenzen von verschiedenen Primzahlen sind. Wir setzen noch

$$m = p_1 m_1 = p_2 m_2 = \ldots,$$

und wählen, was nach dem soeben Bewiesenen stets möglich ist, λ_1 relativ prim zu p_1, λ_2 relativ prim zu p_2 usw., so daß

$$(\varepsilon^{\lambda_1}, \alpha),\ (\varepsilon^{\lambda_2}, \alpha) \ldots$$

von Null verschieden sind. Dann ist nach dem Theorem § 64, XII.

$$(\varepsilon^{\lambda}, \alpha)\,(\varepsilon^{\lambda_1}, \alpha)^{m_1\nu}\,(\varepsilon^{\lambda_2}, \alpha)^{m_2\nu} \cdots = \chi_\lambda \tag{5}$$

eine in $\Omega(\varepsilon)$ enthaltene Größe, wenn

$$\lambda \equiv -\nu(\lambda_1 m_1 + \lambda_2 m_2 + \cdots) \pmod{m}. \tag{6}$$

Es ist aber $(\varepsilon^{\lambda_1}, \alpha)^{m_1}$ eine Wurzel $p_1$ten Grades einer Funktion φ_1 in $\Omega(\varepsilon)$, und wir setzen also:

$$(\varepsilon^{\lambda_1}, \alpha)^{m_1} = \sqrt[p_1]{\varphi_1},\quad (\varepsilon^{\lambda_2}, \alpha)^{m_2} = \sqrt[p_2]{\varphi_2} \ldots,\quad (\varepsilon^{\lambda}, \alpha) = \sqrt[m]{\psi_\lambda}. \tag{7}$$

Dann wird nach (5) und 2:

$$\sqrt[m]{\psi_\lambda} = \frac{\chi_\lambda}{\left(\sqrt[p_1]{\varphi_1}\,\sqrt[p_2]{\varphi_2} \ldots\right)^{\nu}}. \tag{8}$$

Nun ist $\lambda_1 m_1 + \lambda_2 m_2 + \cdots$ relativ prim zu m, da $m_2 \ldots$ durch p_1 teilbar, $\lambda_1 m_1$ zu p_1 relativ prim ist, und also erhält man aus (6) für jedes λ eine nach dem Modul m völlig bestimmte Zahl ν.

Wenn wir also die Ausdrücke (8) in (3) einsetzen und den Radikalen $\sqrt[p_1]{\varphi_1}$, $\sqrt[p_2]{\varphi_2} \ldots$ alle ihre Werte beilegen, so erhalten wir für α genau m verschiedene Werte und nicht mehr.

Die letzten Resultate können wir benutzen, um eine Form der Darstellung der Wurzeln α in etwas verallgemeinerter Gestalt abzuleiten, die Abel an der angeführten Stelle mitteilt, und die sich auf den Fall bezieht, wo der Körper Ω reell ist, d. h. aus lauter reellen Zahlen besteht. Die Funktionen φ, ψ, χ, wie wir sie oben benutzt haben, sind dann zusammengesetzt aus reellen Zahlen und aus der Einheitswurzel ε, die man durch die Teilung der Kreisperipherie in m gleiche Teile findet; man kann etwa

$$\varepsilon = e^{\frac{2\pi i}{m}} = \cos\frac{2\pi}{m} + i \sin\frac{2\pi}{m}$$

setzen.

Die Funktion φ_1 geht, wenn ε in ε^{-1} verwandelt wird, in den konjugiert imaginären Wert über, den wir mit φ_1' bezeichnen.

Wir wollen eine positive Größe ϱ_1 und einen Winkel Θ_1 so annehmen, daß

(9) $$\varphi_1 = \varrho_1 e^{i\Theta_1}, \qquad \varphi_1' = \varrho_1 e^{-i\Theta_1},$$

oder

(10) $$\begin{aligned} \varphi_1 &= \varrho_1(\cos\Theta_1 + i\sin\Theta_1) \\ \varphi_1' &= \varrho_1(\cos\Theta_1 - i\sin\Theta_1), \end{aligned}$$

woraus noch folgt:

(11) $$\varrho_1^2 = \varphi_1\varphi_1'.$$

Nun ist

(12) $$(\varepsilon^{\lambda_1}, \alpha)(\varepsilon^{-\lambda_1}, \alpha) = \pm a_1$$

eine Größe des Körpers $\Omega(\varepsilon)$ (nach § 64, XII.), und zwar ist es, da sie sich beim Übergang zum konjugiert imaginären Wert, d. h. bei der Vertauschung von ε mit ε^{-1} nicht ändert, eine reelle Größe. Das Vorzeichen wollen wir so bestimmen, daß a_1 positiv ist. Es ist also nach (11) und (7), da ϱ_1^2 positiv ist,

$$\varrho_1^2 = (\pm a_1)^m = a_1^m,$$

und es ergibt sich daraus, daß bei ungeradem m jedenfalls das obere Zeichen gilt; bei geradem m kann auch das untere eintreten. Es ist also

(13) $$\varrho_1 = \sqrt[m]{a_1^m},$$

wo die Quadratwurzel positiv zu nehmen ist.

Ferner sind

(14) $$\frac{\varphi_1 + \varphi_1'}{2} = b_1, \qquad \frac{\varphi_1 - \varphi_1'}{2i} = c_1$$

reelle Größen in $\Omega(\varepsilon)$, und es ergibt sich aus (10):

(15) $$\cos\Theta_1 = \frac{b_1}{\sqrt[m]{a_1^m}}, \qquad \sin\Theta_1 = \frac{c_1}{\sqrt[m]{a_1^m}},$$

woraus noch die Relation folgt:

$$a_1^m = b_1^2 + c_1^2.$$

Demnach ergibt sich:

$$\sqrt[p_1]{\varphi_1} = \sqrt[m_1]{a_1}\, e^{\frac{i\Theta_1}{p_1}};$$

und nun verfährt man mit den Funktionen φ_2 usf. ebenso. Man bestimmt also $\Theta_2 \ldots$ aus einem System von Gleichungen wie (15) und erhält, wenn man noch

(16) $$m_1\Theta_1 + m_2\Theta_2 + \cdots = \Theta$$

setzt, nach (8):

(17) $$\sqrt[m]{\psi_\lambda} = (\sqrt{a_1^{m_1} a_2^{m_2} \ldots})^{-\nu}\, e^{-\frac{i\Theta}{m}\nu}\, \chi_\lambda.$$

Nach (15) können wir setzen:

$$e^{i\vartheta_1} = \frac{b_1 + ic_1}{\sqrt{a_1^m}}, \quad e^{i\vartheta_2} = \frac{b_2 + ic_2}{\sqrt{a_2^m}}, \cdots,$$

und wenn wir also durch Zerlegung in den reellen und imaginären Bestandteil

$$(18) \qquad (b_1 + ic_1)^{m_1}\,(b_2 + ic_2)^{m_2} \cdots = B + iC$$

erhalten und

$$(19) \qquad a_1^{m_1} a_2^{m_2} \cdots = A$$

setzen, so folgt nach (16):

$$e^{i\Theta} = \frac{B + iC}{\sqrt{A^m}},$$

also

$$(20) \qquad \sqrt{A^m}\cos\Theta = B, \qquad \sqrt{A^m}\sin\Theta = C,$$

und

$$(21) \qquad \sqrt[m]{\psi_\lambda} = \sqrt{A}^{-\nu}\left(\cos\frac{\Theta\nu}{m} - i\sin\frac{\Theta\nu}{m}\right)\chi_\lambda.$$

Der Winkel Θ ist durch (20) nur bis auf ein Vielfaches von 2π bestimmt, und wenn man also $\Theta + 2h\pi$ für Θ setzt und h von 0 bis $m - 1$ gehen läßt, so erhält man aus (21) die m verschiedenen Werte der Wurzelgröße. Die Funktionen $\cos\frac{\Theta\nu}{m}$ und $\sin\frac{\Theta\nu}{m}$ können noch rational durch $\cos\frac{\Theta}{m}$, $\sin\frac{\Theta}{m}$ ausgedrückt werden. Die Auflösung der zyklischen Gleichungen in dem reellen Körper Ω ist also auf folgendes zurückgeführt.

Man adjungiert zunächst dem Körper Ω die $m$te Einheitswurzel ε (Teilung der Kreisperipherie in m gleiche Teile). Hierauf sind A, B, C bekannt. Man adjungiert ferner die positive Quadratwurzel $\sqrt{A}$, dann sind $\cos\Theta$ und $\sin\Theta$ durch (20) bekannt. Endlich adjungiert man $\cos\frac{\Theta}{m}$, $\sin\frac{\Theta}{m}$ (Teilung des Winkels Θ in m Teile). Dann ist die zyklische Gleichung durch (21) gelöst.

Diese Betrachtungen führen noch zu einem interessanten Resultat über die Realitätsverhältnisse der Wurzeln zyklischer Gleichungen.

Stellen wir in der Reihe der Wurzeln $\alpha, \alpha_1, \ldots \alpha_{m-1}$ jede rational durch die vorangehende dar:

$$\alpha_1 = \Theta(\alpha),\ \alpha_2 = \Theta(\alpha_1) \ldots \alpha_{m-1} = \Theta(\alpha_{m-2}),\ \alpha = \Theta(\alpha_{m-1}),$$

worin, da hier Ω reell vorausgesetzt ist, $\Theta(x)$ eine reelle rationale Funktion von x bedeutet, so folgt zunächst, daß, wenn eine der Wurzeln reell ist, auch alle übrigen reell sein müssen. Dies findet immer bei ungeradem m statt, da eine reelle Gleichung ungeraden Grades immer wenigstens eine reelle Wurzel haben muß.

Bei geradem m können auch imaginäre Wurzeln vorhanden sein, und wenn eine Wurzel imaginär ist, so müssen es alle sein, da, wenn eine reelle Wurzel vorkommt, alle anderen auch reell sind. Bezeichnen wir mit $\Theta^\nu(x)$ die νmalige Wiederholung der Funktion Θ, so ist

$$\alpha_{\nu+k} = \Theta^\nu(\alpha_k).$$

Ist also $\alpha = \alpha_0$ mit α_k konjugiert imaginär, so sind auch für jedes ν die Funktionen $\Theta^\nu(\alpha_0)$ und $\Theta^\nu(\alpha_k)$, d. h. α_ν und $\alpha_{\nu+k}$ konjugiert imaginär.

Daraus folgt, daß $2k = m$ sein muß, und wir schließen, daß im Falle imaginärer Wurzeln

$$\alpha_k \text{ und } \alpha_{k+\frac{m}{2}}$$

für jeden Index k ein Paar konjugiert imaginärer Wurzeln bilden.

Die kubischen Gleichungen werden durch Adjunktion der Quadratwurzel aus der Diskriminante zyklische Gleichungen. Wenn die Diskriminante positiv ist, so sind die Wurzeln in Übereinstimmung mit diesem Satze reell.

§ 66.

Teilung des Winkels.

Zu den zyklischen Gleichungen gehören die Gleichungen, von denen die Teilung eines Winkels in m gleiche Teile abhängt, auf die wir die allgemeinen zyklischen Gleichungen in einem reellen Körper zurückgeführt haben.

Die Aufgabe kann so formuliert werden:

Wenn $\cos m\varphi$ und $\sin m\varphi$ gegeben sind, so sollen daraus $\cos\varphi$ und $\sin\varphi$ gefunden werden.

Die Werte von $\cos m\varphi$ und $\sin m\varphi$ denken wir uns kleiner als 1 und so, daß ihre Quadratsumme $= 1$ ist, gegeben.

Wir adjungieren noch $m$te Einheitswurzeln, mit denen wir uns im nächsten Abschnitte eingehender beschäftigen werden, die aber jedenfalls von Gleichungen abhängen, deren Grad niedriger als m ist. Um aber den Körper reell zu behalten, wollen wir nicht die Einheitswurzeln selbst, sondern

$$\sin\frac{2\pi}{m}, \qquad \cos\frac{2\pi}{m} \tag{1}$$

adjungieren.

Es möge also der Körper Ω aus allen rationalen Zahlen und aus den rationalen Funktionen von $\cos m\varphi$, $\sin m\varphi$, $\cos\frac{2\pi}{m}$, $\sin\frac{2\pi}{m}$ bestehen.

Die Gleichung $m$ten Grades, von der $x = 2\cos\varphi$ abhängt, läßt sich auf die Form bringen:

$$2\cos m\varphi = A_m(x), \tag{2}$$

worin $A_m(x)$ eine ganze Funktion $m$ten Grades von x ist, und es ist ferner

$$\sin\varphi = \frac{\sin m\varphi}{B_m(x)}, \tag{3}$$

wenn $B_m(x)$ wieder eine rationale Funktion von x ist, durch die $\sin\varphi$ rational durch x ausgedrückt ist. In diesen Formeln sind die Einheitswurzeln noch nicht enthalten.

Die m Wurzeln von (2) haben nun folgende Bedeutung:

$$x_0 = 2\cos\varphi, \quad x_1 = 2\cos\left(\varphi + \frac{2\pi}{m}\right)\ldots,$$

$$x_{m-1} = 2\cos\left(\varphi + \frac{2(m-1)\pi}{m}\right),$$

und mit Hilfe von (2) und (3) und unter Adjunktion der Größen (1) kann jede von ihnen als rationale Funktion einer anderen dargestellt werden, und zwar so:

$$x_1 = f(x_0),\ x_2 = f(x_1),\ \ldots\ x_0 = f(x_{m-1}).$$

Irgend eine Funktion der $x_0, x_1, \ldots x_{m-1}$ kann also dargestellt werden als rationale Funktion $F(x_0)$, und bei einer zyklischen Permutation geht

$$F(x_0) \text{ in } F(x_1),\ F(x_1) \text{ in } F(x_2),\ \ldots\ F(x_{m-1}) \text{ in } F(x_0)$$

über. Für eine zyklische Funktion ist also

$$F(x_0) = F(x_1) = \cdots = F(x_{m-1})$$
$$= \frac{1}{m}[F(x_0) + F(x_1) + \cdots + F(x_{m-1})],$$

und es ist folglich $F(x_0)$ rational, da es als symmetrische Funktion der Wurzeln dargestellt ist.

Die Teilung des Winkels hängt also von einer zyklischen Gleichung ab.

Diese zyklische Gleichung ist irreduzibel; denn ersetzt man in irgend einer rationalen Gleichung $\Phi(x_0, \cos m\varphi, \sin m\varphi) = 0$ den Winkel φ durch $\varphi + \frac{2\pi k}{m}$, so geht sie in

$$\Phi(x_k, \cos m\varphi, \sin m\varphi) = 0$$

über und ist also für alle Wurzeln von (2) befriedigt.

Wollte man nur $\cos m\varphi$ adjungieren, nicht zugleich $\sin m\varphi$, dann würde die Gleichung keine zyklische mehr sein, was man leicht an dem Beispiel der Dreiteilung bestätigt, wo eben $\sin m\varphi$ die Quadratwurzel aus der Diskriminante wird.

Elfter Abschnitt.

Kreisteilung.

§ 67.

Die Einheitswurzeln.

Wenn die $n$te Potenz einer Zahl r gleich 1 ist, wenn also die Gleichung

$$r^n = 1 \tag{1}$$

für ein ganzes positives n befriedigt ist, so heißt r eine $n$te Einheitswurzel oder eine Einheitswurzel vom Grade n. Es sind also alle $n$ten Einheitswurzeln, und nur diese, Wurzeln der Gleichung $n$ten Grades:

$$f(x) = x^n - 1 = 0. \tag{2}$$

Es ist

$$f'(x) = n x^{n-1}, \tag{3}$$

und folglich hat $f(x)$ mit $f'(x)$ keinen Teiler gemein; also hat $f(x)$ keine mehrfachen Wurzeln, und es gibt nach dem Fundamentalsatz der Algebra n und nicht mehr voneinander verschiedene $n$te Einheitswurzeln.

Ist r eine $n$te Einheitswurzel, so ist es auch jede ganze Potenz von r, denn aus $r^n = 1$ folgt $r^{kn} = 1$, wenn k eine beliebige positive oder negative ganze Zahl ist (auch $k = 0$ nicht ausgeschlossen); also ist auch r^k eine $n$te Einheitswurzel. Da es aber nur n Einheitswurzeln vom Grade n gibt, so sind die Potenzen r^k nicht alle voneinander verschieden. Hierüber gilt nun folgendes:

Wenn sich zwei Zahlen k, k' um ein Vielfaches von n unterscheiden, wenn also

$$k' \equiv k \pmod{n} \tag{4}$$

ist, so ist auch

$$r^k = r^{k'}. \tag{5}$$

Denn ist $k' = k + hn$, so ist

$$r^{k'} = r^k r^{hn},$$

woraus, da $r^n = 1$ ist, die Gleichung (5) folgt.

Es sind also in der Reihe der Zahlen

$$1,\ r,\ r^2,\ r^3 \ldots r^{n-1} \tag{6}$$

gewiß alle voneinander verschiedenen r^k enthalten; aber es müssen nicht umgekehrt die Größen (6) alle voneinander verschieden sein.

Nehmen wir an, es seien k und $k' = k + \mu$ zwei Zahlen der Reihe

$$0,\ 1,\ 2, \ldots n - 1,$$

und

$$r^k = r^{k+\mu},$$

so folgt, daß $r^\mu = 1$ sein muß. Es kann also in der Reihe (6) kein früher dagewesenes Glied wiederkehren, ehe das erste Glied 1 zum zweiten Male vorkommt, und wenn μ die kleinste positive Zahl ist, für die $r^\mu = 1$ ist, so sind die Zahlen

$$1,\ r,\ r^2 \ldots r^{\mu-1} \tag{7}$$

alle voneinander verschieden.

Es muß dann μ ein Teiler von n sein. Denn durch Division lassen sich die ganzen Zahlen h, μ' so bestimmen, daß

$$n = h\mu + \mu'; \qquad 0 \leqq \mu' < \mu.$$

Dann ist aber auch, wie aus

$$r^n = r^{h\mu} r^{\mu'}$$

hervorgeht, $r^{\mu'} = 1$, d. h. da μ die kleinste positive Zahl sein soll, für die $r^\mu = 1$ ist, $\mu' = 0$, und folglich n durch μ teilbar.

Es ist also r zugleich μte Einheitswurzel, aber nicht Einheitswurzel von noch niedrigerem Grade.

Die $n$ten Einheitswurzeln r, die nicht zugleich Einheitswurzeln eines niedrigeren Grades sind, haben wir bereits in § 64 primitive $n$te Einheitswurzeln genannt.

Aus dieser Definition folgt, daß die Zahlen der Reihe (6), wenn r eine primitive $n$te Einheitswurzel ist, alle voneinander verschieden sind, und daß sämtliche $n$ten Einheitswurzeln darunter enthalten sind, daß sie aber nur einen Teil der $n$ten Einheitswurzeln ausmachen, wenn r eine imprimitive $n$te Einheitswurzel ist.

Jede Einheitswurzel, deren Grad ein von n verschiedener Teiler von n ist, ist zugleich imprimitive $n$te Einheitswurzel.

Ist r zugleich $n$te und $m$te Einheitswurzel, so ist es auch μte Einheitswurzel, wenn μ der größte gemeinschaftliche Teiler von n und m ist.

Denn die ganzen Zahlen x, y lassen sich aus der diophantischen Gleichung

$$mx + ny = \mu$$

bestimmen, und folglich ist

$$r^{\mu} = r^{mx} r^{ny} = 1.$$

Dem entspricht der andere Satz:

Sind $r_1, r_2 \ldots$ Einheitswurzeln der Grade $n_1, n_2 \ldots$, so sind sie alle zugleich Einheitswurzeln des Grades m, wenn m irgend ein gemeinschaftliches Vielfaches von $n_1, n_2 \ldots$ bedeutet.

Wir haben noch zu untersuchen, ob für jeden Grad n primitive Einheitswurzeln existieren und ihre Anzahl festzustellen.

Sei der Grad n in zwei Faktoren a, b zerlegt, die zueinander relativ prim sind, also

$$n = ab,$$

und sei α eine $a$te, β eine $b$te Einheitswurzel. Dann ist das Produkt

$$r = \alpha\beta \tag{8}$$

eine $n$te Einheitswurzel. Sind α', β' zwei andere $a$te und $b$te Einheitswurzeln, so ist $r' = \alpha'\beta'$ auch eine $n$te Einheitswurzel, und es ist zu zeigen, daß r' von r verschieden ist, wenn nicht gleichzeitig $\alpha = \alpha'$ und $\beta = \beta'$ ist.

Da nämlich a, b relativ prim sind, so lassen sich die ganzen Zahlen x, y durch die Diophantische Gleichung

$$ax + by = 1$$

bestimmen, und dann folgt aus (8) wegen $\alpha^a = 1$, $\beta^b = 1$

$$\alpha = r^{by}, \qquad \beta = r^{ax}.$$

Demnach ist α und β durch r vollständig bestimmt, und wenn $\alpha'\beta'$ auch gleich r sein soll, so muß $\alpha = \alpha'$; $\beta = \beta'$ sein.

Läßt man also α alle $a$ten, β alle $b$ten Einheitswurzeln durchlaufen, so erhält $r = \alpha\beta$ genau $ab = n$ verschiedene Werte, und es folgt:

I. In der Form $\alpha\beta$ sind alle $n$ten Einheitswurzeln darstellbar,

und weiter:

II. r ist dann und nur dann eine primitive $n$te Einheitswurzel, wenn α eine primitive $a$te und β eine primitive $b$te Einheitswurzel ist.

Denn erstens sei μ der kleinste positive Exponent, für den $r^\mu = 1$ ist; dann ist auch $\alpha^\mu \beta^\mu = 1$ und daraus folgt, wenn man beiderseits zur Potenz $by = 1 - ax$ und $ax = 1 - by$ erhebt,

$$\alpha^\mu = 1, \qquad \beta^\mu = 1.$$

Wenn nun $\mu < n$ ist, so kann es nicht zugleich durch a und durch b teilbar sein, und also können auch a und b nicht beide die kleinsten positiven Exponenten der Potenzen von α und β sein, die gleich 1 werden, d. h. also, wenn r nicht primitive $n$te Einheitswurzel ist, so sind auch α und β nicht zugleich primitive $a$te und $b$te Einheitswurzeln, oder wenn α und β primitive $a$te und $b$te Einheitswurzeln sind, so ist ihr Produkt r primitive $n$te Einheitswurzel. Auf der anderen Seite ist klar, daß, wenn α oder β Einheitswurzel von niedrigerem Grade als a oder b ist, auch r Einheitswurzel von niedrigerem als dem $n$ten Grade sein wird.

Zerfällt n in mehrere Faktoren $a, b, c \ldots$, von denen je zwei zueinander relativ prim sind, und sind $\alpha, \beta, \gamma \ldots$ Einheitswurzeln der Grade $a, b, c \ldots$, und setzt man

$$r = \alpha\beta\gamma \ldots, \tag{9}$$

so schließt man durch mehrmalige Anwendung der vorigen Sätze, daß in (9) alle $n$ten Einheitswurzeln, und jede nur einmal, enthalten sind, und ferner, daß r dann und nur dann primitive $n$te Einheitswurzel ist, wenn $\alpha, \beta, \gamma \ldots$ primitive Einheitswurzeln der Grade $a, b, c \ldots$ sind.

Bezeichnen wir jetzt die Anzahl der primitiven $n$ten Einheitswurzeln durch $\varphi(n)$, so folgt aus dem hier Bewiesenen

$$\varphi(n) = \varphi(a)\varphi(b)\varphi(c) \ldots, \tag{10}$$

wenn

$$n = abc \ldots,$$

und $a, b, c \ldots$ Zahlen sind, die, je zwei und zwei, zueinander relativ prim sind.

Nun kann man jede Zahl n auf eine und nur auf eine Weise in ein Produkt von Primzahlpotenzen zerlegen:

$$n = p^\pi p_1^{\pi_1} p_2^{\pi_2} \ldots,$$

worin $p, p_1, p_2 \ldots$ verschiedene Primzahlen und $\pi, \pi_1, \pi_2 \ldots$ positive Exponenten sind, so daß aus der Formel (10) folgt:

$$\varphi(n) = \varphi(p^\pi)\ \varphi(p_1^{\pi_1})\ \varphi(p_2^{\pi_2}) \ldots,$$

und daß es also nur noch darauf ankommt, zu entscheiden, ob und wie viele primitive Einheitswurzeln des Grades p^π existieren.

Diese Frage ist aber sehr einfach zu beantworten. Wenn nämlich ϱ eine Einheitswurzel vom Grade p^π ist, so ist der niedrigste Grad, zu dem ϱ als Einheitswurzel gehört, ein Teiler von p^π, also eine Potenz von p, und wenn er also nicht gleich p^π ist, ein Teiler von $p^{\pi-1}$, d. h. jede nicht primitive Einheitswurzel vom Grade p^π ist zugleich Einheitswurzel vom Grade $p^{\pi-1}$. Da es aber p^π Einheitswurzeln vom Grade p^π und nur $p^{\pi-1}$ Einheitswurzeln vom Grade $p^{\pi-1}$ gibt, so müssen

$$p^\pi - p^{\pi-1} = p^\pi \left(1 - \frac{1}{p}\right)$$

primitive Einheitswurzeln des Grades p^π vorhanden sein. Daraus erhält man nach (10) die Anzahl aller primitiven $n$ten Einheitswurzeln

$$\varphi(n) = n\,\Pi\left(1 - \frac{1}{p}\right), \tag{11}$$

worin das Produktzeichen Π sich auf alle voneinander verschiedenen in n aufgehenden Primzahlen bezieht. Nur für den Fall $n = 1$ paßt die Formel (11) nicht mehr; in diesem Falle ist $\varphi(1) = 1$ zu setzen. Die Zahl $\varphi(n)$ ist also niemals gleich Null.

Da es hiernach für jeden Grad n wenigstens eine primitive Einheitswurzel r gibt, so lassen sich alle $n$ten Einheitswurzeln durch die Potenzen von r darstellen:

$$1,\ r,\ r^2,\ r^3 \ldots r^{n-1}. \tag{12}$$

Ist r^k irgend eine Potenz von r, so wird dann und nur dann $r^{km} = 1$ sein, wenn km durch n teilbar ist. Ist also $n = n'n''$ und n' der größte gemeinsame Teiler von k und n, so muß m durch n'' teilbar sein, und n'' ist der Exponent der niedrigsten Potenz von r^k, die gleich 1 wird, d. h. r^k ist eine primitive n''te Einheitswurzel. Es folgt hieraus der Satz:

III. Ist r primitive $n$te Einheitswurzel, so ist r^k dann und nur dann primitive $n$te Einheitswurzel, wenn k relativ prim zu n ist.

Nehmen wir k aus der Reihe der Zahlen 1, 2, 3, ... n, so ergibt sich der Satz der Zahlentheorie, daß $\varphi(n)$ gleich der

Anzahl der Zahlen ist, die nicht größer als n und relativ prim zu n sind. Dies ist die ursprüngliche Definition des in der Zahlentheorie allgemein gebrauchten Zeichens $\varphi(n)$.

Ist k relativ prim zu n, so kann man eine ganze Zahl x so bestimmen, daß $kx \equiv 1 \pmod{n}$ wird. Sind dann r, r_1 zwei $n$te Einheitswurzeln, so kann nur dann $r^k = r_1^k$ sein, wenn $r = r_1$ ist, wie sich durch Erheben zur Potenz x ergibt. Daraus folgt nach III.:

IV. Ist k relativ prim zu n und durchläuft r die Reihe der primitiven $n$ten Einheitswurzeln, so durchläuft r^k dieselbe Zahlenreihe, wenn auch in anderer Ordnung.

Endlich führen wir noch den Satz an:

V. Ist n eine Primzahl, so ist jede $n$te Einheitswurzel mit Ausnahme von 1 primitive $n$te Einheitswurzel.

§ 68.

Die Kreisteilungsgleichungen.

Alle $n$ten Einheitswurzeln sind, wie wir gesehen haben, Wurzeln einer Gleichung $f_n(x) = 0$, wenn

$$f_n(x) = x^n - 1 \tag{1}$$

ist; wenn wir die Funktion $f_n(x)$ von allen Faktoren befreien die sie mit anderen Funktionen derselben Form $f_{n_1}(x)$ gemein hat, was durch rationale Operationen geschieht, so erhalten wir eine Gleichung $X_n = 0$, der die primitiven $n$ten Einheitswurzeln und nur diese genügen, und X_n hat die Form

$$X_n = x^\nu + a_1 x^{\nu-1} + \cdots + a_\nu. \tag{2}$$

Der Grad ν ist gleich $\varphi(n)$, und die Koeffizienten $a_1, a_2, \ldots a_\nu$ sind rationale Zahlen.

Beim Aufsuchen der gemeinschaftlichen Faktoren von f_n und f_{n_1} können wir uns für n_1 auf die Teiler von n beschränken. Wie man die Funktion X_n einfach bilden kann, werden wir gleich noch näher sehen. Wir beweisen aber zunächst einen allgemeinen Satz über diese Funktionen.

Die $n$ten Einheitswurzeln umfassen alle primitiven μten Einheitswurzeln, worin μ irgend ein Teiler von n ist, n selbst und 1 eingeschlossen, da als primitive erste Einheitswurzel eben die Einheit 1 selbst zu betrachten ist. Lassen wir also μ alle Divi-

soren von n durchlaufen und beachten, daß zwei verschiedene X_μ niemals einen gemeinschaftlichen Teiler haben, und daß sowohl $f_n(x)$ als X_μ keine mehrfachen Faktoren enthalten, so folgt:

$$f_n(x) = \Pi X_\mu, \tag{3}$$

worin sich das Produktzeichen Π auf alle Teiler μ von n bezieht.

Daraus erhalten wir nebenbei einen Beweis des zahlentheoretischen Satzes:

$$n = \sum^{\mu} \varphi(\mu), \tag{4}$$

wenn wir den Grad der Funktionen auf der rechten und der linken Seite von (3) einander gleich setzen.

Andererseits schließen wir nach dem Gaußschen Theorem (§ 20, 8.), daß die Koeffizienten der sämtlichen X_μ ganze Zahlen sind. Denn kämen darunter auch gebrochene Zahlen vor, so könnte das Produkt nicht lauter ganzzahlige Koeffizienten enthalten, wie es doch nach (3) und (1) sein muß.

Alle Funktionen $f_n(x)$ haben den Teiler $x - 1$, und wenn wir die Teilung ausführen, so ergibt sich:

$$\frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1. \tag{5}$$

Hieraus schließen wir, daß, wenn r irgend eine primitive oder nicht primitive $n$te Einheitswurzel ist, mit alleiniger Ausnahme von $r = 1$, immer

$$1 + r + r^2 + \cdots + r^{n-1} = 0, \tag{6}$$

während für $r = 1$ die Summe auf der linken Seite von (6) offenbar den Wert n hat [§ 64, (1)].

Wenn n eine Primzahl ist, so gibt es außer 1 keine imprimitiven $n$ten Einheitswurzeln, und daher ist, wenn n eine Primzahl ist, was wir dadurch andeuten wollen, daß wir p dafür setzen,

$$X_p = x^{p-1} + x^{p-2} + \cdots + x + 1. \tag{7}$$

Ebenso einfach läßt sich X_n bilden, wenn n eine Primzahlpotenz ist, also

$$n = p^\pi,$$

oder, wenn wir zur Abkürzung $p^{\pi-1} = p'$ setzen,

$$n = p'p.$$

Die $n$ten Einheitswurzeln bestehen in diesem Falle aus den primitiven $n$ten Einheitswurzeln und aus den p'ten Einheitswurzeln, und es ist also

$$(8)\quad X_n = \frac{x^{pp'} - 1}{x^{p'} - 1} = x^{p'(p-1)} + x^{p'(p-2)} + \cdots + x^{p'} + 1.$$

Ist $p' > 1$, also $\pi > 1$, so fehlt in dieser Gleichung νten Grades das Glied mit der $(\nu - 1)$ten Potenz der Unbekannten, dessen Koeffizient der negativen Summe der Wurzeln gleich ist. Wir haben also den Satz:

VI. Die Summe aller primitiven $n$ten Einheitswurzeln ist, wenn n eine höhere Potenz einer Primzahl ist, immer gleich Null.

Zur allgemeinen Bildung von X_n wollen wir ein rekurrentes Verfahren anwenden; wir nehmen X_n als schon gebildet an, bezeichnen mit p eine in n nicht aufgehende Primzahl, mit p' wie oben die $(\pi - 1)$te Potenz von p, und bilden nun $X_{np'p}$, worin natürlich p' auch gleich 1 sein kann.

Bezeichnen wir mit r die primitiven $n$ten Einheitswurzeln, mit α jede Einheitswurzel des Grades pp', und mit α' jede Einheitswurzel des Grades p', so erhält man die sämtlichen primitiven Einheitswurzeln des Grades npp', wenn man von den sämtlichen $r\alpha$ die $r\alpha'$ wegnimmt. Die $r\alpha$ sind aber die Wurzeln der Gleichung:

$$X_n(x^{pp'}) = 0,$$

weil die Größen

$$(r\alpha)^{pp'} = r^{pp'},$$

von der Reihenfolge abgesehen, mit den r selbst übereinstimmen (§ 67, IV.). Ebenso sind die $r\alpha'$ die Wurzeln der Gleichung

$$X_n(x^{p'}) = 0,$$

und daraus ergibt sich:

$$(9)\quad X_{npp'}(x) = \frac{X_n(x^{pp'})}{X_n(x^{p'})}.$$

Hiervon macht auch der Wert $n = 1$ keine Ausnahme, wenn wir unter X_1 die Funktion $x - 1$ verstehen. Danach können wir X_n in allen Fällen verhältnismäßig einfach bilden. Wir betrachten einige besondere Fälle.

Nehmen wir an, es enthalte n nur zwei voneinander verschiedene Primzahlen p, q und sei also

$$n = pp'qq',$$

dann ergibt die Formel (8) und (9):

$$X_n = \frac{(x^n - 1)\left(x^{\frac{n}{pq}} - 1\right)}{\left(x^{\frac{n}{p}} - 1\right)\left(x^{\frac{n}{q}} - 1\right)},$$

eine Formel, die sich leicht durch vollständige Induktion folgendermaßen verallgemeinern läßt.

Bezeichnet man mit μ_1 alle Zahlen, die aus n entstehen, wenn man n durch eine gerade Zahl verschiedener Primteiler von n dividiert (n selbst eingeschlossen), mit μ_2 die Zahlen, die aus n entstehen, wenn man n durch eine ungerade Zahl solcher Primteiler dividiert, so ist

$$X_n = \frac{\prod^{\mu_1}(x^{u_1} - 1)}{\prod^{\mu_2}(x^{u_2} - 1)}. \tag{10}$$

Der Beweis ergibt sich aus (9) für npp', wenn man annimmt, die Richtigkeit sei für n schon bewiesen.

Bedeutet r jede der primitiven $n$ten Einheitswurzeln, so sind bei ungeradem n die $-r$ die primitiven $2n$ten Einheitswurzeln (§ 67, I., II.). Demnach ist bei ungeradem n

$$X_{2n}(x) = X_n(-x). \tag{11}$$

Ist n eine Potenz von 2, so ist nach (8):

$$X_n = \frac{x^n - 1}{x^{\frac{n}{2}} - 1} = x^{\frac{n}{2}} + 1. \tag{12}$$

Setzen wir in der Formel (8) $x = 1$, so erhält X_n den Wert p, wenn n eine Potenz von p ist. Dagegen ergibt die Formel (9), wenn $n > 1$ ist, für $X_{npp'}$ den Wert 1 (für $n = 1$ würde auf der rechten Seite Zähler und Nenner $= 0$).

Wir erhalten also den Satz:

VII. Die Funktion X_n erhält für $x = 1$ den Wert p, wenn n eine Potenz der Primzahl p ist, und den Wert 1, wenn n mehr als eine Primzahl als Teiler hat.

Ist p eine in n aufgehende Primzahl und

$$n = pn',$$

so ist

$$\frac{x^n - 1}{x^{n'} - 1} = x^{n'(p-1)} + x^{n'(p-2)} + \cdots + x^{n'} + 1, \tag{13}$$

und diese Funktion verschwindet, wenn x gleich irgend einer primitiven $n$ten Einheitswurzel α gesetzt wird. Sie ist daher durch X_n teilbar, und es muß also eine ganze Funktion Y_n von x geben, die nach § 20, 8. ganzzahlige Koeffizienten hat, so daß

$$x^{n'(p-1)} + x^{n'(p-2)} + \cdots + x^{n'} + 1 = X_n Y_n \tag{14}$$

wird. Ist nun α' irgend eine n'te Einheitswurzel, also $\alpha'^{n'} = 1$, so folgt aus (14):

(15) $$X_n(\alpha')\, Y_n(\alpha') = p,$$

also eine Zerlegung der Primzahl p in zwei Faktoren, die ganze Funktionen einer Einheitswurzel von beliebigem Grade n' sind.

§ 69.

Die Diskriminante der Kreisteilungsgleichung.

Die Einheitswurzeln vom Grade n lassen sich in transzendenter Form darstellen durch

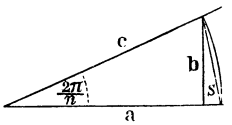
(1) $$r = e^{\frac{2\pi i k}{n}} = \cos\frac{2\pi k}{n} + i \sin\frac{2\pi k}{n}.$$

Diese Einheitswurzel ist primitiv oder nicht primitiv, je nachdem k teilerfremd zu n ist oder nicht. Die einfachste unter den primitiven erhält man für $k = 1$, nämlich

(2) $$r_0 = e^{\frac{2\pi i}{n}} = \cos\frac{2\pi}{n} + i \sin\frac{2\pi}{n} = a + b\,i.$$

Der Winkel $2\pi : n$ ist der $n$te Teil der ganzen Kreisperipherie. Die verschiedenen Teilpunkte, die man erhält, wenn man diesen Winkel von einem beliebigen Anfangspunkte an auf der Peripherie aufträgt, bestimmen das dem Kreise eingeschriebene reguläre n-Eck.

Fig. 9.



Die Teilpunkte lassen sich konstruieren, wenn man r und damit a und b (oder auch nur eine dieser beiden Größen) kennt. Insbesondere ergibt sich für die Seite s des regulären n-Eckes, wenn der Radius c des umgeschriebenen Kreises gleich 1 angenommen wird,

$$s = 2 \sin\frac{\pi}{n} = \sqrt{2\left(1 - \cos\frac{2\pi}{n}\right)}.$$

Die übrigen Größen r stehen in derselben Beziehung zu den anderen Teilpunkten. Wegen dieser geometrischen Bedeutung heißt die Gleichung $X_n = 0$, deren Wurzeln die primitiven unter den Größen (1) sind, die Kreisteilungsgleichung.

Wir bestimmen jetzt die Diskriminante der Kreisteilungsgleichung $X_n = 0$, beschränken uns aber auf den Fall, daß n eine Potenz einer Primzahl q sei:

(3) $$n = q^k.$$

Die Anzahl der primitiven $n$ten Einheitswurzeln r ist dann

$$\mu = q^{k-1}(q-1). \tag{4}$$

Also immer eine gerade Zahl (außer für $n = 2$, was wir ohnehin ausschließen). Die Funktion, deren Wurzeln die r sind, ist

$$X_n = f(x) = \frac{x^n - 1}{x^{n/q} - 1} \tag{5}$$

und daraus

$$f'(r) = \frac{n r^{n-1}}{r^{n/q} - 1} \tag{6}$$

und die Diskriminante ist nach § 26 (weil $\mu - 1$ ungerade ist):

$$D = (-1)^{\frac{\mu}{2}} \prod^{r} f'(r).$$

Das Produkt der Zähler von (6) ist, da das Produkt aller μ Größen r gleich 1 ist, gleich n^μ. Im Nenner von (6) kommt jede primitive $q$te Einheitswurzel $n : q$mal vor und folglich ist das Produkt der Nenner nach § 68, VII. gleich $q^{n/q}$, und daraus ergibt sich die Diskriminante:

$$D = (-1)^{\frac{\mu}{2}} q^{k\mu - \frac{n}{q}} = (-1)^{\frac{\mu}{2}} q^{q^{k-1}[k(q-1)-1]} \tag{7}$$

und dies gibt für $k = 1$, also für eine ungerade Primzahl:

$$D = (-1)^{\frac{q-1}{2}} q^{q-2}. \tag{8}$$

In diesem Fall $n = q$ sind die Wurzeln von X_n, wenn r eine von ihnen ist, alle in der Form enthalten

$$r, r^2, r^3 \ldots r^{n-1},$$

und wenn wir also das Differenzenprodukt

$$\begin{aligned} P = (r - r^2)\,(r - r^3) \ldots (r - r^{n-1}) \\ (r^2 - r^3) \ldots (r^2 - r^{n-1}) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ (r^{n-2} - r^{n-1}) \end{aligned} \tag{9}$$

einführen, so ist (§ 26)

$$D = P^2.$$

Es ist also auch nach (8)

$$P = n^{\frac{n-3}{2}} \sqrt{(-1)^{\frac{n-1}{2}} n}, \tag{10}$$

wodurch P, abgesehen vom Vorzeichen, bestimmt ist. Es ist P reell, wenn $n \equiv 1$, und imaginär, wenn $n \equiv -1 \pmod 4$ ist also reell für $n = 5, 13, 17, 29, 37 \ldots$, imaginär für $n = 3, 7$ 11, 19, 23, 31 ... Das Vorzeichen, das wir in (10) der Wurzel

zu geben haben, hängt davon ab, welches r wir in dem Ausdruck (9) gewählt haben.

Wählen wir ein bestimmtes r, z. B. $r = r_0$, so können wir das Vorzeichen in (10) noch bestimmen. Um dies auszuführen, teilen wir die binomischen Faktoren von P,

$$r^\nu - r^\mu, \qquad \nu < \mu,$$

deren Anzahl $\frac{1}{2}(n-1)(n-2)$ beträgt, in zwei Klassen. Die Differenzen der einen Klasse bilden das Produkt

$$Q = (r - r^{n-1})(r^2 - r^{n-2}) \dots \left(r^{\frac{n-1}{2}} - r^{\frac{n+1}{2}}\right), \tag{11}$$

das also alle die Faktoren enthält, in denen $\mu + \nu = n$ ist.

Die übrigen Faktoren lassen sich in Paaren von folgender Form zusammenfassen:

$$R = (r^\nu - r^\mu)(r^{n-\mu} - r^{n-\nu}). \tag{12}$$

Die Anzahl der Faktoren in Q ist $\frac{1}{2}(n-1)$, und also ist die Anzahl der Paare R

$$\frac{1}{2}\left(\frac{(n-1)(n-2)}{2} - \frac{n-1}{2}\right) = \frac{(n-1)(n-3)}{4},$$

und diese Zahl ist immer gerade, da einer der beiden Faktoren $n-1$, $n-3$ durch 4 teilbar ist.

Nun ist, wenn

$$r = e^{\frac{2\pi i k}{n}}$$

ist,

$$R = -2 + r^{\mu-\nu} + r^{\nu-\mu} = -2\left(1 - \cos\frac{2\pi(\mu-\nu)}{n}k\right),$$

d. h. R ist immer negativ. Folglich ist das Produkt aller R eine positive Größe, denn die Anzahl der Faktoren von R ist, wie oben gezeigt, gerade.

Es ist ferner nach (11)

$$Q = (2i)^{\frac{n-1}{2}} \sin\frac{2\pi k}{n} \sin\frac{4\pi k}{n} \cdots \sin\frac{(n-1)\pi k}{n}, \tag{13}$$

und das Vorzeichen hiervon ist von k abhängig. Nehmen wir aber $k = 1$, also $r = r_0$, so sind alle die Winkel

$$\frac{2\pi}{n}, \frac{4\pi}{n}, \dots \frac{(n-1)\pi}{n}$$

zwischen Null und π gelegen und die Sinus alle positiv. Wir schließen hieraus nach (10), daß in diesem Falle

$$P = i^{\frac{n-1}{2}} n^{\frac{n-3}{2}} \sqrt{n} \tag{14}$$

und $\sqrt{n}$ positiv zu nehmen ist.

Um Q zu bestimmen, multiplizieren wir den Ausdruck (11) mit

$$r \cdot r^2 \cdot r^3 \dots r^{\frac{n-1}{2}} = r^{\frac{n^2-1}{8}}$$

und sodann mit

$$r^{-1} \cdot r^{-2} \cdot r^{-3} \dots r^{-\frac{n-1}{2}} = r^{-\frac{n^2-1}{8}};$$

so erhalten wir:

$$r^{\frac{n^2-1}{8}} Q = (r^2 - 1)(r^4 - 1) \dots (r^{n-1} - 1)$$

$$r^{-\frac{n^2-1}{8}} Q = (1 - r^{n-2})(1 - r^{n-4}) \dots (1 - r).$$

Da nun die Exponenten $2, 4, \dots n-1, n-2, n-4, \dots 1$ zusammen alle Zahlen $1, 2 \dots n-1$ umfassen, so ergibt sich durch Multiplikation dieser beiden Ausdrücke:

$$Q^2 = (-1)^{\frac{n-1}{2}} \Pi(r-1) = (-1)^{\frac{n-1}{2}} n,$$

also

$$Q = \pm i^{\frac{n-1}{2}} \sqrt{n}. \tag{15}$$

Die Vergleichung mit (13) ergibt:

$$2^{\frac{n-1}{2}} \sin\frac{2\pi k}{n} \sin\frac{4\pi k}{n} \cdots \sin\frac{(n-1)\pi k}{n} = \pm\sqrt{n}. \tag{16}$$

Das Vorzeichen in dieser Formel hängt, wie in (15), noch von k ab; es ist aber das positive, wenn $k = 1$ ist. Im übrigen wollen wir über dies Vorzeichen, das weiterhin noch genauer untersucht werden wird, noch einen wichtigen Satz ableiten.

Nach der Definition (11) ist Q eine ganze rationale Funktion von r, die, wenn man sie nach Potenzen von r ordnet, ganze rationale Zahlenkoeffizienten erhält. Setzen wir also

$$r = r_0^k,$$

so wird

$$Q = F(r_0^k),$$

worin F das Zeichen für eine rationale Funktion ist, deren Koeffizienten von k nicht abhängig sind. Setzen wir nun für k der Reihe nach die Werte

$$k = 1, 2, \dots n-1,$$

so stellt r_0^k alle Wurzeln der Kreisteilungsgleichung $X_n = 0$ dar, und die Summe

$$\sum^{k} Q = F(r_0) + F(r_0^2) + \cdots + F(r_0^{n-1})$$

ist eine symmetrische Funktion der Wurzeln dieser Gleichung; sie läßt sich also rational durch die Koeffizienten, d. h. durch rationale Zahlen ausdrücken und ist mithin selbst eine rationale Zahl. Andererseits ist aber nach der Formel (15)

$$\sum^{k} Q = h i^{\frac{n-1}{2}} \sqrt{n},$$

wenn h eine ganze Zahl bedeutet, nämlich die Anzahl der Fälle, in denen in (15) das positive Zeichen zu nehmen ist, vermindert um die Anzahl der Fälle, in denen das negative Zeichen gilt. Beides ist aber nur dann miteinander verträglich, wenn $h = 0$ ist. Denn es kann $h i^{\frac{n-1}{2}} \sqrt{n}$ nur dann für ein rationales h rational sein, wenn $h = 0$ ist. Damit ist der folgende Satz bewiesen:

VIII. Durchläuft k die Reihe der Zahlen 1, 2, 3, ... $n-1$, so gilt in der Formel (16) ebensooft das positive wie das negative Zeichen.

Es braucht kaum besonders erwähnt zu werden, daß man für k auch ein anderes volles Restsystem von n, mit Ausschluß der durch n teilbaren Zahl, nehmen kann.

§ 70.

Primitive Kongruenzwurzeln.

Parallel mit der Theorie der Einheitswurzeln geht eine Theorie der sogenannten binomischen Kongruenzen, ohne die in der Kreisteilung weitere Schritte nicht gemacht werden können. Die Grundzüge dieser Theorie sollen daher hier eingeschoben werden.

Es sei also n eine Primzahl und

$$(1) \qquad f(x) = a_0 x^m + a_1 x^{m-1} + \cdots + a_{m-1} x + a_m$$

eine ganze Funktion von x, deren Koeffizienten a ganze Zahlen sind, und a_0 nicht durch n teilbar. Setzen wir für x eine solche ganze Zahl α, daß $f(\alpha)$ durch n teilbar wird, so sagen wir (nach Analogie der Gleichungen), α sei eine Wurzel der Kongruenz $m$ten Grades:

$$(2) \qquad f(x) \equiv 0 \pmod{n}.$$

Wird α um ein Vielfaches von n vermehrt, so bleibt es Wurzel der Kongruenz (2). Solche nach dem Modul n kongruente Wurzeln gelten nicht als verschieden. Unter dieser Voraussetzung können wir den Satz aussprechen:

IX. Eine Kongruenz $m$ten Grades für einen Primzahlmodul n kann nicht mehr als m verschiedene Wurzeln haben.

Der Satz ist richtig für $m = 1$, denn die Kongruenz $a_0 x + a_1 \equiv 0 \pmod{n}$ hat nur eine Wurzel, nämlich, wenn a_0' so bestimmt wird, daß $a_0 a_0' \equiv 1 \pmod{m}$ ist, $\alpha \equiv -a_1 a_0' \pmod{n}$. Wir nehmen unseren Satz also jetzt als bewiesen an für den Grad $m - 1$ und leiten seine Richtigkeit für den Grad m daraus her. Sind x und α zwei Variable, so erhalten wir durch Division:

$$\frac{f(x) - f(\alpha)}{x - \alpha} = f_1(x), \tag{3}$$

worin $f_1(x)$ eine ganze Funktion vom $(m-1)$ten Grade ist, die, wenn α eine ganze Zahl ist, ganzzahlige Koeffizienten hat. Wenn also $f(\alpha) \equiv 0$ ist, so folgt aus (3):

$$f(x) \equiv (x - \alpha) f_1(x) \pmod{n}. \tag{4}$$

Jede Wurzel β der Kongruenz (2) muß also der Bedingung $(\beta - \alpha) f_1(\beta) \equiv 0 \pmod{n}$ genügen. Also ist entweder $\beta - \alpha$ oder $f_1(\beta)$ durch n teilbar.

Nun gibt es nach Voraussetzung höchstens $m - 1$ Werte β, für die $f_1(\beta)$ durch n teilbar wird; gibt es also noch eine $m$te Wurzel von (4), so muß diese gleich α sein.

Wenn also eine Kongruenz von der Form $f(x) \equiv 0 \pmod{n}$ mehr Wurzeln hat, als ihr Grad beträgt, so schließen wir, daß die Kongruenz identisch ist, d. h. daß alle Koeffizienten von $f(x)$ durch n teilbar sein müssen.

Dieser Satz ist, wie man sieht, ganz analog dem algebraischen Satze, daß eine Gleichung nicht mehr Wurzeln haben kann, als ihr Grad angibt. Es läßt sich aber nicht der andere Satz übertragen, daß jede Gleichung auch wirklich so viele Wurzeln hat. Eine Kongruenz $m$ten Grades kann weniger, selbst gar keine Wurzeln haben. Um so bemerkenswerter ist eine besondere Kongruenz, bei der die Zahl der Wurzeln immer dem Grade gleichkommt, auf Grund eines Lehrsatzes, der der Fermatsche Lehrsatz genannt wird, und den wir hier so formulieren:

X Die Kongruenzen

(5) $$x^n - x \equiv 0, \quad x^{n-1} - 1 \equiv 0 \pmod{n}$$

haben, wenn n eine Primzahl ist, so viele Wurzeln, als ihr Grad beträgt, nämlich n und $n - 1$.

Beide Behauptungen sind nicht wesentlich verschieden, denn die erste der Kongruenzen (5) hat alle Wurzeln der zweiten und außerdem die Wurzel 0, die der zweiten nicht genügt. Ebenso hat die zweite alle Wurzeln der ersten, mit Ausnahme der Wurzel 0.

Da es nun für den Modul n überhaupt nur n verschiedene Zahlen gibt, so ist also zu beweisen, daß für jede ganze Zahl α die Kongruenz besteht:

(6) $$\alpha^n \equiv \alpha \pmod{n}.$$

Diese Kongruenz ist richtig für $\alpha = 0$ und $\alpha = 1$. Wir beweisen sie also durch vollständige Induktion, indem wir aus der als richtig vorausgesetzten Kongruenz (6) die Richtigkeit von

(7) $$(\alpha + 1)^n \equiv \alpha + 1 \pmod{n}$$

ableiten. Dies ist aber aus dem binomischen Satze zu schließen, wenn man beachtet, daß alle Binomialkoeffizienten, mit Ausnahme des ersten und des letzten, die gleich 1 sind, nämlich:

$$n, \quad \frac{n(n-1)}{1.2}, \quad \frac{n(n-1)(n-2)}{1.2.3} \cdots,$$

durch n teilbare ganze Zahlen sind. Demnach ist

$$(\alpha + 1)^n \equiv \alpha^n + 1 \pmod{n},$$

also folgt die Formel (7) aus der Formel (6).

Die Differenz

$$x(x-1)(x-2)\ldots(x-n+1) - x^n + x$$

ist eine Funktion, höchstens vom $n - 1$ten Grade. Sie ist aber für n Werte von x, nämlich für $x = 0, 1, 2, \ldots n - 1$ kongruent mit Null, und daher haben wir identisch

$$x(x-1)(x-2)\ldots(x-n+1) \equiv x^n - x \pmod{n}.$$

Daraus ergibt sich der Wilsonsche Lehrsatz:

$$1.2.3 \ldots (n-1) \equiv -1 \pmod{n},$$

wenn man die Koeffizienten der ersten Potenz von x beiderseits vergleicht.

Wir beschränken uns jetzt auf die zweite der Kongruenzen (5)

(8) $$x^{n-1} \equiv 1 \pmod{n}$$

und beweisen zunächst den Satz:

XI. Ist a ein Teiler von $n-1$, so hat die Kongruenz

$$x^a \equiv 1 \pmod{n} \tag{9}$$

genau a verschiedene Wurzeln.

Denn es ist, wenn $n - 1 = ab$ ist,

$$x^{n-1} - 1 = (x^a - 1)\,(x^{a(b-1)} + x^{a(b-2)} + \cdots + x^a + 1),$$

und da jede Wurzel der linken Seite Wurzel entweder des einen oder des anderen Faktors auf der rechten Seite sein muß, so folgt, daß, wenn $x^a - 1 \equiv 0$ weniger als a Wurzeln hätte, der zweite Faktor mehr Wurzeln haben müßte, als sein Grad angibt, entgegen dem Satz IX.

Ist α eine durch n nicht teilbare Zahl, und a die kleinste positive Zahl, für die $\alpha^a \equiv 1 \pmod{n}$ wird, ist ferner h irgend eine positive ganze Zahl, für die $\alpha^h \equiv 1 \pmod{n}$ ist, so ist a notwendig ein Teiler von h, insbesondere also a ein Teiler von $n - 1$.

Denn setzen wir $h = Qa + a'$, worin Q eine ganze Zahl und $a' < a$ ist, so ist auch $\alpha^{a'} \equiv 1$; also muß $a' = 0$ sein, weil a die kleinste positive Zahl sein sollte, wofür diese Kongruenz befriedigt ist.

Eine Wurzel α der Kongruenz (9) von der Eigenschaft, daß keine niedrigere als die $a$te Potenz mit positiven Exponenten der Einheit kongruent wird, heißt eine primitive Wurzel dieser Kongruenz oder primitive $a$te Kongruenzwurzel der Primzahl n, und eine primitive Wurzel g der Kongruenz (8) nennen wir auch kurz eine primitive Wurzel der Primzahl n.

Wir wollen beweisen, daß für jede Primzahl primitive Wurzeln existieren und ihre Anzahl bestimmen, und gehen dabei denselben Weg wie bei den Einheitswurzeln.

Es seien a und b zwei Teiler von $n - 1$, die unter sich relativ prim sind, und α eine $a$te, β eine $b$te primitive Kongruenzwurzel von n. Setzen wir $ab = c$, so ist $\gamma = \alpha\beta$ eine primitive $c$te Kongruenzwurzel, und umgekehrt läßt sich auch jede primitive $c$te Kongruenzwurzel in der Form $\alpha\beta$ darstellen.

Denn es ist erstens: $(\alpha\beta)^c = \alpha^c \beta^c \equiv 1 \pmod{n}$, und zweitens: wenn $(\alpha\beta)^h = \alpha^h \beta^h \equiv 1 \pmod{n}$ ist, so ist auch $\alpha^{hb}\,\beta^{hb} \equiv 1$, also $\alpha^{hb} \equiv 1$, also hb durch a teilbar, also auch h durch a teilbar, und ebenso schließt man, daß h durch b, also auch durch $ab = c$ teilbar sein muß, d. h. $\alpha\beta$ ist primitive $c$te Kongruenzwurzel.

Ist umgekehrt γ eine primitive $c$te Kongruenzwurzel, so bestimmen wir die positiven ganzen Zahlen x, y so, daß

$$ay + bx \equiv 1 \pmod{c} \tag{10}$$

wird, indem wir zunächst x aus $bx \equiv 1 \pmod{a}$, und dann y aus $y \equiv \frac{1 - bx}{a} \pmod{b}$ bestimmen. Dann ist y relativ prim zu b, und x relativ prim zu a. Hiernach ist

$$\gamma \equiv \gamma^{bx}\gamma^{ax} = \alpha\beta \pmod{n},$$

und $\gamma^{bx} = \alpha$ ist primitive $a$te Kongruenzwurzel, $\gamma^{ay} = \beta$ primitive $b$te Kongruenzwurzel (weil $\alpha^h = \gamma^{bxh}$ nur dann $\equiv 1$ sein kann, wenn h durch a teilbar ist).

Es ist noch zu zeigen, daß die Produkte $\alpha\beta$ alle voneinander verschieden sind, d. h. daß die Kongruenz

$$\alpha\beta \equiv \alpha'\beta' \pmod{n}, \tag{11}$$

wenn α, α'; β, β' primitive $a$te und $b$te Kongruenzwurzeln sind, nur befriedigt werden kann, wenn

$$\alpha \equiv \alpha', \quad \beta \equiv \beta' \pmod{n}$$

ist. Dies folgt aber nach (10), wenn man (11) in die Potenz

$$bx \equiv 1 \pmod{a}$$

erhebt, woraus sich $\alpha \equiv \alpha'$ ergibt, und dann auch $\beta \equiv \beta'$.

Bezeichnen wir also die Anzahl der primitiven a ten Kongruenzwurzeln mit $\varphi(a)$, so folgt aus dem Bewiesenen:

$$\varphi(c) = \varphi(a)\varphi(b). \tag{12}$$

Es bleibt noch übrig, wenn p eine Primzahl und pp_1 eine in $n - 1$ aufgehende Potenz von p ist, $\varphi(pp_1)$ zu bestimmen.

Ist α eine nicht primitive Kongruenzwurzel vom Grade p^π, und $p^{\pi-1} = p_1$, so ist, wenn α^h die niedrigste nach dem Modul n mit 1 kongruente Potenz von α ist, h ein Teiler von pp_1, d. h. eine Potenz von p; da h aber kleiner als pp_1 sein soll, so ist es auch ein Teiler von p_1, und folglich $\alpha^{p_1} \equiv 1 \pmod{n}$, also: eine nicht primitive Kongruenzwurzel des Grades pp_1 ist zugleich eine Kongruenzwurzel des Grades p_1. Umgekehrt ist jede Kongruenzwurzel des Grades p_1 zugleich eine imprimitive Kongruenzwurzel des Grades pp_1. Da nun die beiden Kongruenzen $x^{pp_1} \equiv 1$, $x^{p_1} \equiv 1$ so viele Wurzeln haben, als ihr Grad beträgt, so bleiben $pp_1 - p_1$ primitive Kongruenzwurzeln der ersten übrig, und es ist also

(13) $$\varphi(pp_1) = p_1(p-1)$$

oder

$$\varphi(p^\pi) = p^\pi\left(1-\frac{1}{p}\right).$$

Die Funktion $\varphi(a)$ hat also hier dieselbe Bedeutung wie in § 67 (11) und es ist damit zugleich die Anzahl der primitiven Wurzeln der Primzahl n gleich $\varphi(n-1)$ gefunden. Wir sprechen also den Satz aus:

XII. Eine Primzahl n hat $\varphi(n-1)$ primitive Wurzeln.

Damit ist die Existenz von primitiven Wurzeln für jede Primzahl nachgewiesen und zugleich ihre Anzahl bestimmt. Zur Auffindung der primitiven Wurzeln haben wir freilich keine andere allgemeine Methode als das Probieren.

Ist g eine primitive Wurzel der Primzahl n, so sind die Reste der Potenzen

(14) $$1, g, g^2, g^3 \ldots g^{n-2}$$

alle voneinander verschieden, da, wenn $g^m \equiv g^{m+\mu}$ wäre, $g^\mu \equiv 1$ sein müßte, was nicht möglich ist, solange μ positiv und kleiner als $n-1$ ist. Es muß also unter den Resten der Reihe (14), unter denen der Rest 0 nicht vorkommt, jede der Zahlen

(15) $$1, 2, 3, \ldots n-1$$

und jede nur einmal enthalten sein, oder mit anderen Worten:

XIII. Ist a eine durch n nicht teilbare Zahl, so gibt es eine und nur eine Zahl α aus der Reihe der Zahlen $0, 1, 2, \ldots n-2$, durch die die Kongruenz

(16) $$g^\alpha \equiv a \pmod{n}$$

befriedigt wird. Dieselbe Kongruenz wird aber auch befriedigt, wenn als Exponent eine mit α nach dem Modul $n-1$ kongruente Zahl gesetzt wird, und man findet in jedem vollen Restsystem nur eine solche Zahl α.

Der Exponent α heißt der Index von a in bezug auf die Basis g, und es wird geschrieben:

(17) $$\alpha \equiv \operatorname{ind} a \pmod{n-1}.$$

Man hat also für jede durch n nicht teilbare Zahl

(18) $$g^{\operatorname{ind} a} \equiv a \pmod{n}.$$

Aus den beiden Formeln

(19) $$g^{\operatorname{ind} a + \operatorname{ind} b} \equiv ab, \quad g^{m \operatorname{ind} a} \equiv a^m \pmod{n}$$

ergibt sich der Satz:

XIV. Der Index eines Produktes ist gleich der Summe der Indices der Faktoren; der Index der $m$ten Potenz von a ist gleich dem $m$fachen des Index von a.

Diese Sätze sind ganz analog den entsprechenden Sätzen, die sich auf die Rechnung mit Logarithmen beziehen, und die Analogie läßt sich noch weiter verfolgen. So wird z. B. der Übergang von einer Basis g zu einer anderen Basis g' durch die Formel

(20) $$\overset{g}{\operatorname{ind}} a \equiv \overset{g}{\operatorname{ind}} g' . \overset{g'}{\operatorname{ind}} a \pmod{n-1}$$

vermittelt.

Wir wollen noch den leicht zu beweisenden Satz anführen, daß unter den durch n nicht teilbaren Zahlen a die und nur die primitive Wurzeln sind, deren Indices relativ prim zu $n-1$ sind.

Für praktische Rechnungen bedient man sich zweckmäßig sogenannter Indextabellen, die den Logarithmentafeln entsprechen, wie überhaupt die Indices das genaue Analogon zu den Logarithmen sind.

Man nennt in der Kongruenz (17) α den Index und a den Numerus, und stellt am besten zwei Tabellen auf, von denen die eine zu jedem Index den Numerus, die andere zu jedem Numerus den Index gibt, wo die zweite durch Umstellung aus der ersten gewonnen wird. Dabei ist zu beachten, daß für die Indices der Modul $n-1$, für die Numeri der Modul n in Betracht kommt und kongruente Zahlen als gleichwertig gelten. So erhält man z. B. für $n = 13$ und die Basis 2:

I	0	1	2	3	4	5	6	7	8	9	10	11
N	1	2	4	8	3	6	12	11	9	5	10	7
N	1	2	3	4	5	6	7	8	9	10	11	12
I	0	1	4	2	9	5	11	3	8	10	7	6

Im Canon Arithmeticus von Jacobi (1839) sind für alle Primzahlen im ersten Tausend Indextabellen zusammengestellt; in kleinerem Umfange in Wertheim: „Anfangsgründe der Zahlenlehre“ (Braunschweig 1902).

Es ist noch zu bemerken, daß der Index von 1 immer gleich 0 ist, und daß der Index von -1 (oder von $n-1$) immer gleich $\frac{1}{2}(n-1)$ ist. Denn da $(-1)^2$ den Index 0 oder $n-1$ hat, so muß -1, da es nicht den Index 0 hat, und das Doppelte seines Index gleich 0 oder $n-1$ ist, den Index $\frac{1}{2}(n-1)$ haben; in Formeln:

(21) $\quad \text{ind } 1 \equiv 0, \quad \text{ind } (-1) \equiv \frac{1}{2}(n-1) \pmod{n-1}.$

Ist n eine Primzahl und m nicht durch $n-1$ teilbar, so gibt es immer wenigstens eine durch n nicht teilbare Zahl k, so daß k^m nicht mit 1 kongruent ist (mod n). Denn ist m' der Rest von m nach dem Modul $n-1$, so ist $k^m \equiv k^{m'} \pmod{n}$, und die Kongruenz $x^{m'} \equiv 1$ hat weniger als $n-1$ Wurzeln. Folglich gibt es eine durch n unteilbare Zahl k, die nicht Wurzel dieser Kongruenz ist. Nun durchläuft aber kx zugleich mit x ein volles Restsystem nach dem Modul n, und wir haben daher

$$\Sigma x^m \equiv \Sigma (kx)^m \pmod{n},$$

oder

$$(k^m - 1)\,\Sigma x^m \equiv 0 \pmod{n}.$$

Daraus ergibt sich für jeden durch $n-1$ nicht teilbaren Exponenten m:

(22) $\quad \Sigma x^m \equiv 0 \pmod{n}, \quad [m \text{ nicht} \equiv 0 \pmod{n-1}],$

wenn x ein volles Restsystem für den Modul n durchläuft.

Ist aber m durch $n-1$ teilbar, so ist x^m für jedes durch n unteilbare x mit 1 kongruent, und wir erhalten direkt:

(23) $\quad \Sigma x^m \equiv -1 \pmod{n}, \quad [m \equiv 0 \pmod{n-1}].$

Diese Sätze lassen sich auch dadurch beweisen, daß man x durch eine Potenz einer Primitivwurzel g darstellt, und zeigen eine weitere Analogie der Potenzreste mit den Einheitswurzeln.

§ 71.

Multiplikation und Teilung der Winkel.

Eine allgemeinere Aufgabe ist die Theorie der Winkelteilung die sich nicht mit der Teilung der ganzen Kreisperipherie, sondern mit der Teilung eines beliebigen Winkels befaßt.

Nach dem Moivreschen Satze ist, wenn n eine beliebige positive ganze Zahl und φ ein beliebiger Winkel ist,

(1) $\quad \cos n\varphi + i \sin n\varphi = (\cos \varphi + i \sin \varphi)^n,$

und wenn wir auf der rechten Seite den binomischen Satz anwenden, so folgt durch Trennung des Reellen vom Imaginären:

$$\begin{aligned}\cos n\varphi &= \cos\varphi^n - B_2^{(n)}\cos^{n-2}\varphi\sin^2\varphi \\ &\quad + B_4^{(n)}\cos^{n-4}\varphi\sin^4\varphi - \cdots, \\ \frac{\sin n\varphi}{\sin\varphi} &= n\cos^{n-1}\varphi - B_3^{(n)}\cos^{n-3}\varphi\sin^2\varphi \\ &\quad + B_5^{(n)}\cos^{n-5}\varphi\sin^4\varphi - \cdots\end{aligned} \tag{2}$$

worin $B_\nu^{(n)}$ die Binomialkoeffizienten sind, und die Summen auf der rechten Seite so weit fortzusetzen sind, als keine negativen Exponenten von $\cos\varphi$ vorkommen. Da nun $\sin^2\varphi = 1 - \cos^2\varphi$ ist, so lassen sich diese beiden Ausdrücke rational durch $\cos\varphi$ darstellen; wir setzen:

$$2\cos\varphi = x, \tag{3}$$

und führen die Bezeichnung ein:

$$2\cos n\varphi = A_n(x), \quad \frac{\sin n\varphi}{\sin\varphi} = B_n(x), \tag{4}$$

worin also $A_n(x)$ und $B_n(x)$ ganze Funktionen von x sind, $A_n(x)$ vom Grade n, $B_n(x)$ vom Grade $n-1$. Man findet in den ersten Fällen:

$$\begin{aligned} A_1(x) &= x, & B_1(x) &= 1 \\ A_2(x) &= x^2 - 2, & B_2(x) &= x \\ A_3(x) &= x^3 - 3x, & B_3(x) &= x^2 - 1. \end{aligned} \tag{5}$$

Nun ist aber nach bekannten Formeln

$$\begin{aligned}\cos(n+1)\varphi + \cos(n-1)\varphi &= 2\cos\varphi\cos n\varphi \\ \sin(n+1)\varphi + \sin(n-1)\varphi &= 2\cos\varphi\sin n\varphi\end{aligned}$$

woraus für A_n und B_n die Rekursionsformeln folgen:

$$\begin{aligned} A_{n+1}(x) &= xA_n(x) - A_{n-1}(x) \\ B_{n+1}(x) &= xB_n(x) - B_{n-1}(x). \end{aligned} \tag{6}$$

Daraus kann man A_n, B_n für jedes beliebige n berechnen, wenn diese Funktionen für $n = 1$ und $n = 2$ bekannt sind. So findet man:

$$\begin{aligned} A_4(x) &= x^4 - 4x^2 + 2 \\ A_5(x) &= x^5 - 5x^3 + 5x \\ A_6(x) &= x^6 - 6x^4 + 9x^2 - 2 \end{aligned}$$

$$\begin{aligned} B_4(x) &= x^3 - 2x \\ B_5(x) &= x^4 - 3x^2 + 1 \\ B_6(x) &= x^5 - 4x^3 + 3x, \end{aligned}$$

und so kann man fortfahren.

Nehmen wir $\cos n\varphi$ und $\sin n\varphi$ als gegeben an, so dient die Gleichung $n$ten Grades

$$A_n(x) - 2\cos n\varphi = 0 \tag{7}$$

zur Bestimmung der Unbekannten $x = 2\cos\varphi$. Die Bedeutung der n Wurzeln dieser Gleichung ergibt sich daraus, daß der Kosinus sich nicht ändert, wenn der Winkel um ein Vielfaches von 2π wächst. Demnach genügt der Gleichung (7) jeder Wert

$$x = 2\cos\left(\varphi + \frac{2\nu\pi}{n}\right), \tag{8}$$

wenn ν eine beliebige ganze positive oder negative Zahl bedeutet. Man erhält alle voneinander verschiedenen Werte (8), wenn man ν ein volles Restsystem in bezug auf n durchlaufen läßt, also z. B. $\nu = 0, 1, 2, \ldots n - 1$ setzt, und man sieht auch leicht, daß diese n Werte alle voneinander verschieden sind, wenn man von dem besonderen Falle absieht, daß φ ein Vielfaches von $\pi : n$ ist. Denn zwei Kosinus sind nur dann einander gleich, wenn die Summe oder die Differenz der Winkel ein Vielfaches von 2π ist.

Die Gleichung

$$\sin n\varphi = \sin\varphi\, B_n(x) \tag{9}$$

bestimmt, wenn $B_n(x)$ nicht gleich Null ist, $\sin\varphi$ eindeutig durch x.

Suchen wir in der Gleichung (7) das von x unabhängige Glied, so ergibt sich für ein ungerades n nach (7):

$$-2\cos n\varphi,$$

und wenn man dies dem negativen Produkt der Wurzeln gleichsetzt, so erhält man:

$$2^{n-1}\overset{\nu}{\Pi}\cos\left(\varphi + \frac{2\nu\pi}{n}\right) = \cos n\varphi, \tag{10}$$

und diese Formel gilt, da rechts und links stetige Funktionen von φ stehen, auch noch für den zunächst ausgeschlossenen Fall, wo φ ein Vielfaches von $\pi : n$ ist.

Ersetzt man φ durch $\varphi + \frac{\pi}{2}$, so folgt aus (10):

$$(-1)^{\frac{n-1}{2}}\, 2^{n-1}\overset{\nu}{\Pi}\sin\left(\varphi + \frac{2\nu\pi}{n}\right) = \sin n\varphi. \tag{11}$$

In beiden Formeln (10), (11) kann unter den Produktzeichen νm für ν gesetzt werden, wenn m eine beliebige, zu n teilerfremde Zahl ist, weil dann νm zugleich mit ν ein volles Restsystem nach dem Modul n durchläuft.

Sondern wir in den so erhaltenen Formeln den Faktor, der dem Wert $\nu = 0$ entspricht, ab und lassen dann ν sowohl die positiven als die negativen Zahlen durchlaufen, die absolut kleiner als $\frac{1}{2}n$ sind, so erhalten wir aus (10) und (11):

$$(12)\quad \begin{aligned} 2^{n-1} \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \cos\left(\frac{2\nu m\pi}{n} + \varphi\right) \cos\left(\frac{2\nu m\pi}{n} - \varphi\right) &= \frac{\cos n\varphi}{\cos\varphi} \\ 2^{n-1} \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \sin\left(\frac{2\nu m\pi}{n} + \varphi\right) \sin\left(\frac{2\nu m\pi}{n} - \varphi\right) &= \frac{\sin n\varphi}{\sin\varphi}, \end{aligned}$$

und diese Formeln lassen sich durch das Additionstheorem der trigonometrischen Funktionen auch so darstellen:

$$(13)\quad \begin{aligned} \frac{\cos n\varphi}{\cos\varphi} &= 2^{n-1} \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \left(\cos^2\frac{2\nu m\pi}{n} - \sin^2\varphi\right) \\ \frac{\sin n\varphi}{\sin\varphi} &= 2^{n-1} \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \left(\sin^2\frac{2\nu m\pi}{n} - \sin^2\varphi\right). \end{aligned}$$

Setzt man hierin $\sin^2\varphi = 1 - \cos^2\varphi = 1 - \frac{x^2}{4}$, so erhält man die Faktorenzerlegung von $A_n(x)$ und $B_n(x)$:

$$(14)\quad \begin{aligned} A_n(x) &= x \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \left(x^2 - 4\sin^2\frac{2\nu m\pi}{2}\right) \\ B_n(x) &= \overset{\prime}{\underset{1,\frac{n-1}{2}}{\Pi}} \left(x^2 - 4\cos^2\frac{2\nu m\pi}{2}\right). \end{aligned}$$

§ 72.

Quadratische Reste.

Setzen wir $\varphi = 0$, so ergibt sich aus der Formel (13) des vorigen Paragraphen durch Ziehen der Quadratwurzel:

$$(1)\quad 2^{\frac{n-1}{2}} \overset{\prime}{\Pi} \cos\frac{2\nu m\pi}{n} = \pm 1,$$

$$(2)\quad 2^{\frac{n-1}{2}} \overset{\prime}{\Pi} \sin\frac{2\nu m\pi}{n} = \pm\sqrt{n},$$

worin m relativ prim zu der ungeraden Zahl n ist, und ν die Reihe der Zahlen

$$\nu = 1, 2, \cdots \frac{n-1}{2}$$

durchläuft. Wir wollen das Zeichen ν für die Zahlen dieser Reihe hier festhalten.

Es bleibt das Vorzeichen in den Formeln (1) und (2) zu bestimmen.

Jede beliebige ganze Zahl gibt bei der Teilung durch n als Rest eine der Zahlen ν oder 0 oder $n - \nu$. Statt $n - \nu$ können wir, wenn wir auch negative Reste zulassen, $-\nu$ wählen, und wenn wir also von den durch n teilbaren Zahlen absehen, so bleibt eine der Zahlen $\pm\nu$ als Rest. Wir nennen diese den absolut kleinsten Rest (zum Unterschied von dem kleinsten Rest im gewöhnlichen Sinne, der aus den Zahlen $1, 2, \ldots n - 1$ genommen ist).

Es sei nun also m eine zu n teilerfremde positive oder negative Zahl; wir betrachten die Reihe der Zahlen

$$(m\nu) \qquad m,\ 2m,\ 3m, \cdots \frac{n-1}{2} m,$$

und bilden zu jeder den absolut kleinsten Rest ϱ:

$$(\varrho) \qquad \pm\nu_1, \pm\nu_2, \pm\nu_3, \ldots \pm\nu_{\frac{1}{2}(n-1)};$$

unter diesen Zahlen ϱ kommen nicht zwei gleiche und auch nicht zwei entgegengesetzte vor; denn wenn die Summe oder die Differenz zweier Zahlen (ϱ), also $\varrho + \varrho'$ oder $\varrho - \varrho'$ gleich Null wäre, so müßte für zwei verschiedene Zahlen ν, ν' aus (ν)

$$m(\nu \pm \nu')$$

durch n teilbar sein, also müßte auch $\nu \pm \nu'$ durch n teilbar sein, und dies ist unmöglich, da jede der Zahlen ν, ν' kleiner als $n:2$ ist.

Die Gesamtheit der Zahlen (ϱ) stimmt also, vom Vorzeichen und von der Reihenfolge abgesehen, mit der Gesamtheit der Zahlen (ν) überein.

In den Formeln (1) und (2) können wir aber, wegen der Periodizität von Sinus und Kosinus, νm durch ϱ ersetzen, und da $\cos(-\varphi) = \cos\varphi$ ist, so können wir in der Formel (1) $m\nu$ auch durch ν ersetzen, d. h. das Vorzeichen ist von m nicht abhängig. Um es zu bestimmen, berücksichtigen wir, daß der Kosinus eines Winkels im ersten Quadranten positiv, im zweiten Quadranten negativ ist, daß also das Vorzeichen in (1) positiv oder negativ ist, je nachdem von den Winkeln

$$\frac{2\nu\pi}{n}$$

eine gerade oder eine ungerade Anzahl im zweiten Quadranten liegt, oder je nachdem eine gerade oder eine ungerade Anzahl von Zahlen ν zwischen $\frac{1}{4}n$ und $\frac{1}{2}n$ liegt.

Wir bezeichnen, wenn x irgend eine nicht ganze Zahl ist, mit $E(x)$ die größte ganze Zahl, die kleiner als x ist, so daß x zwischen $E(x)$ und $E(x)+1$ liegt. Die Anzahl der ganzen Zahlen, die zwischen zwei Zahlen x und y liegt, ist dann gleich $E(y) - E(x)$, und wir haben also zu untersuchen, ob

$$E\left(\frac{n}{2}\right) - E\left(\frac{n}{4}\right)$$

eine gerade oder eine ungerade Zahl ist. Wir müssen vier Fälle unterscheiden, wie sich aus der folgenden Zusammenstellung ergibt, worin k eine nicht negative ganze Zahl bedeutet:

$$\begin{aligned}
&n = 8k+1, \quad E\left(\frac{n}{2}\right) = 4k, && E\left(\frac{n}{4}\right) = 2k,\\
&n = 8k+3, \quad E\left(\frac{n}{2}\right) = 4k+1, && E\left(\frac{n}{4}\right) = 2k,\\
&n = 8k+5, \quad E\left(\frac{n}{2}\right) = 4k+2, && E\left(\frac{n}{4}\right) = 2k+1,\\
&n = 8k+7, \quad E\left(\frac{n}{2}\right) = 4k+3, && E\left(\frac{n}{4}\right) = 2k+1.
\end{aligned}$$

Es ist also in (1) das positive Zeichen zu nehmen, wenn n von der Form $8k+1$ oder $8k+7$, das negative, wenn n von der Form $8k+3$ oder $8k+5$ ist. In den ersten Fällen ist $\frac{n^2-1}{8}$ eine gerade, in den beiden letzten eine ungerade Zahl, und demnach erhalten wir die genaue Formel (1):

$$2^{\frac{n-1}{2}} \overset{\nu}{\Pi} \cos\frac{2\nu m\pi}{n} = (-1)^{\frac{n^2-1}{8}}. \tag{3}$$

In der Formel (2) hängt das Vorzeichen von m ab; es ist positiv oder negativ, je nachdem unter den Zahlen (ϱ) eine gerade oder eine ungerade Anzahl negativer vorkommt.

Wir setzen nun, um die Abhängigkeit des Vorzeichens von m und n in der Bezeichnung auszudrücken, für (2):

$$2^{\frac{n-1}{2}} \overset{\nu}{\underset{1,\frac{n-1}{2}}{\Pi}} \sin\frac{2\nu m\pi}{n} = \left(\frac{m}{n}\right)\sqrt{n}, \tag{4}$$

worin

$$\left(\frac{m}{n}\right) = \pm 1$$

ist.

Das Symbol $\left(\frac{m}{n}\right)$ ist von einer anderen Seite her und in speziellerer Fassung von Legendre in die Zahlentheorie eingeführt und von Jacobi verallgemeinert worden. Wir wollen es das Legendresche Symbol nennen. Seine Bedeutung für die Zahlentheorie wird sich gleich ergeben; zunächst haben wir den Satz:

1. Es ist $\left(\frac{m}{n}\right)$ gleich $+1$ oder gleich -1, je nachdem die Anzahl der negativen unter den absolut kleinsten Resten von $m\nu$ eine gerade oder eine ungerade ist.

In der Formel (2) gilt für $m = 1$ das positive Zeichen, weil für diesen Fall die Winkel $2\nu\pi/n$ alle zwischen 0 und π liegen und ihre sin daher positiv sind. Also haben wir:

2. $$\left(\frac{1}{n}\right) = +1.$$

Verwandeln wir m in $-m$, so ändert sich in allen Faktoren des Produktes auf der linken Seite von (4) das Vorzeichen; da die Anzahl der Faktoren $\frac{n-1}{2}$ beträgt, so ergibt sich:

3. $$\left(\frac{-m}{n}\right) = (-1)^{\frac{n-1}{2}} \left(\frac{m}{n}\right),$$

und für $m = 1$:

4. $$\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}.$$

Setzen wir $2m$ für m und benutzen die Formel

$$\sin\frac{4\nu m\pi}{n} = 2\sin\frac{2\nu m\pi}{n}\cos\frac{2\nu m\pi}{n},$$

so folgt aus (3):

5. $$\left(\frac{2m}{n}\right) = (-1)^{\frac{n^2-1}{8}} \left(\frac{m}{n}\right),$$

und für $m = 1$:

6. $$\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}.$$

Ändert man m um ein Vielfaches von n, so bleiben alle Faktoren des Produktes (4) ungeändert, und daraus folgt:

7. Es ist

$$\left(\frac{m}{n}\right) = \left(\frac{m'}{n}\right),$$

wenn $m \equiv m' \pmod{n}$.

Um das Reziprozitätsgesetz zu erhalten, wenden wir die zweite Formel (13) des vorigen Paragraphen an, indem wir zur Abkürzung

$$\alpha = \sin^2 \frac{2\nu m\pi}{n}, \quad \nu = 1, 2, \cdots \frac{n-1}{2} \tag{5}$$

setzen:

$$\frac{\sin n\varphi}{\sin \varphi} = 2^{n-1} \overset{\alpha}{\Pi} (\alpha - \sin^2 \varphi). \tag{6}$$

Es sei jetzt m gleichfalls ungerade wie n, und es werde gesetzt:

$$\beta = \sin^2 \frac{2\mu n\pi}{m}, \quad \mu = 1, 2, \cdots \frac{m-1}{2}. \tag{7}$$

Dann ergibt sich aus (6) durch Vertauschung von m und n:

$$\frac{\sin m\varphi}{\sin \varphi} = 2^{m-1} \overset{\beta}{\Pi} (\beta - \sin^2 \varphi), \tag{8}$$

und wenn m' eine zu n teilerfremde Zahl ist und man

$$\varphi = \frac{2\nu m'\pi}{n},$$

setzt, so durchläuft $\sin^2 \varphi$ dieselbe Wertreihe wie α in (5). Bildet man dann aus (8) das Produkt über alle ν, so folgt:

$$\frac{\overset{\nu}{\Pi} \sin \frac{2\nu m m'\pi}{n}}{\overset{\nu}{\Pi} \sin \frac{2\nu m'\pi}{n}} = 2^{\frac{(m-1)(n-1)}{2}} \overset{\alpha}{\Pi} \overset{\beta}{\Pi} (\beta - \alpha). \tag{9}$$

Die rechte Seite dieser Formel ist aber von m' ganz unabhängig. Der Wert der linken Seite ist also derselbe, als ob $m' = 1$ wäre, und wenn man für die Produkte auf der linken Seite die Werte aus (4) setzt, so folgt:

$$\left(\frac{mm'}{n}\right) : \left(\frac{m'}{n}\right) = \left(\frac{m}{n}\right) : \left(\frac{1}{n}\right),$$

oder nach 2.:

8. $$\left(\frac{mm'}{n}\right) = \left(\frac{m}{n}\right)\left(\frac{m'}{n}\right).$$

Hier war zunächst vorausgesetzt, daß m und m' positiv und ungerade seien. Nach 3. und 5. aber bleibt 8. auch noch richtig, wenn m oder m' negativ oder gerade ist, wenn nur m und m' relativ prim zu n sind.

Nehmen wir wieder m ungerade und positiv an und setzen $m' = 1$, so folgt jetzt aus (9):

(10) $$\left(\frac{m}{n}\right) = 2^{\frac{(m-1)(n-1)}{2}} \prod_{\alpha}\prod_{\beta}(\beta - \alpha).$$

Wenn wir hierin m mit n vertauschen, so ändert in dem Doppelprodukt auf der rechten Seite jeder Faktor sein Vorzeichen. Die Anzahl dieser Faktoren ist aber

$$\frac{n-1}{2} \cdot \frac{m-1}{2}$$

und daraus folgt:

9. $$\left(\frac{m}{n}\right) = (-1)^{\frac{(m-1)(n-1)}{4}} \left(\frac{n}{m}\right),$$

oder $\left(\frac{m}{n}\right)$ ist gleich $\left(\frac{n}{m}\right)$, wenn von den beiden ungeraden Zahlen m, n wenigstens eine von der Form $4k+1$ ist, und $\left(\frac{m}{n}\right)$ ist entgegengesetzt zu $\left(\frac{n}{m}\right)$, wenn beide Zahlen von der Form $4k+3$ sind.

Dieser berühmte Satz ist das Reziprozitätsgesetz. Der hier gegebene Beweis rührt von Eisenstein her. Man kann mit seiner Hilfe und nach 5. und 6. den Wert des Symbols $\left(\frac{m}{n}\right)$ sehr schnell ermitteln, indem man so verfährt, als ob es sich um die Bestimmung des größten gemeinschaftlichen Teilers von m und n handelte.

Aus 8. und 9. ergibt sich noch ein letzter Satz, der gilt, wenn n und n' zwei positive ungerade zu m teilerfremde Zahlen sind:

10. $$\left(\frac{m}{n}\right)\left(\frac{m}{n'}\right) = \left(\frac{m}{nn'}\right).$$

Seine Richtigkeit ergibt sich, wenn m ungerade und positiv ist, auf Grund der Kongruenz

(11) $$(n-1)(n'-1) = (nn'-1)-(n-1)-(n'-1) \equiv 0 \pmod 4,$$

wenn man auf beide Seiten von 10. die Formel 9. und dann 8. anwendet, und wenn m gerade oder negativ ist, aus 4. und 6.

Nach diesen Sätzen kann man $\left(\frac{m}{n}\right)$ auf ein Produkt von Werten $\left(\frac{q}{p}\right)$ zurückführen, worin p, q Primzahlen sind. Die Berechnung von $\left(\frac{m}{n}\right)$ geschieht aber leichter nach 9. ohne diese Zerlegung.

Wir wollen nun noch eine andere, die ursprüngliche Legendresche Bedeutung des Symbols $\left(\frac{m}{p}\right)$ kennen lernen, für den Fall, daß p eine ungerade Primzahl ist.

Wenn wir alle durch p nicht teilbaren ganzen Zahlen x ins Quadrat erheben, und die Reste der Division durch p aufsuchen, so erhalten wir im ganzen nur $\frac{p-1}{2}$ verschiedene Reste, denn erstens geben die Quadrate kongruenter Zahlen dieselben Reste, und zweitens geben die Quadrate entgegengesetzter Zahlen auch dieselben Reste. Wir bekommen also gewiß alle Reste wenn wir die $\frac{p-1}{2}$ Zahlen ν

(ν) $$1, 2, 3, \ldots \frac{p-1}{2}$$

ins Quadrat erheben. Auf der anderen Seite ergeben auch die Quadrate dieser Zahlen ν lauter verschiedene Reste; denn es kann, wenn ν und ν' verschiedene von ihnen sind, niemals

$$\nu^2 - \nu'^2 = (\nu - \nu')(\nu + \nu')$$

durch p teilbar sein, weil sowohl $\nu - \nu'$ als $\nu + \nu'$ kleiner als p ist. Von den Zahlen

(s) $$1, 2, 3 \ldots p-1$$

kommt also die Hälfte unter den Resten von x^2 vor, die Hälfte nicht. Die ersteren Zahlen und alle mit ihnen nach p kongruenten heißen die quadratischen Reste von p, die anderen die quadratischen Nichtreste von p.

Wenn nun m quadratischer Rest ist, so ist die Kongruenz

(12) $$x^2 \equiv m \pmod p$$

möglich, und aus 7. und 8. folgt:

$$\left(\frac{m}{p}\right) = \left(\frac{x^2}{p}\right) = \left(\frac{x}{p}\right)^2 = +1.$$

Es ist also $\left(\frac{m}{p}\right) = +1$, wenn m quadratischer Rest von p ist.

Es ist aber in § 69, VIII. der Satz ausgesprochen, daß, wenn für m die Reihe der Zahlen $1, 2, \ldots p-1$ gesetzt wird, $\left(\frac{m}{p}\right)$ ebenso oft das positive wie das negative Zeichen hat, und daraus ersieht man, daß das negative Zeichen gilt, wenn m quadratischer Nichtrest ist, es folgt also:

11. Ist p eine ungerade Primzahl, und m durch p nicht teilbar, so ist $\left(\frac{m}{p}\right) = +1$ oder $= -1$, je nachdem m quadratischer Rest oder quadratischer Nichtrest von p ist.

Darin sind alle Hauptsätze über die quadratischen Reste enthalten, z. B. nach 8. der Satz:

12. Das Produkt zweier quadratischer Reste oder zweier Nichtreste ist ein Rest; das Produkt aus einem Rest und einem Nichtrest ist ein Nichtrest.

Nach dem Fermatschen Satze ist für jede durch p nicht teilbare Zahl m

$$m^{p-1} - 1 = \left(m^{\frac{p-1}{2}} - 1\right)\left(m^{\frac{p-1}{2}} + 1\right) \equiv 0 \pmod{p}.$$

Es muß also einer der beiden Faktoren

$$m^{\frac{p-1}{2}} - 1, \quad m^{\frac{p-1}{2}} + 1$$

durch p teilbar sein, und es kann auch nur einer von ihnen sein, weil sonst auch ihre Differenz, die gleich 2 ist, durch p teilbar sein müßte. Wenn nun die Kongruenz (12) möglich ist, so ist

$$m^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod{p},$$

woraus folgt, daß alle quadratischen Reste, deren Anzahl $\frac{1}{2}(p-1)$ ist, Wurzeln der Kongruenz

$$x^{\frac{p-1}{2}} \equiv 1 \pmod{p}$$

sind. Da aber diese Kongruenz nach § 70, IX. nicht mehr als $\frac{1}{2}(p-1)$ Wurzeln haben kann, so sind die quadratischen Nichtreste Wurzeln der Kongruenz

$$x^{\frac{p-1}{2}} \equiv -1 \pmod{p},$$

und hieraus folgt der für jede durch p nicht teilbare Zahl m gültige Satz von Euler:

$$m^{\frac{p-1}{2}} \equiv \left(\frac{m}{p}\right) \pmod{p}, \tag{13}$$

oder in Worten:

13. Die durch p nicht teilbare Zahl m ist quadratischer Rest oder Nichtrest von p, je nachdem

$$m^{\frac{p-1}{2}} - 1 \quad \text{oder} \quad m^{\frac{p-1}{2}} + 1$$

durch p teilbar ist.

Wir können noch hinzufügen, daß die quadratischen Reste in jedem Indexsysteme gerade Indices haben, die Nichtreste ungerade.

Hieraus lassen sich noch einige weitere bemerkenswerte Folgerungen ziehen. Wir bezeichnen mit a die quadratischen Reste, mit b die Nichtreste der ungeraden Primzahl p, so daß die Anzahl der a und b je $\frac{1}{2}(p-1)$ beträgt. Mit $\Pi(a)$ und $\Pi(b)$ bezeichnen wir die Produkte der sämtlichen a und der sämtlichen b. Die a und b zusammen enthalten alle Zahlen $1, 2, \ldots p-1$, und es ist daher nach dem Wilsonschen Satze (§ 66)

$$\Pi(a)\,\Pi(b) \equiv -1 \pmod{p}. \tag{14}$$

Ist b_0 ein fester Nichtrest, so durchläuft $b_0 a$ nach 12. die sämtlichen Nichtreste b, wenn a die sämtlichen Reste durchläuft, und daraus ergibt sich nach 13.:

$$\Pi(b) \equiv b_0^{\frac{p-1}{2}}\,\Pi(a) \equiv -\Pi(a) \pmod{p}. \tag{15}$$

Ferner ist, da jedes a dem Quadrate einer der Zahlen $1, 2, \cdots \frac{1}{2}(p-1)$ kongruent ist,

$$\Pi(a) \equiv \left(1.2.3 \cdots \frac{p-1}{2}\right)^2 \pmod{p}. \tag{16}$$

Die in dem Wilsonschen Satze vorkommenden Zahlen $1, 2, \ldots p-1$ kann man aber auch durch die Zahlen $\pm 1, \pm 2, \cdots \pm\frac{1}{2}(p-1)$ ersetzen, und daraus ergibt sich:

$$\left(1.2 \cdots \frac{p-1}{2}\right)^2 \equiv (-1)^{\frac{p+1}{2}} \pmod{p},$$

und es folgen also aus (15) und (16) die Formeln

$$\Pi(a) \equiv (-1)^{\frac{p+1}{2}}, \quad \Pi(b) \equiv (-1)^{\frac{p-1}{2}} \pmod{p}. \tag{17}$$

Zwölfter Abschnitt.

Auflösung der Kreisteilungsgleichung.

§ 73.

Irreduzibilität der Kreisteilungsgleichung.

Ist n eine beliebige natürliche Zahl, Primzahl oder zusammengesetzt, so gibt es, wie wir in § 67 gesehen haben, $\varphi(n) = \nu$ primitive $n$te Einheitswurzeln, die einer ganzzahligen Gleichung

$$X_n = x^\nu + a_1 x^{\nu-1} + \cdots + a_\nu = 0 \tag{1}$$

genügen. Die nächste Frage, die nun zu beantworten ist, ist die nach der Irreduzibilität dieser Gleichung.

Die Irreduzibilität von X_n ist für den Fall, daß n eine Primzahl ist, zuerst von Gauß bewiesen (Disq. arithmeticae art. 341). Später sind noch viele andere Beweise gegeben worden (von Kronecker, Schönemann, Eisenstein, Arndt, Dedekind, Mertens), die sich zum Teil auf denselben einfachen Fall oder den ebenso zu behandelnden Fall, wo n eine Potenz einer Primzahl ist, beziehen, zum Teil den allgemeinen Fall eines beliebig zusammengesetzten n behandeln. Wir wollen hier einem besonders einfachen Beweise für die Irreduzibilität von X_n folgen, den Dedekind gegeben hat, der sich gleich auf den allgemeinen Fall eines ganz beliebigen n bezieht[1]).

Erheben wir irgend ein Polynom in die $p$te Potenz, so ergibt sich ein Ausdruck der Form

$$(u + v + w + \cdots)^p = \Sigma P_{\alpha,\beta,\gamma} \ldots u^\alpha v^\beta w^\gamma \ldots,$$

worin α, β, γ ... alle nicht negativen ganzen Zahlen zu durchlaufen hat, die der Bedingung

[1]) Über die Geschichte der Irreduzibilitätsbeweise sehe man die Straßburger Dissertation von M. Ruthinger von 1907: „Die Irreduzibilitätsbeweise der Kreisteilungsgleichung".

$$\alpha + \beta + \gamma + \cdots = p$$

genügen, und die Polynomialkoeffizienten

$$P_{\alpha,\beta,\gamma\ldots} = \frac{p!}{\alpha!\,\beta!\,\gamma!\ldots}$$

sind ganze Zahlen. Ist nun p eine Primzahl, so ist der Zähler dieses Ausdruckes $p!$ durch p teilbar. Im Nenner kommt aber die Primzahl p nicht vor, außer wenn eine der Zahlen $\alpha, \beta, \gamma \cdots = p$ und die anderen gleich 0 sind. Es sind daher die Koeffizienten $P_{\alpha,\beta,\gamma\ldots}$ alle durch p teilbar, außer in dem erwähnten Ausnahmefalle, in dem $P_{p,0,0\ldots} = 1$ ist. Demnach ist für unbestimmte $u, v, w \ldots$:

$$(2) \qquad (u + v + w + \cdots)^p \equiv u^p + v^p + w^p + \cdots \pmod{p}.$$

Hieraus ergibt sich nun ein Hilfssatz, auf den wir unseren Beweis stützen wollen. Es sei

$$(3) \qquad f(x) = x^m + a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_m$$

irgend eine ganze Funktion von x mit ganzzahligen rationalen Koeffizienten $a_1, a_2, \ldots a_m$, und die Wurzeln dieser Funktion seien $\alpha, \beta, \gamma, \ldots$, so daß

$$(4) \qquad -a_1 = \Sigma\alpha, \quad a_2 = \Sigma\alpha\beta, \quad -a_3 = \Sigma\alpha\beta\gamma, \ldots$$

Wir setzen nun, wenn p eine beliebige Primzahl ist,

$$(5) \qquad \begin{aligned} F(x) &= (x - \alpha^p)\,(x - \beta^p)\,(x - \gamma^p)\ldots \\ &= x^m + A_1 x^{m-1} + A_2 x^{m-2} + \cdots + A_m, \end{aligned}$$

worin die Koeffizienten $A_1, A_2, \ldots A_m$ als ganzzahlige symmetrische Funktionen der $\alpha, \beta, \gamma \ldots$ nach dem Fundamentalsatz von den symmetrischen Funktionen ganze ganzzahlige Funktionen der $a_1, a_2, \ldots a_m$ und also auch ganze rationale Zahlen sind. Zugleich ist

$$-A_1 = \Sigma\alpha^p, \quad A_2 = \Sigma\alpha^p\beta^p, \quad -A_3 = \Sigma\alpha^p\beta^p\gamma, \ldots$$

und folglich nach dem Satze (2)

$$A_1 \equiv a_1^p, \quad A_2 \equiv a_2^p, \ldots A_m \equiv a_m^p \pmod{p},$$

und nach dem Fermatschen Satze

$$(6) \qquad A_1 \equiv a_1, \quad A_2 \equiv a_2, \ldots A_m \equiv a_m \pmod{p},$$

was wir auch so ausdrücken können, wenn wir wieder die Kongruenz nur auf die Koeffizienten entsprechender Potenzen von x beziehen:

$$(7) \qquad f(x) \equiv F(x) \pmod{p}.$$

Zur Berechnung von $F(x)$ kann auch die Darstellung dienen:

$$F(x) = \overset{\varepsilon}{\Pi} f(\varepsilon \sqrt[p]{x}), \tag{8}$$

worin das Produkt über alle $p$ten Einheitswurzeln ε zu nehmen ist. Wir fügen noch hinzu, daß, wenn $f(x)$ im Körper R irreduzibel ist, und wenn die α^p, β^p, γ^p, ... voneinander verschieden sind, auch $F(x)$ irreduzibel ist. Denn ist $F_1(x)$ ein irreduzibler Faktor von F, der für α^p verschwindet, so hat $F_1(x^p)$ eine Wurzel mit $f(x)$ gemein, und ist daher wegen der vorausgesetzten Irreduzibilität von $f(x)$ durch $f(x)$ teilbar (§ 53, I.); es ist also

$$F_1(\alpha^p) = 0, \quad F_1(\beta^p) = 0, \quad F_1(\gamma^p) = 0 \ldots,$$

d. h. $F_1(x)$ verschwindet für alle Wurzeln von $F(x)$, woraus die Irreduzibilität von $F(x)$ folgt.

Hieraus ergibt sich nun der Beweis der Irreduzibilität von X_n auf folgende Weise:

Es sei α eine primitive $n$te Einheitswurzel, und $f(x)$ der im Körper R irreduzible Faktor von X_n, der für $x = \alpha$ verschwindet. Hat die höchste in $f(x)$ vorkommende Potenz von x den Koeffizienten 1, so müssen auch die übrigen Koeffizienten $a_1, a_2 \ldots a_m$ nach dem Satze von Gauß (§ 21) ganze Zahlen sein.

Wenn μ eine zu n teilerfremde Zahl ist, so ist α^μ ebenfalls primitive $n$te Einheitswurzel, und jede primitive $n$te Einheitswurzel, also jede Wurzel der Gleichung $X_n = 0$ ist in dieser Form enthalten. Wenn also bewiesen werden kann, daß für alle Exponenten μ, die mit n keinen gemeinsamen Teiler haben, $f(\alpha^\mu) = 0$ ist, so folgt, daß $f(x)$ durch X_n teilbar und folglich mit X_n identisch ist, wodurch dann bewiesen ist, daß X_n in R irreduzibel ist.

Es genügt aber hierzu, nachzuweisen, daß für jede Primzahl p, die in n nicht aufgeht, $f(\alpha^p) = 0$ ist. Denn ist $f(\alpha^p) = 0$, so hat $f(x^p)$ mit $f(x)$ eine gemeinsame Wurzel α und folglich ist $f(x^p)$ wegen der vorausgesetzten Irreduzibilität von $f(x)$ durch $f(x)$ teilbar. Wenn also für eine zweite in n nicht aufgehende Primzahl q, die auch mit p identisch sein kann, $f(\alpha^q) = 0$ ist, so folgt, daß auch $f(\alpha^{pq}) = 0$ ist, und wenn man diesen Schluß wiederholt, so erkennt man, daß, wenn $f(\alpha^p)$ für jede nicht in n aufgehende Primzahl verschwindet, auch $f(\alpha^\mu)$ verschwinden muß, wenn μ irgend ein Produkt von Primzahlen ist, die in n nicht aufgehen.

Ist p eine in n nicht aufgehende Primzahl und sind α und β zwei voneinander verschiedene Wurzeln von X_n, so sind auch α^p und β^p voneinander verschiedene primitive $n$te Einheitswurzeln, denn bestimmt man p' aus der Kongruenz $p'p \equiv 1 \pmod{n}$, so folgt aus $\alpha^p = \beta^p$ durch Erhebung in die Potenz p', daß $\alpha = \beta$ sein muß.

Sind nun $\alpha, \beta, \gamma, \ldots$ die Wurzeln von $f(x)$, so bilden wir nach (8) die Funktion $F(x)$, deren Wurzeln $\alpha^p, \beta^p, \gamma^p, \ldots$ sind, und die also nach dem oben Bewiesenen gleichfalls irreduzibel ist.

Wenn $f(x)$ von $F(x)$ verschieden ist, so haben diese beiden Funktionen auch keinen gemeinschaftlichen Teiler, und beide sind Teiler der Funktion X_n und folglich auch Teiler von $x^n - 1$. Wir setzen

$$x^n - 1 = f(x) F(x) \varphi(x),$$

worin $\varphi(x)$ eine ganzzahlige Funktion ist. Hierfür können wir aber wegen (7) auch setzen:

$$x^n - 1 = f(x)^2 \varphi(x) + p\psi(x),$$

und durch Bildung der abgeleiteten Funktionen:

$$n x^{n-1} = f(x) \chi(x) + p\Theta(x),$$

worin $\varphi, \psi, \chi, \Theta$ ganze ganzzahlige Funktionen sind. Wenn wir die vorletzte dieser Gleichungen mit $-n$, die letzte mit x multiplizieren und addieren, so ergibt sich eine Gleichung, die wir als Kongruenz so darstellen können:

$$n \equiv f(x) \Phi(x) \pmod{p}, \tag{9}$$

worin $\Phi(x)$ wieder eine ganzzahlige Funktion ist. Da n nicht durch p teilbar ist, so können hier auch nicht alle Koeffizienten von $\Phi(x)$ durch p teilbar sein, und wir können annehmen, daß der Koeffizient der höchsten Potenz von Φ nicht durch p teilbar ist. Dann aber ist die Kongruenz (9) offenbar unmöglich, da auf der rechten Seite gewiß ein von x abhängiges Glied mit nicht durch p teilbarem Koeffizienten vorkommt, was auf der linken Seite nicht der Fall ist.

Es muß also $F(x)$ mit $f(x)$ identisch sein, und es ist daher $f(\alpha^p) = 0$. W. z. b. w.

Hiermit ist bewiesen:

I. Die Funktion X_n, deren Wurzeln die primitiven $n$ten Einheitswurzeln sind, ist im Körper der rationalen Zahlen irreduzibel.

Die Irreduzibilität von X_n gilt aber, wie hieraus sofort folgt, in einem noch weiteren Sinne.

II. Die Funktion X_n ist auch dann noch irreduzibel, wenn dem Körper R beliebige Einheitswurzeln adjungiert werden, deren Grad relativ prim zu n ist.

Es sei nämlich a zu n teilerfremd, r eine primitive $n$te und α eine primitive $a$te Einheitswurzel, also nach § 67, I., II.

$$\varrho = r\alpha$$

eine primitive nate Einheitswurzel. Bestimmen wir die ganzen Zahlen x, y so, daß

$$ax + ny = 1$$

wird, so ist

$$r = \varrho^{ax}, \qquad \alpha = \varrho^{ny}, \tag{10}$$

und wenn $\varphi(x, \alpha)$ ein Teiler von X_n im Körper $R(\alpha)$ ist, so gibt es wenigstens ein r, für das

$$\varphi(r, \alpha) = 0$$

ist. Daraus ergibt sich aber nach (10)

$$\varphi(\varrho^{ax}, \varrho^{ny}) = 0. \tag{11}$$

Nun aber haben wir die Irreduzibilität der Gleichung für die primitiven naten Einheitswurzeln im Körper R bewiesen. Es folgt daher aus (11) für jede zu an teilerfremde Zahl h

$$\varphi(\varrho^{hax}, \varrho^{hny}) = 0, \tag{12}$$

und wenn s eine beliebige zu n teilerfremde Zahl ist, so können wir h so bestimmen, daß

$$h \equiv s \pmod{n}, \qquad h \equiv 1 \pmod{a}.$$

Dann folgt aus (12)

$$\varphi(r^s, \alpha) = 0,$$

und folglich ist $\varphi(x, \alpha)$ durch X_n teilbar. Darin liegt die Irreduzibilität von X_n im Körper $R(\alpha)$.

Dieser Satz rührt von Kronecker her, der gewissermaßen den umgekehrten Weg geht, indem er zunächst die Irreduzibilität von X_n in $R(\alpha)$ unter der Voraussetzung beweist, daß n nur durch eine Primzahl teilbar ist, und daraus dann die Irreduzibilität von X_n allgemein ableitet[1]).

---

[1]) Kronecker, Mém. sur les facteurs irréducibles de l'expression $(x^n - 1)$. Liouvilles Journal, Bd. 19 (1854).

§ 74.

Die Kreisteilungsperioden und die Periodengleichungen.

Wir beschäftigen uns nun zunächst mit den $n$ten Einheitswurzeln, unter der Voraussetzung, daß n eine ungerade Primzahl ist (die Primzahl 2 bietet zu keinen weiteren Fragen Anlaß, da die einzigen zweiten Einheitswurzeln, ± 1, in R enthalten sind).

Außer der rationalen $n$ten Einheitswurzel 1 existieren noch $n-1$ primitive, die durch die transzendenten Ausdrücke

$$e^{\frac{2\pi i}{n}}, \quad e^{\frac{4\pi i}{n}}, \quad e^{\frac{6\pi i}{n}} \ldots e^{\frac{2(n-1)\pi i}{n}}$$

dargestellt werden können. Wenn wir eine von ihnen mit r bezeichnen, so ist das ganze System auch durch

$$r, \quad r^2, \quad r^3 \ldots r^{n-1} \tag{1}$$

darzustellen. Im Exponenten von r kommt es nur auf den Rest nach dem Modul n an, da immer und nur dann $r^h = r^k$ ist, wenn $h \equiv k \pmod{n}$ ist.

Die Größen (1) sind die Wurzeln der Gleichung $(n-1)$ten Grades:

$$X = x^{n-1} + x^{n-2} + x^{n-3} + \cdots + x + 1 = 0, \tag{2}$$

deren Irreduzibilität im vorigen Paragraphen bewiesen ist. Wir wollen jetzt zunächst nachweisen, daß diese Gleichung zyklisch ist. Diese Eigenschaft ergibt sich aus der Existenz primitiver Wurzeln der Primzahl n, mit denen wir uns im § 65 beschäftigt haben. Es wurde dort nachgewiesen, daß es für jede Primzahl n gewisse Zahlen g gibt, die man primitive Wurzeln von n nennt, die durch die Eigenschaft charakterisiert sind, daß unter den Resten der Potenzen

$$1, \quad g, \quad g^2, \quad g^3, \ldots g^{n-2},$$

bei der Teilung durch n jede der Zahlen

$$1, \quad 2, \ldots, n-1$$

ein und nur einmal vorkommt. Der Rest einer Potenz g^h bleibt derselbe, wenn h um ein Vielfaches von $n-1$ verändert wird. Ist

$$g^\alpha \equiv a \pmod{n},$$

so heißt α der Index von a ($\alpha = \operatorname{ind} a$) und a die Zahl oder der Numerus. Der Index wird nach dem Modul $n-1$ genommen, der Numerus nach dem Modul n (§ 70).

Nehmen wir also eine solche primitive Wurzel g von n an, so können wir, von der Reihenfolge abgesehen, die Größen (1) so darstellen:

$$r, \quad r^g, \quad r^{g^2} \dots r^{g^{n-2}},$$

oder, wenn wir

(3) $$r^{g^h} = r_h$$

setzen,

(4) $$r, \quad r_1, \quad r_2, \dots r_{n-2}.$$

Jede Zahl des Körpers $R(r)$, d. h. jede rationale Funktion von r läßt sich in die Form bringen:

(5) $$\varphi(r) = b_0 + b_1 r + b_2 r^2 + \cdots + b_{n-2} r^{n-2},$$

oder, da nach (2)

$$1 + r + r^2 + \cdots + r^{n-1} = 0$$

ist,

(6) $$\varphi(r) = (b_1 - b_0) r + (b_2 - b_0) r^2 + \cdots + (b_{n-2} - b_0) r^{n-2} - b_0 r^{n-1},$$

oder, da die Potenzen $r, r^2 \dots r^{n-1}$, von der Reihenfolge abgesehen, mit $r, r_1, r_2 \dots r_{n-2}$ übereinstimmen, in die Form

(7) $$\varphi(r) = a_0 r + a_1 r_1 + a_2 r_2 + \cdots + a_{n-2} r_{n-2},$$

worin die Koeffizienten b und a rationale Zahlen sind. Wenn von den Zahlensystemen a, b das eine ganzzahlig ist, so ist es auch das andere.

III. Eine solche Funktion $\varphi(r)$ kann nur dann gleich Null sein, wenn alle ihre Koeffizienten $a_0, a_1, \dots a_{n-2}$ verschwinden;

denn der Ausdruck (5) zeigt, da X irreduzibel ist, also r keiner Gleichung von niedrigerem als den $(n-1)$ten Grade genügen kann und (6) nach Division mit r auf den $(n-2)$ten Grad kommt, daß $\varphi(r)$ nicht anders verschwinden kann, als wenn die $b_0, b_1, \dots b_{n-2}$ Null sind. Sind aber diese Null, so zeigt (6), daß auch die Koeffizienten a alle Null sein müssen. Es kann also $\varphi(r)$ nur auf eine Weise in die Form (7) gebracht werden.

Treffen wir die Festsetzung, daß sich a_h und r_h nicht ändern sollen, wenn der Index h um ein Vielfaches von $n-1$ wächst, so können wir auch

(8) $$\varphi(r) = \sum^{h} a_h r_h$$

setzen, und darin h ein volles Restsystem nach dem Modul $n-1$ durchlaufen lassen.

Wir haben nun nachzuweisen, daß die Gruppe der Gleichung $X = 0$ keine andere ist, als die Periode der zyklischen Permutation

$$\pi = (r, r_1, r_2 \dots r_{n-2}).$$

Diese Periode oder zyklische Gruppe bezeichnen wir mit

$$C = 1, \pi, \pi^2 \dots \pi^{n-2}. \tag{9}$$

Die Gleichung $X = 0$ ist jedenfalls eine Normalgleichung, weil sie irreduzibel ist, und weil alle ihre Wurzeln nach (3) rational durch eine unter ihnen ausdrückbar sind; sie ist also ihre eigene Galoissche Resolvente (§ 56). Ihre Substitutionen sind

$$(r, r), \quad (r, r_1), \quad (r, r_2) \dots (r, r_{n-2}), \tag{10}$$

und nach (3) ist

$$(r, r_h) = (r_k, r_{k+h}).$$

Daraus folgt, daß (r, r_1) unter den Wurzeln (4) die zyklische Permutation π hervorruft, und daß also die Substitutionsgruppe (10) mit der Permutationsgruppe C isomorph ist. Also ist C die Galoissche Gruppe von $X = 0$.

Nach § 63 läßt sich die Gleichung $X = 0$ auf eine Reihe von zyklischen Gleichungen niedrigeren Grades zurückführen, wenn man Funktionen aufsucht, die zu den Teilern der Gruppe C gehören. Die Teiler von C sind aber, wenn $n - 1$ in zwei Faktoren e, f zerlegt, also

$$n - 1 = ef$$

gesetzt wird, die zyklischen Gruppen

$$C_e = 1, \pi^e, \pi^{2e} \dots \pi^{(f-1)e}, \tag{11}$$

und die Permutation π^e ist darin gleichbedeutend mit der Substitution (r, r_e).

Um zu einer zu C gehörigen Funktion der Wurzeln von möglichst einfacher Form zu gelangen, betrachten wir nach Gauß die f-gliedrigen Perioden

$$\begin{aligned}
\eta &= r + r_e + r_{2e} + \dots + r_{(f-1)e} \\
\eta_1 &= r_1 + r_{e+1} + r_{2e+1} + \dots + r_{(f-1)e+1} \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\eta_{e-1} &= r_{e-1} + r_{2e-1} + r_{3e-1} + \dots + r_{n-2},
\end{aligned} \tag{12}$$

die aus der ersten von ihnen durch die Substitutionen (r, r), (r, r_1), (r, r_2), $\dots$ (r, r_{e-1}) hervorgehen. Wir bezeichnen die Größen (12) auch als ein System konjugierter Perioden.

Jede dieser Funktionen bleibt durch die Substitution (r, r_e) ungeändert, und sie sind alle voneinander verschieden, da sich eine Gleichung $\eta_h - \eta_k = 0$, wenn h und k verschieden sind, in der Form

$$a_0 r + a_1 r_1 + \cdots + a_{n-2} r_{n-2} = 0$$

schreiben ließe, wo die $a_0, a_1, \ldots a_{n-2}$ nicht alle zugleich verschwindende ganze Zahlen sind, die die Werte $+1$, 0 oder -1 annehmen, was nach dem Theorem III nicht möglich ist. Jede der Perioden (12) ist also eine zu der Gruppe C_e gehörige Funktion, und diese Größen sind daher die Wurzeln einer irreduzibeln ganzzahligen Gleichung $e$ten Grades. Adjungiert man eine dieser Größen, so hängt die Bestimmung von r noch von einer Gleichung vom Grade f ab.

Nach den allgemeinen Sätzen des neunten Abschnittes kann jede Funktion der r, die die Substitutionen der Gruppe C_e gestattet, rational durch jede der Größen η dargestellt werden. Hier können wir aber noch den folgenden Satz aufstellen:

IV. Jede Zahl in $R(r)$, die die Substitution (r, r_e) gestattet, also auch jede Zahl des Körpers $R(\eta)$, läßt sich als homogene lineare Funktion der Perioden $\eta, \eta_1 \ldots \eta_{e-1}$ darstellen.

Denn nach III. können wir jede Zahl von $R(r)$ auf eine Weise in die Form

$$\varphi(r) = \Sigma a_h r_h \tag{13}$$

setzen, wo die Koeffizienten a_h rationale Zahlen sind, und h ein volles Restsystem nach dem Modul $n - 1$ durchläuft. Es ist aber

$$\varphi(r_e) = \Sigma a_h r_{h+e} = \Sigma a_{h-e} r_h,$$

und daher muß, wenn $\varphi(r) = \varphi(r)$ sein soll, $a_{h-e} = a_h$ sein für jedes h; also auch $a_h = a_{h+e} = a_{h+2e} \ldots$ Danach läßt sich (13) in die Form setzen:

$$\varphi(r) = \sum_{0,e-1}^{h} a_h r_h + \sum_{0,e-1}^{h} a_h r_{h+e} + \cdots + \sum_{0,e-1}^{h} a_h r_{h+(f-1)e},$$

oder

$$\varphi(r) = a_0 \eta + a_1 \eta_1 + \cdots + a_{e-1} \eta_{e-1}, \tag{14}$$

was zu beweisen war. Zu bemerken ist dabei noch, daß, wenn φ ganzzahlige Koeffizienten hat, auch die Koeffizienten der η^h ganze Zahlen werden.

Da jede Zahl des Körpers $R(\eta)$ auf eine und nur auf eine Art in der linearen Form (14) dargestellt werden kann, so nennen wir das Größensystem η, η_1, η_2, ... η_{e-1} eine Basis des Körpers $R(\eta)$.

Wenn man die Darstellung des Produktes je zweier Perioden $\eta_h \eta_k$ in der Form (14) kennt, so kann man daraus durch wiederholte Anwendung alle rationalen Funktionen der Perioden η in derselben Form darstellen. Ebenso kann man auch leicht die Gleichung bilden, deren Wurzeln die e Größen η sind, wenn man das Verfahren auf das Produkt

$$F_e(x) = (x - \eta)\,(x - \eta_1) \ldots (x - \eta_{e-1}) \tag{15}$$

anwendet. Diese Gleichung läßt sich auch in Determinantenform darstellen.

Man setze nämlich

$$\eta\,\eta_h = a_{0,h}\,\eta + a_{1,h}\,\eta_1 + \cdots + a_{e-1,h}\,\eta_{e-1}, \tag{16}$$

worin die $a_{0,h}$, $a_{1,h}$, ... $a_{e-1,h}$ ganze Zahlen sind, über deren Berechnung später das Nähere folgen wird. Stellt man diese Gleichung für $h = 0, 1, \ldots e - 1$ in der Form auf:

$$\begin{array}{llll}
(a_{0,0} - \eta)\,\eta + & a_{1,0}\,\eta_1 + \cdots\cdots\cdots + & a_{e-1,0}\,\eta_{e-1} = 0 \\
a_{0,1}\,\eta + & (a_{1,1} - \eta)\,\eta_1 + \cdots\cdots\cdots + & a_{e-1,1}\,\eta_{e-1} = 0 \\
\cdots\cdots & \cdots\cdots & \cdots\cdots \\
a_{0,e-1}\,\eta + & a_{1,e-1}\,\eta_1 + \cdots & (a_{e-1,e-1} - \eta)\,\eta_{e-1} = 0,
\end{array}$$

so sieht man, da die η, η_1, ... η_{e-1} nicht verschwinden, daß die Determinante dieses Systems Null sein muß, also

$$\begin{vmatrix}
a_{0,0} - \eta, & a_{1,0}, & \ldots & a_{e-1,0} \\
a_{0,1}, & a_{1,1} - \eta, & \ldots & a_{e-1,1} \\
\cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\
a_{0,e-1}, & a_{1,e-1}, & \ldots & a_{e-1,e-1} - \eta
\end{vmatrix} = 0, \tag{17}$$

was, wenn man x für η schreibt, vom Vorzeichen abgesehen, mit $F_e(x)$ übereinstimmen muß.

Die Funktion $F_e(x)$ hat also nicht bloß rationale, sondern ganzzahlige Koeffizienten.

Man kann nun auch die Gleichung $f$ten Grades bilden, deren Wurzeln r, r_e, r_{2e} ... $r_{(f-1)e}$ sind; denn es ist für ein unbestimmtes x

$$\Phi_e(x) = (x - r)\,(x - r_e) \ldots (x - r_{(f-1)e}) \tag{18}$$

eine Funktion im Körper $R(\eta)$. Ihre Koeffizienten sind in der Form $\varphi(r)$, (5), mit ganzzahligen b enthalten und lassen sich

also in der Form (14) mit ganzzahligen a darstellen. Sie lassen sich leicht aus den Newtonschen Formeln berechnen, durch die die Koeffizienten einer Gleichung mittels der Potenzsummen der Wurzeln ausgedrückt werden. Denn es ist nach (12) und (3)

$$r^{g^h} + r_e^{g^h} + \cdots + r_{(f-1)e}^{g^h} = \eta_h,$$

also geradezu einer der Potenzsummen gleich, und jede Potenzsumme ist einer der Perioden gleich. Man sieht daher aus den Newtonschen Formeln zwar unmittelbar, daß die Koeffizienten a rational sind, nicht aber, daß es ganze Zahlen sind.

Die Gleichung (18) läßt sich aber ebenso behandeln, wie die Gleichung (2); denn die Gruppe von $\Phi_e(x) = 0$ ist C_e, und wenn e' ein Teiler von f ist, etwa $f = e'f'$, so ist $C_{ee'}$ ein Teiler von C_e. Zu $C_{ee'}$ gehört aber die Periode

$$\eta' = r + r_{ee'} + r_{2ee'} + \cdots + r_{(f'-1)ee'},$$

die also von einer Gleichung des Grades e' im Körper $R(\eta)$ abhängt.

Wenn man die Zahlen e, e', ... als Primzahlen annimmt, so besteht die Reihe der Resolventen aus einer Reihe von Gleichungen, deren Grade die Primfaktoren von $n - 1$ sind.

Was die Gruppe der Gleichung $e$ten Grades betrifft, deren Wurzeln die e Größen η sind, so finden wir diese sehr einfach aus der Bemerkung, daß die Substitution (r, r_h) unter den η_k die Substitution (η_k, η_{k+h}) hervorruft, wenn der Index von η nach dem Modul e genommen wird. Die Gruppe von $F_e(x) = 0$ besteht daher aus den Potenzen der zyklischen Permutation

$$(\eta, \eta_1, \eta_2 \cdots \eta_{e-1}).$$

§ 75.

Die Gaußsche Methode zur Berechnung der Resolventen.

Um in konkreten Fällen die Auflösung der Kreisteilungsgleichung wirklich durchzuführen, kommt es nach dem, was wir im vorigen Paragraphen entwickelt haben, nur darauf an, die Produkte je zweier konjugierter Perioden als lineare Funktionen derselben Perioden darzustellen. Dafür hat Gauß ein einfaches Verfahren angegeben, das wir jetzt kennen lernen wollen. Wir müssen dazu die Bezeichnung der Perioden ein wenig verändern.

Wir schreiben zwei beliebige der Perioden § 74, (12) in der Weise:

$$\eta^{(\lambda)} = r^{\lambda} + r^{\lambda'} + r^{\lambda''} + \cdots$$
(1)
$$\eta^{(\mu)} = r^{\mu} + r^{\mu'} + r^{\mu''} + \cdots$$

so daß, wenn $\lambda \equiv g^h \pmod{n}$ ist, $\eta^{(\lambda)} = \eta_h$ oder auch

$$\eta^{(\lambda)} = \eta_{\mathrm{ind}\,\lambda};$$

es ist dann

$$\lambda' \equiv \lambda g^e, \quad \lambda'' \equiv \lambda g^{2e}, \ldots$$
(2)
$$\mu' \equiv \mu g^e, \quad \mu'' \equiv \mu g^{2e}, \ldots \pmod{n};$$

lassen wir also die Zeichen s und t je ein volles Restsystem nach dem Modul f durchlaufen, so ist

$$\eta^{(\lambda)} = \sum^{s} r^{\lambda g^{se}}, \qquad \eta^{(\mu)} = \sum^{t} r^{\mu g^{te}}.$$

Das Produkt davon ist

$$\eta^{(\lambda)}\eta^{(\mu)} = \sum^{s}\sum^{t} r^{\lambda g^{se} + \mu g^{te}}.$$

Halten wir bei der Bildung dieser Summe zunächst s fest, und summieren in bezug auf t, so dürfen wir t durch $t + s$ ersetzen, weil beide gleichzeitig ein volles Restsystem nach dem Modul f durchlaufen. Demnach wird

$$\eta^{(\lambda)}\eta^{(\mu)} = \sum^{s}\sum^{t} r^{(\lambda + \mu g^{te})g^{se}}.$$

Hierin aber darf die Reihenfolge der Summation vertauscht werden, so daß auch

$$\eta^{(\lambda)}\eta^{(\mu)} = \sum^{t}\sum^{s} r^{(\lambda + \mu g^{te})g^{se}}$$
(3)

ist. Die nach s genommene Summe

$$\sum^{s} r^{(\lambda + \mu g^{te})g^{se}}$$

ist aber selbst eine der Perioden, nämlich nach der Bezeichnung (1) die Periode

$$\eta^{(\lambda + \mu g^{te})}.$$

Es ergibt sich also aus (2) und (3)

$$\eta^{(\lambda)}\eta^{(\mu)} = \eta^{(\lambda+\mu)} + \eta^{(\lambda+\mu')} + \eta^{(\lambda+\mu'')} + \cdots$$
(4)

Die rechte Seite dieses Ausdruckes enthält f Glieder; darunter kann aber dieselbe Periode mehrmals auftreten; auch kann darunter die uneigentliche Periode $\eta^{(0)}$ vorkommen, die gleich der ganzen Zahl f zu setzen ist. Will man die homogene Form des Ausdruckes wieder herstellen, so benutzt man die Relation

$$\eta + \eta_1 + \eta_2 + \cdots + \eta_{e-1} = -1,$$

die ja nur eine andere Schreibweise der Gleichung § 74, (2) ist.

Wenn man nach (4) den Ausdruck für ein Produkt $\eta\,\eta_h$ berechnet hat, dem wir schon oben (S. 338) die Form gegeben haben:

$$\eta\,\eta_h = a_{0,h}\,\eta + a_{1,h}\,\eta_1 + \cdots + a_{e-1,h}\,\eta_{e-1}, \tag{5}$$

so erhält man das Produkt von zwei beliebigen der Perioden durch die Substitution $(\eta,\ \eta_k)$,

$$\eta_k\,\eta_{h+k} = a_{0,h}\,\eta_k + a_{1,h}\,\eta_{k+1} + \cdots + a_{e-1,h}\,\eta_{k-1}. \tag{6}$$

Um an einem einfachen Beispiel diese Regeln zu erläutern, nehmen wir $n = 13$. Für 13 ist 2 eine primitive Wurzel, und wir können die im § 70 gegebene Indextabelle anwenden:

(7)

I	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11
N	1,	2,	4,	8,	3,	6,	12,	11,	9,	5,	10,	7

Wenn wir zuerst $e = 3$, $f = 4$ annehmen, so erhalten wir eine kubische Resolvente, deren Wurzeln

$$\begin{aligned} \eta &= r + r_3 + r_6 + r_9 \\ \eta_1 &= r_1 + r_4 + r_7 + r_{10} \\ \eta_2 &= r_2 + r_5 + r_8 + r_{11} \end{aligned} \tag{8}$$

sind. Aus der Tabelle (7) ergibt sich dafür auch

$$\begin{aligned} \eta &= r + r^{-5} + r^{-1} + r^5 = \eta^{(1)} \\ \eta_1 &= r^2 + r^3 + r^{-2} + r^{-3} = \eta^{(2)} \\ \eta_2 &= r^4 + r^6 + r^{-4} + r^{-6} = \eta^{(4)}. \end{aligned} \tag{9}$$

Wendet man die Formel (4) an, so erhält man z. B.:

$$\begin{aligned} \eta\,\eta &= \eta^{(2)} + \eta^{(-4)} + \eta^{(0)} + \eta^{(6)} = 4 + \eta^{(2)} + 2\,\eta^{(4)} \\ &= -4\,\eta - 3\,\eta_1 - 2\,\eta_2, \end{aligned}$$

und so ergeben sich die Formeln:

$$\begin{aligned} \eta^2 &= -4\,\eta - 3\,\eta_1 - 2\,\eta_2, \\ \eta\,\eta_1 &= \eta + 2\,\eta_1 + \eta_2, \\ \eta\,\eta_2 &= 2\,\eta + \eta_1 + \eta_2, \end{aligned}$$

woraus nach § 74, (17) die Gleichung für η

$$\begin{vmatrix} -4-\eta, & -3, & -2 \\ 1, & 2-\eta, & 1 \\ 2, & 1, & 1-\eta \end{vmatrix} = 0,$$

oder

$$\eta^3 + \eta^2 - 4\,\eta + 1 = 0 \tag{10}$$

folgt. Die Diskriminante dieser Gleichung ergibt sich gleich $169 = 13^2$, also positiv. Die Gleichung hat folglich drei reelle,

und zwar, da das letzte Glied positiv und das vorletzte negativ ist, zwei positive und eine negative Wurzel.

Setzen wir $r = e^{\frac{2\pi i}{13}}$, so wird

$$\eta = 2\cos\frac{2\pi}{13} + 2\cos\frac{10\pi}{13} = 2\left(\cos\frac{2\pi}{13} - \cos\frac{3\pi}{13}\right),$$

$$\eta_1 = 2\cos\frac{4\pi}{13} + 2\cos\frac{6\pi}{13} = 2\left(\cos\frac{4\pi}{13} + \cos\frac{6\pi}{13}\right),$$

$$\eta_2 = 2\cos\frac{8\pi}{13} + 2\cos\frac{12\pi}{13} = -2\left(\cos\frac{\pi}{13} + \cos\frac{5\pi}{13}\right),$$

oder auch

$$\eta = 4\cos\frac{4\pi}{13}\cos\frac{6\pi}{13}, \quad \eta_1 = 4\cos\frac{\pi}{13}\cos\frac{5\pi}{13},$$

$$\eta_2 = -4\cos\frac{2\pi}{13}\cos\frac{3\pi}{13}.$$

Es ist also η_2 die negative, η_1 die größere, η die kleinere der beiden positiven Wurzeln.

Die Gleichung vierten Grades, deren Wurzeln r, r^{-1}, r^5, r^{-5} sind, läßt sich nun leicht bilden, wenn man

$$\xi = r + r^{-1} = 2\cos\frac{2\pi}{13}, \quad \xi' = r^5 + r^{-5} = 2\cos\frac{10\pi}{13}$$

setzt. Es ist nämlich

$$\xi + \xi' = \eta, \quad \xi\xi' = \eta_2,$$

also

$$\xi^2 - \eta\xi + \eta_2 = 0 \tag{11}$$

die quadratische Gleichung für ξ, aus der man die biquadratische Gleichung für r erhält,

$$r^4 - \eta r^3 + (\eta_2 + 2)r^2 - \eta r + 1 = 0. \tag{12}$$

Die quadratische Gleichung (11) hat eine positive und eine negative Wurzel. Die positive Wurzel ist $2\cos\frac{2\pi}{13}$. Anstatt r selbst zu berechnen, berechnet man auch $2\sin\frac{2\pi}{13}$, was man als die positive Quadratwurzel $\sqrt{4 - \xi^2}$ erhält.

An Stelle der drei viergliedrigen Perioden η hätte man auch zuerst zwei Perioden von sechs Gliedern betrachten können:

$$\begin{aligned} \xi &= r + r^3 + r^4 + r^{-1} + r^{-3} + r^{-4} \\ \xi_1 &= r^2 + r^5 + r^6 + r^{-2} + r^{-5} + r^{-6}. \end{aligned} \tag{13}$$

Für das Produkt $\zeta\zeta_1$ findet man nach der Formel (4) den Wert $3(\zeta + \zeta_1)$ oder -3, so daß ζ, ζ_1 die Wurzeln der quadratischen Gleichung

$$\zeta^2 + \zeta - 3 = 0 \tag{14}$$

sind, woraus, da aus dem Ausdruck durch die Kosinus leicht zu sehen ist, daß ζ positiv ist:

$$\zeta = \frac{-1+\sqrt{13}}{2}, \quad \zeta_1 = \frac{-1-\sqrt{13}}{2}. \tag{15}$$

Nach Adjunktion der Werte ζ, ζ_1 kann man eine kubische Gleichung für die zweigliedrigen Perioden bilden. Setzt man nämlich

$$\begin{aligned} \xi &= r + r^{-1}, & \xi_1 &= r^2 + r^{-2}, \\ \xi_2 &= r^4 + r^{-4}, & \xi_3 &= r^5 + r^{-5}, \\ \xi_4 &= r^3 + r^{-3}, & \xi_5 &= r^6 + r^{-6}, \end{aligned}$$

so folgt

$$\zeta = \xi + \xi_2 + \xi_4, \quad \zeta_1 = \xi_1 + \xi_3 + \xi_5, \quad \zeta + \zeta_1 = -1,$$

ferner

$$\xi\xi_2 = \xi_3 + \xi_4, \quad \xi\xi_4 = \xi_1 + \xi_2, \quad \xi_2\xi_4 = \xi + \xi_5,$$
$$\xi\xi_2\xi_4 = 2 + \zeta_1 = 1 - \zeta,$$

also sind ξ, ξ_2, ξ_4 die Wurzeln der kubischen Gleichung

$$x^3 - \zeta x^2 - x - 1 + \zeta = 0. \tag{16}$$

§ 76.

Zurückführung der Kreisteilungsgleichung auf reine Gleichungen.

Gauß hat schon in seiner ersten Darstellung in den Disq. arithm. die Kreisteilungsgleichungen durch Benutzung der Resolventen direkt auf reine Gleichungen zurückgeführt[1]). Um dies durchzuführen, setzen wir

$$n - 1 = m,$$

und bezeichnen mit ε eine primitive Einheitswurzel $m$ten Grades, mit r, r_1, r_2, ... r_{m-1} die $n$ten Einheitswurzeln in der oben (S. 335)

[1]) Gauß, disq. arithm. art. 359, 360; disq. circa aequationes puras ulterior evolutio, Werke, Bd. II. Lagrange, rés. des équations numériques. Jakobi, „Über die Kreisteilung und ihre Anwendung in der Zahlentheorie". Werke, Bd. 6. Kummer, Crelles Journal, Bd. 35 und Abhandl. d. Berl. Akademie 1856. Zu erwähnen sind noch Eisenstein und Cauchy. Die Lehre von der Kreisteilung ist im Zusammenhange dargestellt und von den historischen Nachweisen begleitet in dem Buche von Bachmann, „Die Lehre von der Kreisteilung". Leipzig 1872. Siehe auch Bachmann, Über Gauß' zahlentheoretische Arbeiten. Göttinger Nachrichten 1911.

festgesetzten Reihenfolge. Die Lagrangeschen Resolventen sind dann

$$(\varepsilon^\lambda, r) = \sum_{0,\, n-2}^{h} \varepsilon^{\lambda h} r_h = \sum_{0,\, n-2}^{h} \varepsilon^{\lambda h} r^{g^h},$$

worin λ ein beliebiger, nach dem Modul m genommener Exponent ist.

Setzen wir $g^h = s$, also $h \equiv \operatorname{ind} s \pmod m$, so können wir auch setzen

$$(1) \qquad (\varepsilon^\lambda, r) = \sum_{1,\, n-1}^{s} \varepsilon^{\lambda \operatorname{ind} s} r^s.$$

Als Grundlage für die weitere Reduktion dieser Ausdrücke dient die Berechnung des Produktes zweier solcher Resolventen. Wir bezeichnen mit μ eine zweite Zahl wie λ und verstehen unter s und t zwei Zeichen, die voneinander unabhängig je ein Restsystem nach dem Modul n durchlaufen, mit Ausschluß der Null.

Dann ist

$$(\varepsilon^\lambda, r)\,(\varepsilon^\mu, r) = \sum^{s} \sum^{t} r^{(s+t)}\, \varepsilon^{\lambda \operatorname{ind} s + \mu \operatorname{ind} t}.$$

Wenn die Summation in bezug auf t zuerst ausgeführt wird, so kann st an Stelle von t gesetzt werden, da s von Null verschieden ist, und also st und t zugleich ein volles Restsystem nach dem Modul n durchlaufen. Also wird

$$(2) \qquad (\varepsilon^\lambda, r)\,(\varepsilon^\mu, r) = \sum^{s} \sum^{t} r^{s(t+1)}\, \varepsilon^{(\lambda+\mu) \operatorname{ind} s}\, \varepsilon^{\mu \operatorname{ind} t}.$$

Wir erledigen zuerst den speziellen Fall $\mu = -\lambda$, den wir aus (2) erhalten, wenn wir die Summation nach s zuerst ausführen:

$$(\varepsilon^\lambda, r)\,(\varepsilon^{-\lambda}, r) = \sum^{t} \varepsilon^{-\lambda \operatorname{ind} t} \sum^{s} r^{s(t+1)}.$$

Hierin ist

$$\sum^{s} r^{s(t+1)} = -1, \text{ wenn } t = 1, 2 \ldots n-2$$
$$= n-1, \text{ wenn } t = n-1,$$

also, da [nach § 70 (21)] $\operatorname{ind}(n-1) = \frac{1}{2}(n-1)$ und

$$\varepsilon^{\frac{n-1}{2}} = -1$$

ist,

$$(\varepsilon^\lambda, r)\,(\varepsilon^{-\lambda}, r) = -\sum_{1,\, n-1}^{t} \varepsilon^{-\lambda \operatorname{ind} t} + (-1)^\lambda n.$$

In der nach t genommenen Summe durchläuft $\operatorname{ind} t$ ein volles Restsystem nach dem Modul $n-1$, und also ist, wenn λ durch

$n-1$ nicht teilbar angenommen wird,

$$\sum_{1,n-1}^{t} \varepsilon^{-\lambda \operatorname{ind} t} = 0,$$

und daher

$$(\varepsilon^{\lambda}, r)\,(\varepsilon^{-\lambda}, r) = (-1)^{\lambda} n, \tag{3}$$

während, wenn λ durch $n-1$ teilbar ist, $(\varepsilon^{\lambda}, r) = (1, r) = -1$ wird.

Wir behandeln also nun die Summe (2) weiter unter der Voraussetzung, daß $\lambda + \mu$ nicht durch $(n-1)$ teilbar ist. Dann ist $\sum_{1,n-1}^{s} \varepsilon^{(\lambda+\mu)\operatorname{ind} s} = 0$, und es können in der Summe auf der rechten Seite von (2) die dem Werte $t = n-1$ entsprechenden Glieder weggelassen werden. Man erhält so

$$\begin{aligned}(\varepsilon^{\lambda}, r)\,(\varepsilon^{\mu}, r) &= \sum_{1,n-2}^{t} \varepsilon^{\mu \operatorname{ind} t} \sum_{1,n-1}^{s} \varepsilon^{(\lambda+\mu)\operatorname{ind} s}\, r^{s(t+1)} \\ &= \sum_{1,n-2}^{t} \varepsilon^{\mu \operatorname{ind} t - (\lambda+\mu)\operatorname{ind}(t+1)} \sum_{1,n-1}^{s} \varepsilon^{(\lambda+\mu)\operatorname{ind} s(t+1)}\, r^{s(t+1)}.\end{aligned} \tag{4}$$

Nun durchläuft $s(t+1)$ bei feststehendem t zugleich mit s ein volles Restsystem nach dem Modul n, und es ist also

$$\sum_{1,n-1}^{s} \varepsilon^{(\lambda+\mu)\operatorname{ind} s(t+1)}\, r^{s(t+1)} = \sum_{1,n-1}^{s} \varepsilon^{(\lambda+\mu)\operatorname{ind} s}\, r^{s} = (\varepsilon^{\lambda+\mu}, r),$$

und die Formel (4) ergibt

$$\frac{(\varepsilon^{\lambda}, r)\,(\varepsilon^{\mu}, r)}{(\varepsilon^{\lambda+\mu}, r)} = \psi_{\lambda,\mu}(\varepsilon), \tag{5}$$

wenn zur Abkürzung

$$\sum_{1,n-2}^{t} \varepsilon^{\mu \operatorname{ind} t - (\lambda+\mu)\operatorname{ind}(t+1)} = \psi_{\lambda,\mu}(\varepsilon) \tag{6}$$

gesetzt und $\lambda + \mu$ durch $n - 1 = m$ nicht teilbar angenommen wird. Diese Funktionen ψ sind dann aus den $(n-1)$ten Einheitswurzeln ε mit ganzzahligen Koeffizienten zusammengesetzt, und sind also bekannt, wenn die $(n-1)$ten Einheitswurzeln als bekannt vorausgesetzt werden. Es sind Zahlen des Körpers $R(\varepsilon)$.

Wir nehmen jetzt λ durch μ teilbar an, ersetzen also λ durch $\mu\lambda$ und setzen

$$\alpha = \varepsilon^{\mu}, \tag{7}$$

so daß α auch eine nicht primitive $(n-1)$te Einheitswurzel, z. B. eine $e$te Einheitswurzel sein kann (wenn, wie oben, $m = ef$ ist und μ durch f teilbar, etwa $= f$ genommen wird), wobei jedoch der Fall $\alpha = 1$, also $e = 1$ auszuschließen ist.

Dann setzen wir

$$(8)\qquad \sum_{1,\,n-2}^{t} \alpha^{\mathrm{ind}\, t-(\lambda+1)\,\mathrm{ind}\,(t+1)} = \psi_\lambda(\alpha);$$

darin ist $\psi_\lambda(\alpha)$ eine Zahl des Körpers $R(\alpha)$, die mit Hilfe einer Indextabelle leicht berechnet werden kann; λ kann dabei nach dem Modul e genommen werden. Die Formel (5) gibt für diese Annahme

$$(9)\qquad (\alpha^\lambda, r)\ (\alpha, r) = (\alpha^{\lambda+1}, r)\,\psi_\lambda(\alpha),$$

und diese Formel gilt, solange $\mu(\lambda + 1)$ nicht durch $n - 1$ teilbar ist, also, wenn μ und m teilerfremd, und daher α eine primitive $m$te Einheitswurzel ist, für $\lambda = 1, 2, \ldots n - 3$, und wenn f der größte gemeinschaftliche Teiler von μ und m und $m = ef$ ist, für $\lambda = 1, 2, \ldots e - 2$.

Die Formel (3) gibt jetzt, wenn α^λ nicht $= 1$ ist,

$$(10)\qquad (\alpha^\lambda, r)\ (\alpha^{-\lambda}, r) = (-1)^{u\lambda} n.$$

Die Formeln (9), (10) genügen, um nicht nur die r selbst, sondern auch die sämtlichen Perioden η durch Radikale zu bestimmen, deren Grade die in m aufgehenden Primzahlen sind, und unter denen die Funktionen ψ_λ als die bekannten Größen vorkommen.

Da die f-gliedrigen Perioden (für $f = 1$) die Größen r mit umfassen, so wollen wir gleich zur Bestimmung dieser Perioden übergehen.

Wir wollen also unter α eine primitive $e$te Einheitswurzel verstehen und $\mu = f$ setzen.

Dann ist

$$(\alpha^\lambda, r) = r + \alpha^\lambda r_1 + \alpha^{2\lambda} r_2 + \cdots + \alpha^{(n-2)\lambda} r_{n-2},$$

oder nach § 74, (12):

$$(11)\qquad (\alpha^\lambda, r) = (\alpha^\lambda, \eta) = \eta + \alpha^\lambda \eta_1 + \alpha^{2\lambda}\eta_2 \ldots + \alpha^{(e-1)\lambda}\eta_{e-1},$$

und die Formeln (9) und (10) ergeben:

$$(12)\qquad (\alpha^\lambda, \eta)\ (\alpha, \eta) = (\alpha^{\lambda+1}, \eta)\,\psi_\lambda(\alpha)$$

$$(13)\qquad (\alpha^\lambda, \eta)\ (\alpha^{-\lambda}, \eta) = (-1)^{f\lambda} n. \qquad \lambda = 1, 2, \ldots (e-2).$$

Die Gleichung (13) zeigt, daß keine der Zahlen (α^λ, η) verschwindet, und nach (12) sind auch die $\psi_\lambda(\alpha)$ von Null verschieden, und wenn wir die Gleichung (12) für $\lambda = 1, 2, \ldots \lambda - 1$ bilden, so folgt:

$$\begin{aligned}
(\alpha, \eta)\ (\alpha, \eta) &= (\alpha^2, \eta)\,\psi_1(\alpha)\\
(\alpha^2, \eta)\ (\alpha, \eta) &= (\alpha^3, \eta)\,\psi_2(\alpha)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
(\alpha^{\lambda-1}, \eta)\ (\alpha, \eta) &= (\alpha^\lambda, \eta)\,\psi_{\lambda-1}(\alpha).
\end{aligned}$$

Daraus erhalten wir durch Multiplikation und Wegheben des Faktors $(\alpha^2, \eta) \dots (\alpha^{\lambda-1}, \eta)$:

$$(14) \qquad (\alpha, \eta)^\lambda = (\alpha^\lambda, \eta)\, \psi_1(\alpha)\, \psi_2(\alpha) \dots \psi_{\lambda-1}(\alpha).$$

Dadurch ist (α^λ, η) rational durch (α, η) ausgedrückt. Setzen wir aber in (14) $\lambda = e - 1$, so ergibt sich:

$$(\alpha, \eta)^{e-1} = (\alpha^{-1}, \eta)\, \psi_1(\alpha)\, \psi_2(\alpha) \dots \psi_{e-2}(\alpha),$$

und durch Multiplikation mit (α, η), wenn (13) in der Form

$$(\alpha, \eta)\ (\alpha^{-1}, \eta) = (-1)^f n$$

benutzt wird,

$$(15) \qquad (\alpha, \eta)^e = (-1)^f n\, \psi_1(\alpha)\, \psi_2(\alpha) \dots \psi_{e-2}(\alpha),$$

wodurch die Bildung von (α, η) auf eine $e$te Wurzel zurückgeführt ist. Setzen wir $e = m$, so erhalten wir (ε, r).

Wir wollen dies Verfahren auf den interessanten Fall $n = 17$ anwenden, der zu dem geometrisch merkwürdigen, von Gauß gefundenen Resultat führt, daß die Teilung der Kreisperipherie in 17 Teile von einer Reihe quadratischer Gleichungen abhängt, und daß daher das reguläre Siebzehneck mit Zirkel und Lineal konstruiert werden kann.

Für die Primzahl 17 ist 3 eine primitive Wurzel, und man findet die Indextabelle:

l	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	1	3	9	10	13	5	15	11	16	14	8	7	4	12	2	6

Die zweigliedrigen Perioden sind, wenn wir $r = e^{\frac{2\pi i}{17}}$ setzen:

$$(16) \quad \begin{aligned} \eta &= 2\cos\frac{2\pi}{17}, & \eta_1 &= 2\cos\frac{6\pi}{17}, \\ \eta_2 &= 2\cos\frac{18\pi}{17} = -2\cos\frac{\pi}{17}, & \eta_3 &= 2\cos\frac{20\pi}{17} = -2\cos\frac{3\pi}{17}, \\ \eta_4 &= 2\cos\frac{26\pi}{17} = 2\cos\frac{8\pi}{17}, & \eta_5 &= 2\cos\frac{10\pi}{17} = -2\cos\frac{7\pi}{17}, \\ \eta_6 &= 2\cos\frac{30\pi}{17} = 2\cos\frac{4\pi}{17}, & \eta_7 &= 2\cos\frac{22\pi}{17} = -2\cos\frac{5\pi}{17}. \end{aligned}$$

Zur Berechnung der Funktionen ψ wendet man am einfachsten folgende Tabelle an:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ind t	0	14	1	12	5	15	11	10	2	3	7	13	4	9	6
ind $(t+1)$.	14	1	12	5	15	11	10	2	3	7	13	4	9	6	8

woraus man, wenn α eine achte Einheitswurzel bedeutet, nach der Formel (8) findet:

$$\begin{aligned}
&\psi_1(\alpha) = \psi_6(\alpha) = 2\alpha + 2\alpha^2 + 3\alpha^4 + 4\alpha^5 + 2\alpha^6 + 2\alpha^7\\
(17)\quad &\psi_2(\alpha) = \psi_5(\alpha) = 2 + 3\alpha + \alpha^3 + \alpha^4 + 3\alpha^5 + 4\alpha^6 + \alpha^7\\
&\psi_3(\alpha) = \psi_4(\alpha) = 3 + 3\alpha + 2\alpha^2 + 3\alpha^3 + \alpha^5 + 2\alpha^6 + \alpha^7.
\end{aligned}$$

Nimmt man

$$\alpha = \frac{1+i}{\sqrt{2}}, \quad \alpha^2 = i, \quad \alpha^4 = -1,$$

so ergibt sich:

$$\begin{aligned}
&\psi_1(\alpha) = \psi_6(\alpha) = -3 - i\sqrt{8}\\
&\psi_2(\alpha) = \psi_5(\alpha) = 1 - 4i\\
&\psi_3(\alpha) = \psi_4(\alpha) = 3 + i\sqrt{8}\\
&\psi_1(i) = -1 + 4i,
\end{aligned} \tag{18}$$

also nach (15):

$$(\alpha, \eta)^8 = 17(3 + i\sqrt{8})^4 (1 - 4i)^2. \tag{19}$$

Aus (13) und (14) aber erhält man:

$$\begin{aligned}
(-1, \eta)^2 &= 17,\\
(i, \eta)^2 &= (-1, \eta)(-1 + 4i),\\
(\alpha, \eta)^2 &= -(i, \eta)(3 + i\sqrt{8});
\end{aligned} \tag{20}$$

außerdem hat man $(1, \eta) = -1$,

$$\begin{aligned}
(i, \eta)(-i, \eta) &= 17,\\
(\alpha, \eta)(\alpha^{-1}, \eta) &= 17,\\
(\alpha^3, \eta)(\alpha^{-3}, \eta) &= 17,
\end{aligned} \tag{21}$$

und nach (9)

$$(\alpha, \eta)(\alpha^3, \eta) = (-1, \eta)(3 + i\sqrt{8}), \tag{22}$$

wodurch $(1, \eta)$, $(-1, \eta)$, $(\pm i, \eta)$, (α, η), (α^{-1}, η), (α^3, η), (α^{-3}, η) durch Quadratwurzeln bestimmt sind.

Bei der Bestimmung, die wir über r und α getroffen haben, ergibt sich leicht aus der Betrachtung der Werte der Kosinus (16), daß $(-1, \eta)$ positiv, also gleich der positiven Quadratwurzel $\sqrt{17}$ ist, und daß ferner (i, η) und (α, η) positive reelle Teile haben, wodurch auch diese Größen völlig bestimmt sind.

Es ist also

$$(-1, \eta) = \sqrt{17},$$

$$(i, \eta) = \sqrt[4]{17}\left\{\sqrt{\frac{\sqrt{17}-1}{2}} + i\sqrt{\frac{\sqrt{17}+1}{2}}\right\},$$

wenn alle Wurzeln positiv genommen werden, und nach der letzten Gleichung (20) läßt sich durch eine, wenn auch etwas lange Formel (α, η) durch Quadratwurzeln aus reellen Größen darstellen. Auf die geometrischen Konstruktionen, die sich hier anknüpfen, gehen wir nicht ein[1]).

Die benutzte Vorzeichenbestimmung erhält man einfach wenn man nach (16) setzt:

$$\frac{1}{2}(-1, \eta) = \cos\frac{2\pi}{17} - \cos\frac{\pi}{17} + \cos\frac{8\pi}{17} + \cos\frac{4\pi}{17}$$
$$- \cos\frac{6\pi}{17} + \cos\frac{3\pi}{17} + \cos\frac{7\pi}{17} + \cos\frac{5\pi}{17}.$$

Nach einer bekannten trigonometrischen Formel ist aber

$$\cos\frac{2\pi}{17} - \cos\frac{\pi}{17} = -2\sin\frac{\pi}{34}\sin\frac{3\pi}{34}$$
$$-\cos\frac{6\pi}{17} + \cos\frac{5\pi}{17} = \quad 2\sin\frac{\pi}{34}\sin\frac{11\pi}{34},$$

und da $\sin\frac{11\pi}{34}$ größer ist als $\sin\frac{3\pi}{34}$, so ist die Summe dieser beiden Ausdrücke und damit die ganze Summe $(-1, \eta)$ positiv.

Es ergibt sich ferner für den reellen Teil von (i, η):

$$\eta - \eta_2 + \eta_4 - \eta_6 = 2\cos\frac{2\pi}{17} + 2\cos\frac{\pi}{17} + 2\cos\frac{8\pi}{17} - 2\cos\frac{4\pi}{17},$$

was unmittelbar als positiv erkannt wird; endlich für den reellen Teil von (α, η):

$$\eta - \eta_4 + \frac{1}{\sqrt{2}}(\eta_1 - \eta_3 - \eta_5 + \eta_7)$$
$$= 2\cos\frac{2\pi}{17} - 2\cos\frac{8\pi}{17} + \sqrt{2}\left(\cos\frac{6\pi}{17} + \cos\frac{3\pi}{17} + \cos\frac{7\pi}{17} - \cos\frac{5\pi}{17}\right),$$

was sich gleichfalls als positiv erweist.

§ 77.
Besondere Perioden.

Im § 74 ist gezeigt, daß, wenn $n - 1 = ef$ ist, die Perioden $\eta, \eta_1, \ldots \eta_{e-1}$ einer ganzzahligen Gleichung $e$ten Grades genügen. Diese Gleichung hat nur reelle Wurzeln, wenn f gerade ist, weil dann $\frac{1}{2}(n-1)$ ein Vielfaches von e ist, und folglich r und

[1]) Vgl. hierüber v. Staudt, Konstruktion des regulären Siebzehnecks. Crelles Journ., Bd. 24. Gérard, Math. Annalen, Bd. 48.

r^{-1} in derselben Periode vorkommen. Ist aber f ungerade, so sind alle Wurzeln imaginär.

Es ist nun vom höchsten Interesse, diese Gleichung $e$ten Grades für einzelne besondere Werte von e ohne eine spezielle Annahme über die Primzahl n, außer der, daß $n-1$ durch e teilbar sein soll, zu untersuchen.

Wir betrachten die ersten speziellen Fälle und nehmen zunächst $e = 2$ an, was bei jeder ungeraden Primzahl n zulässig ist; η, η_1 sind hier also die zwei Perioden von $\frac{1}{2}(n-1)$ Gliedern, die wir jetzt mit A, B bezeichnen wollen, so daß

$$A = r + r_2 + r_4 + \cdots + r_{n-3}$$
$$B = r_1 + r_3 + r_5 + \cdots + r_{n-2}.$$

Die r, r_2, ... r_{n-3} haben die Exponenten g^0, g^2, ... g^{n-3}, d. h. die Exponenten von g sind gerade Zahlen. Die Exponenten von r sind also die quadratischen Reste von n (§ 72). Ebenso sind die Exponenten von r in der Summe B die Nichtreste. Bezeichnen wir also die Reste mit a, die Nichtreste mit b, so ist

(1) $$A = \Sigma r^a, \qquad B = \Sigma r^b,$$

und es ist $(-1, r) = A - B$. Diese Ausdrücke A, B werden die Gaußschen Summen genannt.

Machen wir in der Formel § 76, (10) die Annahme

$$e = 2, \quad \mu = f = \frac{n-1}{2}, \quad \lambda = 1, \quad \alpha = -1,$$

so ergibt sich:

(2) $$A - B = \pm\sqrt{(-1)^{\frac{n-1}{2}} n},$$

während andererseits

$$A + B = -1,$$

also

(3) $$2A = -1 \pm \sqrt{(-1)^{\frac{n-1}{2}} n}, \quad 2B = -1 \mp \sqrt{(-1)^{\frac{n-1}{2}} n}.$$

Das Vorzeichen, das man der Wurzel zu geben hat, hängt von der Wahl von r ab. Ist aber über r verfügt, so ist das Vorzeichen völlig bestimmt. Seine Ermittelung bietet eigentümliche Schwierigkeiten, die zu einer eigenen Abhandlung von Gauß Anlaß gegeben haben[1]). Wir gehen hier nicht näher darauf ein.

---

[1]) Summatio quarundam serierum singularium. Gauß' Werke, Bd. II. Auf andere Weise, mit Anwendung der höheren Analysis, ist das Zeichen von Dirichlet, Cauchy, Kronecker bestimmt worden. Der Weg, den

Wir gehen jetzt zu dem Falle $e = 3$ über, wobei $n - 1$ durch 3 teilbar angenommen werden muß, also $n = 7, 13, 19, 31, 37, 43 \ldots$ Wir bezeichnen mit ϱ eine imaginäre dritte Einheitswurzel:

$$\varrho = \frac{-1 + \sqrt{-3}}{2},$$

und bestimmen die drei Perioden η, η_1, η_2 von je $\frac{1}{3}(n-1)$ Gliedern, die, wie wir in § 74 gesehen haben, die Wurzeln einer zyklischen kubischen Gleichung sind. Nach § 76, (8) ist

$$\psi_1(\varrho) = \sum_{1, n-2}^{t} \varrho^{\operatorname{ind} t - 2\operatorname{ind}(t+1)},$$

wofür, da $-2 \equiv 1 \pmod 3$, auch

$$\psi_1(\varrho) = \sum_{1, n-2}^{t} \varrho^{\operatorname{ind}(t^2+t)} \tag{4}$$

gesetzt werden kann. Diese Zahl $\psi_1(\varrho)$ kann in einer der beiden Formen

$$\psi_1(\varrho) = a + b\varrho = \frac{A + b\sqrt{-3}}{2} \tag{5}$$

dargestellt werden, worin a, b, A ganze Zahlen sind und $A = 2a - b$ ist.

Es ist dann

$$\psi_1(\varrho^2) = a + b\varrho^2 = \frac{A - b\sqrt{-3}}{2}, \tag{6}$$

und die Formeln (13), (15) des § 76 ergeben:

$$\eta + \eta_1 + \eta_2 = -1, \tag{7}$$

$$(\varrho, \eta)^3 = n\psi_1(\varrho), \quad (\varrho^2, \eta)^3 = n\psi_1(\varrho^2), \tag{8}$$

$$(\varrho, \eta)(\varrho^2, \eta) = n, \quad \psi_1(\varrho)\psi_1(\varrho^2) = n, \tag{9}$$

worin

$$(\varrho, \eta) = \eta + \varrho\eta_1 + \varrho^2\eta_2, \quad (\varrho^2, \eta) = \eta + \varrho^2\eta_1 + \varrho\eta_2. \tag{10}$$

Beispielsweise erhält man für $n = 7, 13$, wenn man die primitiven Wurzeln 3, 2 zugrunde legt und die Indextabellen anwendet:

Gauß einschlägt, benutzt nur algebraische Hilfsmittel. Kronecker hat in Liouvilles Journal, Ser. 2, Bd. 1 einen Weg für die Vorzeichenbestimmung angegeben, der mit Benutzung einer Bemerkung von Dedekind (Schlömilchs Zeitschr., Bd. 15, Literaturzeitung, S. 21) gleichfalls zu einem rein algebraischen wird. Einen Weg, der durch sehr einfache Hilfsmittel, durch Benutzung elementarer goniometrischer Formeln zum Ziele führt, hat Mertens eingeschlagen (Sitzungsber. d. Berliner Akademie, 27. Februar 1896).

$n = 7$

N	1	2	3	4	5	6
I	0	2	1	4	5	3

$n = 13$

N	1	2	3	4	5	6	7	8	9	10	11	12
I	0	1	4	2	9	5	11	3	8	10	7	6

für $n = 7$, $a + b\varrho = -(1 + 3\varrho)$,
für $n = 13$, $a + b\varrho = -(4 + 3\varrho)$.

Aus diesen Formeln können wir leicht die kubische Gleichung herleiten, deren Wurzeln die η, η_1, η_2 sind. Sie hat wegen (7) die Form:

$$\eta^3 + \eta^2 + \beta\eta + \gamma = 0, \tag{11}$$

und die ganzen Zahlen β, γ sind zu bestimmen.

Führen wir die Multiplikation in (9) aus, so ergibt sich:

$$\begin{aligned} n &= \eta^2 + \eta_1^2 + \eta_2^2 - \eta\eta_1 - \eta\eta_2 - \eta_1\eta_2 \\ &= (\eta + \eta_1 + \eta_2)^2 - 3\beta, \end{aligned}$$

also

$$3\beta = -(n - 1). \tag{12}$$

Die Ausführung der Kuben in (8) ergibt, wenn wir für den Augenblick

$$s_3 = \eta^3 + \eta_1^3 + \eta_2^3, \qquad s = \eta^2\eta_1 + \eta_1^2\eta_2 + \eta_2^2\eta,$$
$$s' = \eta_1^2\eta + \eta_2^2\eta_1 + \eta^2\eta_2$$

setzen,

$$n\psi_1(\varrho) = s_3 - 6\gamma + 3s\varrho + 3s'\varrho^2.$$

Aus (11) und (12) erhält man aber $s_3 = -n - 3\gamma$ (§ 23), also

$$\begin{aligned} n[\psi_1(\varrho) + 1] &= -9\gamma + 3s\varrho + 3s'\varrho^2, \\ n[\psi_1(\varrho^2) + 1] &= -9\gamma + 3s\varrho^2 + 3s'\varrho, \end{aligned} \tag{13}$$

wozu noch, wenn man (7) in den Kubus erhebt,

$$n - 1 = -9\gamma + 3s + 3s'$$

kommt. Addiert man die drei letzten Gleichungen, so folgt

$$n[\psi_1(\varrho) + \psi_1(\varrho^2)] + 3n - 1 = -27\gamma,$$

oder endlich nach (5) und (6)

$$-27\gamma = nA + 3n - 1. \tag{14}$$

Aus dieser Relation folgt, da γ eine ganze Zahl ist,

$$A \equiv 1 \pmod 3. \tag{15}$$

Nun ist nach (5), (6) und (9)

$$n = a^2 - ab + b^2, \tag{15'}$$

oder

(16) $$4n = A^2 + 3b^2,$$

und wenn wir diesen Wert in (14) substituieren,

$$-4.27\gamma = A^3 + 3Ab^2 + 3A^2 + 9b^2 - 4.$$

Setzen wir nach (15) für den Augenblick $A = 1 + 3m$, so folgt $A^3 + 3A^2 - 4 \equiv 0 \pmod{27}$ und mithin

$$3b^2(A + 3) \equiv 0 \pmod{27},$$

und daraus

$$3b^2 \equiv 0 \pmod{27}.$$

Es ist also b durch 3 teilbar, und wenn wir $b = 3B$ setzen, so folgt aus (16):

(17) $$4n = A^2 + 27B^2.$$

Wir können auch leicht die Quadratwurzel aus der Diskriminante bilden, nämlich

(18) $$\sqrt{D} = (\eta - \eta_1)(\eta - \eta_2)(\eta_1 - \eta_2) = s - s',$$

wenn wir die beiden Gleichungen (13) subtrahieren und (5), (6) benutzen:

(19) $$\sqrt{D} = nB.$$

Die kubische Gleichung, deren Wurzeln die drei Größen η sind, hat also, was übrigens schon aus dem am Anfang dieses Paragraphen Bemerkten folgt, drei reelle Wurzeln. Sie vereinfacht sich, wenn man

$$3\eta + 1 = \xi$$

setzt, und ergibt dann

(20) $$\xi^3 - 3n\xi - nA = 0.$$

Substituiert man auf der linken Seite für ξ der Reihe nach

(21) $$-2\sqrt{n},\quad -\sqrt{n},\quad +\sqrt{n},\quad +2\sqrt{n},$$

so ergeben sich die Werte

$$-n(2\sqrt{n} + A),\quad +n(2\sqrt{n} - A),\quad -n(2\sqrt{n} + A),$$
$$+n(2\sqrt{n} - A),$$

die, weil nach (17) A absolut kleiner als $2\sqrt{n}$ ist, abwechselnde Vorzeichen haben. Es liegt also in jedem der drei durch die Werte (21) begrenzten Intervalle eine Wurzel der Gleichung (20).

Über die Frage, welche der drei Wurzeln der kubischen Gleichung (21) für $\xi = 3\eta + 1$, welche für $3\eta_1 + 1$ oder für $3\eta_2 + 1$ zu setzen ist, läßt sich allgemein so viel sagen, daß das Vorzeichen des Produktes (18) mit dem Vorzeichen von (19),

das durch die Definition (5), (6) bestimmt ist, freilich aber noch von der Wahl der primitiven Wurzel g abhängt, übereinstimmen muß. Dadurch sind von den sechs möglichen Zuordnungen drei ausgeschlossen. Welche Zuordnung unter den drei übrigen zu treffen ist, hängt von der Wahl von r ab, und würde zu ähnlichen Untersuchungen Anlaß geben, wie sie zur Bestimmung des Zeichens der Quadratwurzel in § 76[1]) nötig sind.

Wir heben noch den durch die Formeln (15′), (16) bewiesenen Satz hervor:

> Ist n eine Primzahl von der Form $3k + 1$, so ist n durch die Form $a^2 - ab + b^2$ und $4n$ durch die Form $A^2 + 27B^2$ darstellbar, wo a, b, A, B ganze Zahlen sind.

Daraus ergibt sich noch, daß n auch in der Form $x^2 + 3y^2$ darstellbar ist. Denn wenn von den beiden Zahlen a, b eine, etwa a, gerade ist, so ist

$$n = \left(\frac{a}{2} - b\right)^2 + \tfrac{3}{4}a^2,$$

und wenn a und b beide ungerade sind, so ist

$$n = \tfrac{1}{4}(a + b)^2 + \tfrac{3}{4}(a - b)^2.$$

Auch die zweigliedrigen Perioden der neunten Einheitswurzeln genügen einer kubischen Gleichung, und weil diese Gleichung bei der allgemeinen Theorie der kubischen Kreisteilungskörper eine wichtige Rolle spielt, wollen wir sie der Vollständigkeit halber hier betrachten, obwohl sie eigentlich in ein allgemeineres Gebiet gehört, in dem die Grade der Einheitswurzeln keine Primzahlen mehr sind.

Eine neunte Einheitswurzel r genügt der Gleichung sechsten Grades:

$$r^6 + r^3 + 1 = 0, \tag{22}$$

und wir haben drei konjugierte zweigliedrige Perioden

$$\eta = r + r^{-1}, \quad \eta' = r^2 + r^{-2}, \quad \eta'' = r^4 + r^{-4},$$

die die Wurzeln der kubischen Gleichung

$$\eta^3 - 3\eta + 1 = 0 \tag{23}$$

sind. Ist ϱ eine dritte Einheitswurzel, so können wir $\varrho = r^3$ setzen, und erhalten die Resolventen

---

[1]) Kummer, Journ. f. Mathematik, Bd. 32.

(24) $(\varrho,\eta) = \eta + \varrho\,\eta' + \varrho^2\eta''$, $(\varrho^{-1},\eta) = \eta + \varrho^{-1}\eta' + \varrho^{-2}\eta''$,

für die man mit Benutzung von (22) erhält:

$$(\varrho,\ \eta) = 3r,\quad (\varrho^{-1},\ \eta) = 3r^{-1},$$

also

$$(\varrho,\ \eta)^3 = 27\,\varrho, \tag{25}$$

$$(\varrho,\ \eta)\ (\varrho^{-1},\ \eta) = 9, \tag{26}$$

was den Formeln (8) und (9) entspricht.

Wir gehen noch in der Kürze auf den Fall $e = 4$ ein, der bei den Primzahlen n von der Form $4f + 1$ (f eine ganze Zahl) eintritt, also bei $n = 5, 13, 17, 29, 37, 41 \ldots$

Die vier Perioden von je f Gliedern seien wieder η, η_1, η_2, η_3. Für α in § 76 haben wir i zu setzen, erhalten also:

$$\begin{aligned}(i,\ \eta) &= \eta + i\eta_1 - \eta_2 - i\eta_3,\\ (-1,\ \eta) &= \eta - \eta_1 + \eta_2 - \eta_3,\\ (-i,\ \eta) &= \eta - i\eta_1 - \eta_2 + i\eta_3,\end{aligned} \tag{27}$$

und nach § 76, (8), da $-2 \equiv +2 \pmod 4$ ist,

$$\begin{aligned}\psi_1(i) &= \Sigma\, i^{\mathrm{ind}\,t + 2\,\mathrm{ind}\,(t+1)}\\ \psi_2(i) &= \Sigma\, i^{\mathrm{ind}\,(t^2+t)} = a + bi,\end{aligned} \tag{28}$$

worin a und b ganze Zahlen sind. $\psi_1(i)$ hat dieselbe Form wie $\psi_2(i)$. Wir werden aber sogleich die eine Funktion auf die andere zurückführen.

Da nun $\eta + \eta_2$, $\eta_1 + \eta_3$ die beiden $2f$-gliedrigen Perioden sind, so ist nach (2)

$$(-1,\ \eta) = \eta + \eta_2 - \eta_1 - \eta_3 = \sqrt{n}, \tag{29}$$

und also, da $\eta + \eta_1 + \eta_2 + \eta_3 = -1$ ist,

$$2(\eta + \eta_2) = -1 + \sqrt{n},\quad 2(\eta_1 + \eta_3) = -1 - \sqrt{n}, \tag{30}$$

wo das Vorzeichen von $\sqrt{n}$ bei passender Annahme über r positiv genommen werden darf.

Wir haben ferner nach (10), (14), § 76

$$\begin{aligned}(i,\ \eta)\ (-i,\ \eta) &= (-1)^f n\\ (i,\ \eta)^2 &= (-1,\ \eta)\,\psi_1(i),\end{aligned} \tag{31}$$

also nach (29) und § 76, (15):

$$\begin{aligned}(i,\ \eta)^2 &= \psi_1(i)\sqrt{n},\quad (-i,\ \eta)^2 = \psi_1(-i)\sqrt{n}\\ (i,\ \eta)^4 &= (-1)^f n\,\psi_1(i)\psi_2(i) = n\,\psi_1(i)^2,\end{aligned} \tag{32}$$

woraus folgt:

$$\psi_1(i) = (-1)^f \psi_2(i) = (-1)^f (a + bi), \tag{33}$$

und hieraus nach (31):

$$\psi_1(i)\psi_1(-i) = \psi_2(i)\psi_2(-i) = a^2 + b^2 = n. \tag{34}$$

Um die biquadratische Gleichung zu bilden, deren Wurzeln $\eta, \eta_1, \eta_2, \eta_3$ sind, suchen wir zunächst die quadratische Gleichung mit den Wurzeln η, η_2 und erhalten sie aus

$$2(\eta + \eta_2) = -1 + \sqrt{n}$$
$$2(\eta - \eta_2) = (i, \eta) + (-i, \eta).$$

Quadrieren wir diese beiden Gleichungen und subtrahieren die zweite von der ersten, so folgt wegen (31), (32) und (33):

$$16\eta\eta_2 = 1 + n - 2\sqrt{n} - 2n(-1)^f - 2a(-1)^f\sqrt{n}. \tag{35}$$

Durch (30) und (35) sind aber die Koeffizienten der gesuchten quadratischen Gleichung bestimmt. Sie lautet:

$$\eta^2 + \frac{\eta}{2} + \frac{1 + n - 2n(-1)^f}{16} = \left(\frac{\eta}{2} + \frac{a(-1)^f + 1}{8}\right)\sqrt{n},$$

und die biquadratische Gleichung für die vier Perioden erhält man, wenn man beiderseits quadriert:

$$\left(\eta^2 + \frac{\eta}{2} + \frac{1 + n - 2n(-1)^f}{16}\right)^2 - n\left(\frac{\eta}{2} + \frac{a(-1)^f + 1}{8}\right)^2 = 0. \tag{36}$$

Suchen wir den Koeffizienten der ersten Potenz von η, der eine ganze Zahl sein muß, so finden wir dafür

$$\frac{1 + n - 2n(-1)^f}{16} - n\,\frac{a(-1)^f + 1}{8}; \tag{37}$$

f ist gerade, wenn $n \equiv 1$, ungerade, wenn $n \equiv 5 \pmod 8$ ist, demnach ist

$$1 + n - 2n(-1)^f \equiv 0 \pmod 8.$$

Daraus aber folgt wegen (37), daß $a(-1)^f + 1$ durch 4 teilbar sein muß, also

$$a \equiv -(-1)^f \pmod 4, \tag{38}$$

und daraus ist nach (34) weiter zu schließen, daß b gerade sein muß, also

$$b \equiv 0 \pmod 2. \tag{39}$$

Die biquadratische Gleichung nimmt eine einfachere Gestalt an, wenn wir

$$4\eta + 1 = \xi \tag{40}$$

setzen. Sie erhält dann nach (36) die Gestalt:

$$(41)\quad [\xi^2 + n(1 - 2(-1)^f)]^2 - 4n[\xi + (-1)^f a]^2 = 0,$$

oder wenn wir die beiden Fälle $n \equiv 1$ oder $\equiv 5 \pmod 8$ trennen:

$$(42)\quad \begin{aligned} (\xi^2 - n)^2 - 4n(\xi + a)^2 &= 0 \\ (\xi^2 + 3n)^2 + 4n(\xi - a)^2 &= 0. \end{aligned}$$

Wir wollen auch hier den in der Formel (34) ausgedrückten Satz hervorheben:

Jede Primzahl von der Form $4f + 1$ läßt sich in die Summe zweier Quadrate zerlegen.

§ 78.
Die komplexen Zahlen von Gauß.

Der zuletzt bewiesene Satz bildet die Grundlage für die Theorie des Zahlkörpers, der durch Adjunktion der imaginären Einheit $i = \sqrt{-1}$ aus dem Körper der rationalen Zahlen entsteht, den wir nach unserer Festsetzung mit $R(i)$ zu bezeichnen haben[1]).

Wir haben hier die Primzahlen von der Form $4f + 1$ zu unterscheiden von denen der Form $4f + 3$, und wir wollen der Kürze wegen die ersten mit der Primzahl 2 zusammenfassen und mit p, die zweiten mit q bezeichnen, also

$$p = 4f + 1,\ 2; \quad q = 4f + 3$$

setzen. Es gelten dann folgende Sätze:

1. Für jedes p gibt es zwei ganze Zahlen a und b, so daß
$$p = a^2 + b^2$$
ist.

Dies ist für die Primzahlen von der Form $4f + 1$ im Schlußsatze des letzten Paragraphen ausgesprochen und ist für $p = 2$ aus $2 = 1^2 + 1^2$ unmittelbar ersichtlich.

Dem steht ein zweiter Satz gegenüber:

2. Eine Primzahl q ist niemals in der Form $a^2 + b^2$ darstellbar,

oder der noch allgemeinere:

3. Die Summe zweier Quadrate ganzer Zahlen $a^2 + b^2$ ist nur dann durch eine Primzahl q teilbar, wenn a und b durch q teilbar sind.

[1]) Gauß, Theoria residuorum biquadraticorum, commentatio secunda. Werke, Bd. II.

Der Satz 2. ergibt sich einfach daraus, daß, wenn $a^2 + b^2$ ungerade ist, die eine der beiden Zahlen a, b gerade, also ihr Quadrat durch 4 teilbar, die andere ungerade, also ihr Quadrat $\equiv 1 \pmod 4$, also $a^2 + b^2 \equiv 1 \pmod 4$ sein muß. Der Satz 3, der übrigens den zweiten in sich schließt, wird so bewiesen:

Angenommen, es sei $a^2 + b^2$, aber nicht b, durch q teilbar, so bestimmen wir b' aus der Kongruenz $bb' \equiv 1 \pmod q$ und erhalten aus $a^2 + b^2 \equiv 0 \pmod q$:

$$(ab')^2 \equiv -1 \pmod q.$$

Dies ist aber unmöglich, da nach § 72, 4. für jede Primzahl q die Zahl -1 quadratischer Nichtrest ist.

Der Körper $R(i)$ ist der Inbegriff aller Zahlen von der Form $x + yi$, wenn x, y ganze oder gebrochene rationale Zahlen sind.

Die Zahlen $a + bi$, in denen a und b ganze Zahlen in R sind, heißen die ganzen Zahlen des Körpers $R(i)$.

Das Produkt zweier konjugierter Zahlen

$$\xi = x + yi, \qquad \xi' = x - yi,$$

also

$$\xi\xi' = x^2 + y^2,$$

heißt die Norm von ξ und wird mit $N(\xi)$ bezeichnet. Die Norm einer nicht verschwindenden Zahl ξ ist eine positive rationale Zahl und die Norm einer ganzen Zahl ist eine ganze Zahl. Die Norm eines Produktes oder eines Quotienten ist gleich dem Produkt oder dem Quotienten der Normen.

Eine ganze Zahl, deren Norm gleich 1 ist, heißt eine Einheit.

Da $a^2 + b^2$ (für ganzzahlige a, b) nur dann $= 1$ sein kann, wenn $a = \pm 1$, $b = 0$ oder $a = 0$, $b = \pm 1$ ist, so gibt es in $R(i)$ nur die vier Einheiten

$$+1, \quad -1, \quad +i, \quad -i.$$

Der reziproke Wert einer Einheit ist auch eine Einheit.

Zwei Zahlen, von denen die eine aus der anderen durch Multiplikation mit einer Einheit entsteht, heißen assoziierte Zahlen.

Jede komplexe Zahl gehört zu einem System von vier assoziierten Zahlen

$$a + bi, \quad -a - bi, \quad -b + ai, \quad b - ai.$$

Summe, Differenz und Produkt zweier ganzer Zahlen sind wieder ganze Zahlen.

> Eine ganze Zahl α heißt durch eine ganze Zahl β teilbar, wenn eine dritte ganze Zahl γ existiert, so daß $\alpha = \beta\gamma$ ist.

Ist α durch β teilbar, so ist $N(\alpha)$ durch $N(\beta)$ teilbar. Denn aus $\alpha = \beta\gamma$ folgt $N(\alpha) = N(\beta)\,N(\gamma)$.

Sind α, β, γ ganze Zahlen in $R(i)$, und ist α durch γ teilbar, so ist auch $\alpha\beta$ durch $\gamma\beta$ teilbar, und ist α und β durch γ teilbar, so ist auch $\alpha \pm \beta$ durch γ teilbar.

Jede ganze Zahl ist durch jede Einheit teilbar.

Die Einheiten sind aber nur durch Einheiten teilbar; denn ist $\alpha\beta$ eine Einheit, so ist $N(\alpha)N(\beta) = 1$, also $N(\alpha)$ und $N(\beta) = 1$, d. h. α und β sind Einheiten.

Assoziierte Zahlen sind gegenseitig durcheinander teilbar, und wenn zwei Zahlen gegenseitig teilbar sind, so sind sie assoziiert. Denn sind $\alpha:\beta$ und $\beta:\alpha$, deren Produkt $= 1$ ist, beides ganze Zahlen, so ist das Produkt der Normen und mithin jede der beiden Normen $= 1$; also sind beides Einheiten.

Für die Zahlen des Körpers $R(i)$ gelten dieselben Gesetze über die Zerlegung in Primfaktoren, wie bei den reellen ganzen Zahlen.

> Eine ganze Zahl α des Körpers $R(i)$, die keine Einheit ist, heißt zusammengesetzt, wenn sie sich in mehrere ganzzahlige Faktoren, deren keiner eine Einheit ist, zerlegen läßt.
>
> Läßt sie sich nicht so zerlegen, so soll sie eine Primzahl im Körper $R(i)$ heißen.

Ist $\xi = x + yi$ eine gebrochene Zahl des Körpers $R(i)$, so läßt sich eine ganze Zahl $\mu = m + ni$ so bestimmen, daß die Norm der Differenz $\xi - \mu$, also $(x - m)^2 + (y - n)^2$, kleiner oder wenigstens nicht größer als $^1/_2$ ist; denn man braucht die ganzen rationalen Zahlen m, n nur so zu wählen, daß $x - m$ und $y - n$ absolut genommen nicht größer als $^1/_2$, ihre Quadrate also nicht größer als $^1/_4$ sind.

Sind also α und α_1 zwei ganze Zahlen in $R(i)$, und α_1 von Null verschieden, so kann man hiernach die ganzen Zahlen μ und α_2 so bestimmen, daß

$$\frac{\alpha}{\alpha_1} - \mu = \frac{\alpha_2}{\alpha_1}, \quad \frac{N(\alpha_2)}{N(\alpha_1)} \leqq \frac{1}{2},$$

also

$$\alpha = \mu\alpha_1 + \alpha_2, \quad N(\alpha_2) < N(\alpha_1).$$

Ist α_2 nicht Null, so kann man ebenso mit den Zahlen α_1, α_2 verfahren und erhält

$$\alpha_1 = \mu_1\alpha_2 + \alpha_3, \quad N(\alpha_3) < N(\alpha_2),$$

und so bestimmt man eine Reihe ganzer Zahlen α, α_1, α_2, α_3 ... mit stets abnehmender Norm, und diese Reihe läßt sich so lange fortsetzen, als die Null darin nicht vorkommt. Da diese Normen ganze positive Zahlen sind, die immer abnehmen, so muß nach einer endlichen Zahl von Schritten die Null auftreten und man erhält also ein Gleichungssystem:

$$(1) \qquad \begin{aligned} \alpha &= \mu\alpha_1 + \alpha_2, \\ \alpha_1 &= \mu_1\alpha_2 + \alpha_3, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{h-2} &= \mu_{h-2}\alpha_{h-1} + \alpha_h, \\ \alpha_{h-1} &= \mu_{h-1}\alpha_h. \end{aligned}$$

Daraus schließt man, daß die ganze Reihe der Zahlen α, α_1, α_2 ..., also insbesondere auch α und α_1, durch α_h teilbar sind, und umgekehrt, daß jeder gemeinsame Teiler von α und α_1 Teiler von allen folgenden α, also auch von α_h ist. Jede andere Zahl, die diese beiden Eigenschaften hat, muß mit α_h assoziiert sein, und man nennt also α_h (und jede mit α_h assoziierte Zahl) den **größten gemeinschaftlichen Teiler** von α und α_1.

Wenn wir aus dem Algorithmus (1) die zwischenliegenden α eliminieren, indem wir α_2 aus der ersten in die zweite, dann α_3 aus der zweiten in die dritte Gleichung substituieren usf., so ergibt sich ein Resultat von der Form $\alpha\varkappa + \alpha_1\lambda = \alpha_h$, und wir können also, wenn wir statt α, α_1, α_h setzen α, β, δ, den Satz aussprechen:

4. Wenn α, β irgend zwei ganze Zahlen in $R(i)$ sind, und δ ihr größter gemeinschaftlicher Teiler, so kann man die ganzen Zahlen $\varkappa, \lambda$ so bestimmen, daß

$$(2) \qquad \alpha\varkappa + \beta\lambda = \delta$$

wird.

Ist der größte gemeinschaftliche Teiler δ zweier Zahlen α, β eine Einheit, so heißen α und β relativ prim, und wir können in diesem Falle der Gleichung

(3) $$\alpha\varkappa + \beta\lambda = 1$$

durch ganzzahlige $\varkappa$, λ genügen.

Ist β eine Primzahl, so ist entweder α durch β teilbar, oder α und β sind relativ prim. Im letzteren Falle besteht die Gleichung (3). Ist dann γ eine andere Zahl in $R(i)$, so folgt aus (3)

$$\alpha\gamma\varkappa + \beta\gamma\lambda = \gamma.$$

Es muß demnach, wenn $\alpha\gamma$ durch β teilbar ist, auch γ durch β teilbar sein. Also haben wir folgenden Fundamentalsatz:

5. Ein Produkt aus zwei oder mehr ganzen Zahlen in $R(i)$ ist nur dann durch eine Primzahl in $R(i)$ teilbar, wenn wenigstens einer seiner Faktoren durch diese Primzahl teilbar ist.

Macht man nun noch die Bemerkung, daß eine Primzahl nur dann durch eine andere teilbar sein kann, wenn der Quotient eine Einheit ist, wenn also beide assoziiert sind, und betrachtet assoziierte Primzahlen als nicht wesentlich verschieden, so folgt:

6. Eine ganze Zahl des Körpers $R(i)$ kann immer und wesentlich nur auf eine Art in ein Produkt von Primzahlen zerlegt werden.

Denn sei α irgend eine ganze Zahl in $R(i)$. Ist α nicht selbst eine Primzahl, so ist sie zerlegbar, etwa in $\gamma\beta$; die Normen von γ und von β sind aber kleiner als die Norm von α. Ist β noch keine Primzahl, so ist es wieder zerlegbar, etwa in $\delta\varepsilon$; die Normen dieser Faktoren sind positive ganze rationale Zahlen und nehmen immer ab, und man muß also notwendig einmal auf einen Faktor stoßen, der eine Primzahl ist. Also ist jede Zahl α gewiß durch eine Primzahl π teilbar. Ist demnach $\alpha = \pi\alpha'$, so gilt von α' dasselbe und es ist etwa $\alpha' = \pi'\alpha'' \ldots$ Da die Normen von α, α', $\alpha'' \ldots$ wieder alle abnehmen, so muß man bei der Fortsetzung dieser Reihe schließlich auf eine Einheit stoßen. Es ist daher α in eine endliche Anzahl von Primfaktoren zerlegbar, also

$$\alpha = \pi\pi'\pi'' \ldots,$$

wobei ein zuletzt übrig bleibender Einheitsfaktor mit einer der Primzahlen π vereinigt werden kann. Sind nun $\varkappa$, $\varkappa'$, $\varkappa'' \ldots$ gleichfalls Primzahlen und ist $\alpha = \varkappa\varkappa'\varkappa'' \ldots$, so folgt aus

$$\varkappa\varkappa'\varkappa'' \cdots = \pi\pi'\pi'' \ldots$$

nach 5., daß eine der Primzahlen π, z. B. die erste, durch $\varkappa$ teil-

bar, also von $\varkappa$ nicht wesentlich verschieden ist. Hebt man beiderseits mit $\varkappa = \pi$, so kann man denselben Schluß mit $\varkappa$ wiederholen usf., und findet also, daß jede der Primzahlen $\varkappa$ auch unter den π vorkommen muß, und wenn unter den $\varkappa$ eine Primzahl mehrmals vorkommt, so muß sie mindestens ebenso oft unter den π vorkommen. Derselbe Schluß läßt sich aber auch umgekehrt machen, und daraus folgt, daß die Gesamtheit der $\varkappa$, von Einheitsfaktoren abgesehen, mit der Gesamtheit der π übereinstimmen muß.

Es handelt sich nun noch darum, die Primzahlen des Körpers $R(i)$ wirklich zu ermitteln.

Jede Primzahl ist Teiler von unendlich vielen rationalen ganzen Zahlen, z. B. von ihrer Norm und allen ihren Vielfachen. Unter den rationalen positiven ganzen Zahlen, die durch eine Primzahl π teilbar sind, ist eine, die wir mit n bezeichnen wollen, die kleinste, und diese muß eine Primzahl im Körper R sein. Denn wäre sie in R zerlegbar, so müßte nach 5. einer ihrer Faktoren, also eine noch kleinere Zahl, durch π teilbar sein. Ebenso ist auch umgekehrt jede reelle Primzahl n wenigstens durch eine Primzahl π teilbar. Es ist also $n = \pi\alpha$, woraus durch Bildung der Norm $n^2 = N(\pi)N(\alpha)$ folgt. Da n eine Primzahl ist und $N(\pi)N(\alpha)$ ganze Zahlen, so sind zwei Fälle möglich:

$$1.\quad N(\pi) = n, \qquad N(\alpha) = n,$$
$$2.\quad N(\pi) = n^2, \qquad N(\alpha) = 1.$$

Im ersten Falle ist, wenn π' die zu π konjugierte Zahl bedeutet und $\pi = a + bi$ gesetzt wird, $n = \pi\pi' = a^2 + b^2$, $\alpha = \pi'$, und man sieht, daß dieser Fall nur dann eintritt, wenn n zu den Primzahlen p gehört, die in die Summe von zwei Quadraten zerlegbar sind. Umgekehrt kann jede solche Zahl p in zwei konjugierte Faktoren π, π' zerlegt werden, deren keine eine Einheit ist. Die Primzahlen p sind also im Körper $R(i)$ nicht Primzahlen.

Im zweiten Falle, der hiernach bei den Primzahlen q eintritt, ist α eine Einheit, also n mit π assoziiert, d. h. die reellen Primzahlen q sind auch im Körper $R(i)$ Primzahlen.

Die Primzahl 2 gehört, wie schon bemerkt, zu der ersten Art, und es ist $2 = (1 + i)(1 - i)$. Aber sie nimmt eine besondere Stellung ein, weil $1 - i = -i(1 + i)$, also $2 = -i(1 + i)^2$ ist. Die reelle Primzahl 2 ist also (von dem Faktor $-i$ ab-

gesehen) im Körper $R(i)$ das Quadrat einer Primzahl. Dieser Fall tritt bei keiner der übrigen Primzahlen p ein, weil sonst die beiden konjugierten Faktoren π, π' assoziiert sein müßten, was nicht möglich ist.

Die Gesamtheit der Primzahlen des Körpers $R(i)$ besteht also aus den reellen Primzahlen q und aus den Faktoren der Primzahlen p. Ist p in die Summe zweier Quadrate $a^2 + b^2$ zerlegt, so kennt man auch die konjugierten Faktoren $a \pm bi$ von p. Dazu kommen noch die assoziierten Zahlen. Aus 5. folgt dann noch, daß eine reelle Primzahl p nur auf eine Art in die Summe von zwei Quadraten zerlegt werden kann. Durch diese Zerlegung sind aber die Zahlen a, b noch nicht völlig bestimmt, sondern sie können noch miteinander vertauscht und mit zwei Vorzeichen versehen werden. Das kommt darauf hinaus, daß man π durch jede der vier assoziierten Zahlen $a + bi$, $-a - bi$, $-b + ai$, $b - ai$ ersetzen kann. Ist p ungerade, so muß von den beiden Zahlen a, b die eine gerade, die andere ungerade sein. Wählen wir etwa für b die gerade der beiden Zahlen, und bestimmen das Vorzeichen so, daß $a \equiv 1 \pmod 4$ wird, so ist unter den vier assoziierten Zahlen eine bestimmte ausgewählt, die man die primäre nennen kann[1]).

Diese Definition der primären Zahlen läßt sich auf alle ganzen Zahlen des Körpers $R(i)$ übertragen, deren Norm ungerade ist, und man hat dann das Gesetz, daß das Produkt zweier primärer Zahlen wieder eine primäre Zahl ergibt.

Das System von vier assoziierten Zahlen im Körper $R(i)$ ist analog dem Paar entgegengesetzter Zahlen im Körper R. In R betrachtet man die positiven Zahlen als die primären.

Durch die Formel (28) des vorigen Paragraphen ist für irgend eine gegebene Primzahl p einer der Faktoren $a + bi$ aus der Kreisteilung abgeleitet. Nach den Formeln § 77, (38), (39) ist diese Zahl primär, wenn $p \equiv 5 \pmod 8$, dagegen der primären entgegengesetzt, wenn $p \equiv 1 \pmod 8$ ist. Darüber aber, welche von den beiden konjugierten Zahlen $a \pm bi$ durch diese Formeln dargestellt ist, haben wir kein allgemeines Kennzeichen.

Beispielsweise sind die komplexen Primzahlen in $R(i)$, deren Normen kleiner als 200 sind:

---

[1]) Gauß gibt an der erwähnten Stelle zwei verschiedene Bestimmungen für die primären Zahlen zur Auswahl, von denen dies die erste ist. Er behält weiterhin die zweite bei.

$1+i$, $1+2i$, $3+2i$, $1+4i$, $5+2i$, $1+6i$, $5+4i$,
$7+2i$, $5+6i$, $3+8i$, $5+8i$, $9+4i$, $1+10i$, $3+10i$,
$7+8i$, $11+4i$, $7+10i$, $11+6i$, $13+2i$, $9+10i$, $7+12i$,
$1+14i$.

§ 79.

Biquadratische Abelsche Gleichungen und Kreisteilungskörper.

Die Zerlegung der rationalen ganzen Zahlen in ihre Primfaktoren in einem algebraischen Zahlkörper, von der die Gaußschen komplexen Zahlen ein erstes Beispiel geben, eröffnet den Zugang zu dem Satze von Kronecker, daß alle Abelschen Gleichungen im Körper R durch Kreisteilungszahlen d. h. durch Einheitswurzeln gelöst werden können. Der Beweis dieses Satzes für Gleichungen $n$ten Grades ist verhältnismäßig einfach in den Fällen, in denen die Zerlegung der ganzen Zahlen in Primfaktoren im Körper der $n$ten Einheitswurzeln eindeutig ist, wie z. B. im Fall der biquadratischen Gleichungen, den wir hier durchführen wollen; ähnlich auch für die kubischen Abelschen Gleichungen. Für andere Fälle macht aber der Beweis größere Schwierigkeiten und kann nur durch Heranziehen der sogenannten idealen Primfaktoren erbracht werden.

Der Satz, den wir also jetzt beweisen wollen, lautet so:

> Die Wurzeln einer biquadratischen Abelschen Gleichung im absoluten Rationalitätsbereich sind rationale Funktionen von Einheitswurzeln.

Die Galoissche Gruppe einer biquadratischen Normalgleichung muß vier Permutationen enthalten, und alle Funktionen der Wurzeln dieser Gleichung, die durch diese vier Permutationen ungeändert bleiben, sind rationalen Zahlen gleich. Solcher Gruppen gibt es aber nur die Vierergruppe und die zyklische Gruppe von vier Elementen. Diese beiden Fälle sind einzeln zu betrachten. Bezeichnen wir die vier Wurzeln unserer Gleichung mit x_0, x_1, x_2, x_3, so besteht die Vierergruppe aus den Permutationen:

(1) $\quad 1, \quad (0,1)(2,3), \quad (0,2)(1,3), \quad (0,3)(1,2),$

und wenn also dieses die Gruppe unserer Gleichung ist, so sind die drei Quadrate

$$(2)\qquad \begin{aligned} (x_0+x_1-x_2-x_3)^2 &= a\\ (x_0-x_1+x_2-x_3)^2 &= b\\ (x_0-x_1-x_2+x_3)^2 &= c \end{aligned}$$

und das Produkt

(3) $(x_0+x_1-x_2-x_3)(x_0-x_1+x_2-x_3)(x_0-x_1-x_2+x_3)=c,$

und ferner die Summe

(4) $$x_0+x_1+x_2+x_3=4A$$

rationale Zahlen.

Wenn wir aus (2) die Quadratwurzeln ausziehen und berücksichtigen, daß sich $\sqrt{a}$ nach (3) von $\sqrt{bc}$ nur durch einen rationalen Faktor unterscheidet, so folgt durch Addition:

(5) $$x_0=A+B\sqrt{b}+C\sqrt{c}+D\sqrt{bc},$$

worin A, B, C, D rationale Zahlen sind; und daraus erhält man x_1, x_2, x_3, wenn man die Vorzeichen von $\sqrt{b}$, $\sqrt{c}$ ändert.

Nach § 77 können aber alle Quadratwurzeln, nötigenfalls unter Zuziehung von i und $\sqrt{2}$, was ja selbst Kreisteilungszahlen sind, rational durch die Kreisteilungsperioden (die Gaußschen Summen) ausgedrückt werden, so daß also in (5) schon der Beweis unseres Satzes liegt

$$\left(e^{\frac{\pi i}{2}}=i,\quad e^{\frac{\pi i}{4}}=(1+i)\sqrt{2}\right).$$

Ist die Gruppe der biquadratischen Gleichung zyklisch, so können wir die Wurzeln so anordnen, daß die Gruppe aus den Permutationen

(6) $$1,\quad (0,\,1,\,2,\,3),\quad (0,\,2)\,(1,\,3),\quad (0,\,3,\,2,\,1)$$

besteht, und dann ist zu setzen:

(7)
$$\begin{aligned} 4A &= x_0+x_1+x_2+x_3\\ (i,x_0) &= x_0+ix_1-x_2-ix_3\\ (-1,x_0) &= x_0-x_1+x_2-x_3\\ (-i,x_0) &= x_0-ix_1-x_2+ix_3; \end{aligned}$$

darin ist A und $(-1,x_0)^2=m$ rational, und daher kann $(-1,x_0)$ durch Kreisteilungszahlen ausgedrückt werden. Die vierten Potenzen

(8) $$(i,x_0)^4=a+bi,\quad (-i,x_0)^4=a-bi$$

sind Zahlen des Körpers $R(i)$, und es kommt noch darauf an, $a+bi$ und $a-bi$ als vierte Potenzen von Kreisteilungszahlen darzustellen. Dies ergibt sich daraus, daß das Produkt

(9) $$(i,x_0)\,(-i,x_0)=c$$

durch die Permutationen (6) ungeändert bleibt und folglich eine rationale Zahl ist.

Im Körper $R(i)$ haben wir die Einheiten ± 1, $\pm i$, ferner als Primzahlen die assoziierten Faktoren $1 \pm i$ von 2, die reellen Primzahlen q der Form $4N + 3$ und die beiden komplexen Faktoren π, π' der reellen Primzahlen p von der Form $4N + 1$.

In § 77, (34) haben wir die Zerlegung

$$p = \psi_1(i)\,\psi_1(-i)$$

gefunden, die uns erlaubt,

$$\pi = \psi_1(i), \qquad \pi' = \psi_1(-i) \tag{10}$$

zu setzen, und, wenn η eine N-gliedrige Periode von $p$ten Einheitswurzeln ist [§ 77, (32)],

$$(i, \eta)^4 = p\,\psi_1(i)^2, \qquad (-i, \eta)^4 = p\,\psi_1(-i)^2,$$

oder

$$(i, \eta)^4 = \pi^3\pi', \qquad (-i, \eta)^4 = \pi\,\pi'^3, \qquad (i, \eta)\,(-i, \eta) = p. \tag{11}$$

Zerlegen wir nun die Zahlen $a + bi$, $a - bi$ in ihre Primfaktoren in $R(i)$, so ergibt sich, wenn λ, n, s, t positive Exponenten sind:

$$\begin{aligned} a + bi &= i^\lambda\ (1+i)^n\, q_1^{s_1} q_2^{s_2} \dots\ \pi_1^{t_1}\pi_1'^{t_1'}\pi_2^{t_2}\pi_2'^{t_2'}\dots \\ a - bi &= i^{-\lambda}\,(1-i)^n\, q_1^{s_1} q_2^{s_2} \dots\ \pi_1'^{t_1}\pi_1^{t_1'}\pi_2'^{t_2}\pi_2^{t_2'}\dots \end{aligned} \tag{12}$$

Das Produkt dieser beiden Zahlen muß aber die vierte Potenz einer rationalen Zahl sein [nach (8) und (9)], und daraus folgt:

$$\begin{aligned} n &\equiv 0, && && \pmod 4 \\ s_1 &\equiv 0, && s_2 \equiv 0 && \pmod 2 \\ t_1 + t_1' &\equiv 0, & t_2 + t_2' &\equiv 0 && \pmod 4 \\ t_1 - t_1' &\equiv 0, & t_2 - t_2' &\equiv 0 && \pmod 2. \end{aligned} \tag{13}$$

Es folgt aber jetzt aus (11):

$$\pi^t\pi'^{t'} = (\pi^3\pi')^{\frac{t-t'}{2}}\,(\pi\pi')^{\frac{-t+3t'}{2}} = (i, \eta)^{2(t-t')}\,p^{\frac{-t+3t'}{2}} \tag{14}$$

und ferner

$$(1+i)^n = (2i)^{\frac{n}{2}}, \tag{15}$$

und hiernach können wir also, mit Rücksicht auf (13), wenn wir mit H_1, H_2 zwei Kreisteilungszahlen, nämlich das Produkt aller in der Zerlegung von $a + bi$ vorkommenden Zahlen

$$(i, \eta)^{\frac{t-t'}{2}},$$

und die durch die Vertauschung von i mit $-i$ daraus hervorgehende Zahl, ferner mit c eine rationale Zahl bezeichnen, setzen:

$$a + bi = i^{\lambda}\ c^2 H_1^4$$
$$a - bi = i^{-\lambda} c^2 H_7^4.$$

Aus (8) folgt durch Ausziehen der vierten Wurzel, wodurch eine Potenz von i als Faktor hinzukommen kann,

$$(16) \qquad \begin{aligned} (i, x_0) &= i^h \sqrt[4]{i^{\lambda}}\sqrt{c}\, H_1 \\ (-i, x_0) &= i^{-h}\sqrt[4]{-i^{\lambda}}\sqrt{c}\, H_2 \\ (-1, x_0) &= \sqrt{m}, \end{aligned}$$

und daraus ergibt sich, da $\sqrt[4]{i}$, $\sqrt{c}$ und $\sqrt{m}$ Kreisteilungszahlen sind, die Richtigkeit unseres Satzes auch in diesem Falle.

Dreizehnter Abschnitt.

Algebraische Auflösung von Gleichungen.

§ 80.

Reduktion der Gruppe durch reine Gleichungen.

Eine der ältesten Fragen, an der sich vorzugsweise die neuere Algebra entwickelt hat, ist die nach der sogenannten algebraischen Auflösung der Gleichungen, worunter man eine Darstellung der Wurzeln einer Gleichung durch eine Reihe von Radikalen, oder die Berechnung durch eine endliche Kette von Wurzelziehungen versteht. Auf diese Frage fällt von der Gruppentheorie das hellste Licht.

Präzisieren wir zunächst die Frage, um die es sich handelt, so ist es offenbar die, ob und wie man den Körper Ω, der als ursprünglicher Rationalitätsbereich gilt, durch successive Adjunktion von Wurzelgrößen (Radikalen) so erweitern kann, daß entweder alle oder wenigstens ein Teil der Wurzeln im erweiterten Körper enthalten sind. Eine Wurzelgröße ist eine solche, die zwar nicht selbst, von der aber irgend eine ganze Potenz in Ω enthalten ist, also, wenn a eine Größe in Ω ist, die Wurzel einer Gleichung von der Form

$$y^m - a = 0,$$

d. h. einer reinen Gleichung.

Soll eine irreduzible Gleichung algebraisch auflösbar sein, oder wenigstens eine oder einige algebraisch darstellbare Wurzeln haben, so muß nach einer successiven Adjunktion von Wurzeln reiner Gleichungen in endlicher Anzahl die gegebene Gleichung reduzibel werden, da ein Teil ihrer Wurzeln in dem erweiterten Körper enthalten sein soll. Da die anfängliche Gruppe P transitiv ist, so muß diese Gruppe schließlich intransitiv werden, oder sich auf die Einheitsgruppe reduzieren. Es muß also jedenfalls

einmal der Fall eintreten, daß die Gruppe P durch Adjunktion einer Wurzel einer reinen Gleichung reduziert wird.

Die Untersuchung dieser Frage wird außerordentlich vereinfacht, wenn man sie noch etwas umformt[1]).

Wir haben im § 60 gesehen, daß die reinen Gleichungen zu den Abelschen gehören, daß alle Abelschen Gleichungen durch eine Kette von zyklischen Gleichungen von Primzahlgrad und diese letzteren durch Radikale lösbar sind.

Wir ersetzen also die Frage nach der Lösbarkeit durch Radikale durch die damit gleichbedeutende der Lösbarkeit durch eine Kette von Wurzeln zyklischer Gleichungen von Primzahlgrad. Nach einer endlichen Anzahl solcher Adjunktionen, wenn auch nicht gleich von Anfang, muß also eine weitere Reduktion der Gruppe eintreten, und da es sich hier nur um die notwendigen Bedingungen handelt, so fragen wir zunächst:

> Unter welchen Bedingungen wird die Gruppe P einer Gleichung $n$ten Grades $f(x) = 0$ durch Adjunktion einer Wurzel einer zyklischen Gleichung $\varphi(x) = 0$ von Primzahlgrad m auf einen Teiler Q reduziert?

Wir beschränken uns hierbei nicht auf irreduzible Gleichungen $f(x) = 0$, sondern erörtern die Frage allgemein, immer unter der selbstverständlichen Voraussetzung, daß $f(x)$ keine mehrfachen Wurzeln hat.

Bezeichnen wir die Wurzeln von $\varphi(x) = 0$ mit

$$\varepsilon,\ \varepsilon_1,\ \varepsilon_2 \ldots \varepsilon_{m-1},$$

so sind alle diese Größen rational (in Ω) durch eine beliebige unter ihnen ausdrückbar, und wenn P die Gruppe von $f(x) = 0$ in Ω ist, so ist Q die Gruppe derselben Gleichung in $\Omega(\varepsilon)$, oder was das elbe ist, in $\Omega(\varepsilon_1)$, $\Omega(\varepsilon_2) \ldots \Omega(\varepsilon_{m-1})$. Nun können wir aber den Satz 6., § 60 anwenden. Nach diesem Satze muß der Index j des Teilers Q von P ein Teiler von m sein, und da m als Primzahl vorausgesetzt ist, so ist $m = j$. Außerdem ist nach demselben Satze ε rational durch die Wurzeln der Gleichung $f(x) = 0$ darstellbar:

$$\varepsilon = \psi(x_0,\ x_1 \ldots x_{n-1}).$$

---

[1]) Auf diese Form der Fragestellung hat zuerst C. Jordan hingewiesen, (Traité des substitutions, p. 386).

Diese Funktion gehört zur Gruppe Q, und wenn wir darauf sämtliche Permutationen der Gruppe P anwenden, so erhalten wir die Funktionen ε, ε_1, $\varepsilon_2 \ldots \varepsilon_{m-1}$ und keine anderen. Diese Funktionen gehören zu den konjugierten Gruppen $\pi^{-1} Q \pi$. Da aber jede dieser Funktionen rational durch jede andere ausdrückbar ist, so müssen sie alle zu derselben Gruppe gehören, d. h. Q ist ein Normalteiler von P.

Wir haben also hiermit den ersten Satz bewiesen:

I. Wenn die Gruppe P einer Gleichung durch Adjunktion der Wurzeln einer zyklischen Gleichung reduziert wird, so hat P einen Normalteiler Q von Primzahlindex.

Dieser Satz läßt sich auch umkehren.

Wenn nämlich die Gruppe P einen Normalteiler Q vom Index m hat, so können wir eine zu Q gehörige Funktion ψ wählen, und die damit konjugierten Funktionen $\psi, \psi_1, \psi_2, \ldots \psi_{m-1}$ gehören alle zu derselben Gruppe. Der Körper $\Omega(\psi)$ ist ein Normalkörper, und ψ die Wurzel einer Normalgleichung. Im Körper $\Omega(\psi)$ ist Q die Gruppe von $f(x) = 0$ (§ 63). Wenn aber m eine Primzahl ist, so ist ψ nach § 64 die Wurzel einer zyklischen Gleichung, und damit ist also bewiesen:

II. Wenn die Gruppe P von $f(x) = 0$ einen Normalteiler Q von Primzahlindex m besitzt, so wird die Gruppe P durch Adjunktion der Wurzel einer zyklischen Gleichung $m$ten Grades auf Q reduziert.

§ 81.

Metazyklische Gleichungen.

Wir wollen eine Gleichung, deren vollständige Lösung sich auf eine Kette von zyklischen Gleichungen zurückführen läßt, eine metazyklische Gleichung nennen. Die zyklischen Gleichungen selbst sind als spezieller Fall darunter mit enthalten, und nach dem im § 80 Bemerkten sind die metazyklischen Gleichungen dieselben, wie die durch Radikale lösbaren Gleichungen. Ist P die Gruppe einer solchen Gleichung, so muß sie nach dem vorigen Paragraphen einen Normalteiler von Primzahlindex j_1, den wir jetzt mit P_1 bezeichnen wollen, besitzen. Besteht P_1 aus der einzigen identischen Permutation, so ist P selbst zyklisch und von Primzahlgrad. Ist P_1 nicht die Einheitsgruppe, so muß

P_1 wieder einen Normalteiler P_2 von Primzahlindex j_2 enthalten usf., bis wir endlich zur Einheitsgruppe gelangen.

Daß es auch die für eine metazyklische Gleichung ausreichende Bedingung ist, wenn ihre Gruppe P diese Zusammensetzung hat, ergibt sich aus dem vorigen Paragraphen. Wir sprechen also den Satz aus:

III. Die notwendige und hinreichende Bedingung für eine metazyklische Gleichung ist die, daß es eine Reihe von Gruppen

$$P, P_1, P_2, P_3 \ldots$$

gibt, deren erste die Galoissche Gruppe der Gleichung, deren letzte die Einheitsgruppe ist, von denen jede folgende ein normaler Teiler der nächst vorangehenden von Primzahlindex ist.

Hiernach nennen wir eine Permutationsgruppe P, die diese Eigenschaft hat, zu der sich also eine Kette von Gruppen

$$P, P_1, P_2 \ldots P_{\mu-1}, 1$$

so bestimmen läßt, daß jedes Glied Normalteiler des vorangehenden von Primzahlindex ist, eine metazyklische Gruppe[1]).

Wir haben hier die Bedingung für die vollständige Auflösbarkeit einer Gleichung durch eine Kette von zyklischen Gleichungen erhalten. Es handelt sich aber noch um die Frage, ob eine oder einige der Wurzeln auf diese Weise dargestellt werden können, während andere eine solche Darstellung nicht gestatten. Diese Frage ist nur berechtigt bei irreduziblen Gleichungen, da bei reduziblen Gleichungen alle denkbaren Kombinationen vorkommen können, und hier gilt nun der Satz:

IV. Wenn eine Wurzel einer irreduziblen Gleichung durch Lösung zyklischer Gleichungen bestimmbar ist, so ist die Gleichung metazyklisch.

Wenn eine irreduzible Gleichung $n$ten Grades $f(x) = 0$ auch nur eine Wurzel hat, die durch successive Adjunktion von Wurzeln zyklischer Gleichungen rational wird, so muß sie notwendig durch diese Adjunktion reduzibel werden, da sich ja schließlich ein linearer Faktor absondern muß. Es sei also P

[1]) Der Ausdruck „metazyklische Gruppen“ ist zuerst von Kronecker, wenn auch in beschränkterem Sinne, gebraucht. Sie wird sonst auch auflösbare Gruppe genannt, weil sie die Galoissche Gruppe der sogenannten „auflösbaren Gleichungen“ ist.

die Gruppe unserer Gleichung, nachdem alles Nötige so weit adjungiert ist, daß zwar $f(x)$ noch nicht reduzibel ist, aber durch die nächste Adjunktion der Wurzel ε einer zyklischen Gleichung von Primzahlgrad m in Faktoren zerfällt. Es muß dann P nach § 80 einen Normalteiler Q vom Index m haben, auf den sich die Gruppe der Gleichung nach Adjunktion von ε reduziert, und die Permutationsgruppe Q muß intransitiv sein. Wenn Q die Einheitsgruppe ist, so ist $f(x) = 0$ durch Adjunktion von ε vollständig gelöst. Ist aber Q noch von der Einheitsgruppe verschieden, so sondert sich von $f(x)$ nach Adjunktion von ε ein irreduzibler Faktor $\varphi(x, \varepsilon)$ ab, dessen Grad μ sei; sind $\varepsilon, \varepsilon_1, \varepsilon_2, \ldots \varepsilon_{m-1}$ die zu ε konjugierten Größen, die alle rational durch ε ausdrückbar sind, da sie die Wurzeln einer zyklischen Gleichung sind, so sind die Faktoren $\varphi(x, \varepsilon)$, $\varphi(x, \varepsilon_1)$, $\ldots$ $\varphi(x, \varepsilon_{m-1})$ alle in $f(x)$ enthalten. Wenn zwei dieser Funktionen einen gemeinsamen Teiler haben, so müssen sie, als irreduzibel, ganz identisch sein. Weil aber m eine Primzahl ist, so wären alle diese m Funktionen identisch, d. h. sie wären rational gegen die Voraussetzung.

Setzen wir, indem wir alles durch ε ausdrücken,

$$\varphi(x, \varepsilon_i) = \varphi_i(x, \varepsilon),$$

so folgt:

$$f(x) = \varphi(x, \varepsilon)\ \varphi_1(x, \varepsilon) \ldots \varphi_{m-1}(x, \varepsilon)$$

und es wäre also, durch Vergleichung der Grade in x,

$$n = \mu m.$$

Ist z. B. n eine Primzahl, so muß $\mu = 1$ sein; die Funktionen φ sind linear, und die Gleichung $f(x) = 0$ ist vollständig gelöst; also ist der Satz IV. für Gleichungen von Primzahlgrad richtig. Im allgemeinen Falle muß einer der Faktoren μten Grades $\varphi(x, \varepsilon)$, $\varphi_1(x, \varepsilon)$, $\ldots$ $\varphi_{m-1}(x, \varepsilon)$ etwa $\varphi(x, \varepsilon)$ eine Wurzel haben, die durch Adjunktion der Wurzeln zyklischer Gleichungen rational wird, und wenn wir also annehmen, daß unser Satz für Gleichungen μten Grades schon bewiesen sei, so folgt, daß $\varphi(x, \varepsilon) = 0$ selbst und also auch ihre Gruppe metazyklisch ist.

Da aber nach § 60 die verschiedenen Gleichungen $\varphi = 0$, $\varphi_1 = 0$, $\ldots$ $\varphi_{m-1} = 0$ dieselbe Gruppe haben, so sind sie alle und mithin auch $f(x) = 0$ metazyklisch.

Unter der Voraussetzung also, daß der Satz IV. für Gleichungen μten Grades richtig ist, folgt seine Richtigkeit für Gleichungen μmten Grades; und da er für Gleichungen von Primzahlgrad gilt, so ist er allgemein nachgewiesen.

§ 82.

Einfachheit der alternierenden Gruppe.

Wir haben früher gesehen (§ 58), daß, wenn wir die Koeffizienten einer Gleichung ηten Grades als unabhängige Variable und den Körper aller rationalen Funktionen dieser Koeffizienten als Rationalitätsbereich betrachten, die Galoissche Gruppe der Gleichung die symmetrische Gruppe ist. In der symmetrischen Gruppe ist immer ein Normalteiler vom Index 2 enthalten, die alternierende Gruppe, auf die sich die Gruppe der Gleichung reduziert, wenn die Quadratwurzel aus der Diskriminante adjungiert wird. Bei vier Ziffern hat die alternierende Gruppe, die aus der identischen Permutation, acht dreigliedrigen Zyklen und drei Paaren von Transpositionen besteht, den Normalteiler vom Index 3:

$$1,\quad (0,1)(2,3),\quad (0,2)(1,3),\quad (0,3)(1,2).$$

Diese Gruppe hat drei verschiedene Normalteiler vom Index 2, von denen wir einen 1, (0, 1) (2, 3) bevorzugen, von dem wieder die Einheitsgruppe ein Normalteiler vom Index 2 ist. Die Gruppe der 24 Permutationen von vier Ziffern ist also metazyklisch, und darauf beruht jede Auflösungsmethode der biquadratischen Gleichung.

Wir wollen nun nachweisen, daß, wenn n größer als 4 ist, die alternierende Gruppe außer der Einheitsgruppe überhaupt keine normalen Teiler hat, oder nach der früher eingeführten Bezeichnung einfach ist. Daraus folgt dann, daß die Bedingung, die wir für die algebraische Auflösbarkeit einer Gleichung als notwendig gefunden haben, für die Gleichungen von höherem als dem vierten Grade, deren Gruppe die symmetrische oder die alternierende ist, nicht erfüllt ist, und daß also Gleichungen von höherem als dem vierten Grade, solange die Koeffizienten unabhängige Variable sind, nicht mehr algebraisch lösbar sind.

Der Beweis, daß die alternierende Gruppe einfach ist, läßt sich so führen:

Sei A die alternierende Gruppe der Permutationen von n Ziffern 1, 2, 3, ... n und Q ein normaler Teiler von A. Wir haben im § 49, 7. und 4. gesehen, daß sich alle Permutationen von A aus dreigliedrigen Zyklen zusammensetzen lassen, und daß man, wenn $\varkappa$ und π irgend welche Permutationen sind, die

transformierte Permutation zu $\varkappa$, $\pi^{-1}\varkappa\pi$ erhält, wenn man in den Zyklen von $\varkappa$ die Vertauschungen π vornimmt. Ist nun $\varkappa$ ein dreigliedriger Zyklus, etwa (1, 2, 3), so kann man π aus A so wählen, daß $\pi^{-1}\varkappa\pi$ jeden beliebigen dreigliedrigen Zyklus der n Ziffern darstellt; denn man kann in

$$\pi = \begin{pmatrix} 1, & 2, & 3, & \dots n \\ a_1, & a_2, & a_3, & \dots a_n \end{pmatrix}$$

die drei ersten Ziffern a_1, a_2, a_3 beliebig wählen, und, wenn es nötig ist, damit π zu A gehöre, noch a_1 und a_2 vertauschen. Dadurch geht aus $\varkappa$ einer der beiden Zyklen (a_1, a_2, a_3), $(a_2, a_1\ a_3)$ hervor, von denen jeder die zweite Potenz des anderen ist. Wenn nun Q ein Normalteiler von A ist und $\varkappa$ eine Permutation aus Q, so ist $\pi^{-1}\varkappa\pi$ auch in Q enthalten, wenn π in A enthalten ist, und daraus folgt, daß, wenn in Q ein dreigliedriger Zyklus vorkommt, Q mit A identisch ist.

Unser Beweis beruht nun darauf, daß, wenn $\varkappa$ irgend eine Permutation in Q ist, auch $\pi^{-1}\varkappa\pi$ und folglich auch

$$\lambda = \varkappa^{-1}\pi^{-1}\varkappa\pi \tag{1}$$

in Q vorkommen muß, und es ist dann zu zeigen, daß man, wenn $\varkappa$ irgend eine nicht identische Permutation ist, π immer so aus A wählen kann, daß die Permutation λ ein dreigliedriger Zyklus, und folglich Q mit A identisch wird.

Wir nehmen zu diesem Zwecke sowohl $\varkappa$ als π in ihre Zyklen zerlegt an und bemerken, daß man bei der Bildung von λ solche Zyklen von $\varkappa$ gar nicht zu berücksichtigen braucht, deren Ziffern durch π ungeändert bleiben, weil sie sich in $\varkappa^{-1}$ und $\varkappa$ gegenseitig aufheben. Wir müssen nun die verschiedenen möglichen Formen von $\varkappa$ einzeln betrachten.

1. Es enthalte $\varkappa$ einen Zyklus von mehr als drei Ziffern, etwa $(1, 2, 3, \dots m)$, wir nehmen $\pi = (1, 2, 3)$ an und erhalten, indem wir $\varkappa^{-1}\pi^{-1}\varkappa$ zuerst (nach § 49, 4.) bilden,

$$\lambda = \varkappa^{-1}\pi^{-1}\varkappa\pi = (2, 4, 3)\ (1, 2, 3) = (1, 2, 4).$$

In Q kommt also ein dreigliedriger Zyklus vor.

2. Es enthalte $\varkappa$ zwei dreigliedrige Zyklen $(1, 2, 3)$ $(4, 5, 6)$. Wir nehmen $\pi = (1, 3, 4)$ an und erhalten

$$\lambda = (2, 5, 1)\ (1, 3, 4) = (1, 2, 5, 3, 4).$$

Diese Permutation λ, die in Q enthalten ist, fällt aber unter den Fall 1.

3. Es enthalte $\varkappa$ einen dreigliedrigen und einen zweigliedrigen Zyklus (1, 2, 3) (4, 5). (Daß in $\varkappa$, wenn es zu A gehört, noch ein zweiter Zyklus von gerader Gliederzahl vorkommen muß, ist hier gleichgültig.) Für $\pi = (1, 2, 4)$ ergibt sich
$$\lambda = (2, 5, 3)\,(1, 2, 4) = (1, 2, 5, 3, 4),$$
was wieder unter den Fall 1 fällt.
4. Es enthalte $\varkappa$ drei Transpositionen (1, 2) (3, 4) (5, 6). Für $\pi = (1, 3, 5)$ folgt
$$\lambda = (2, 6, 4)\,(1, 3, 5),$$
was unter den Fall 2 fällt.
5. Es enthalte $\varkappa$ zwei Transpositionen und ein unverändertes Element (1, 2) (3, 4) (5). Man setzt $\pi = (1, 2, 5)$ und erhält
$$\lambda = (1, 2, 5)\,(1, 2, 5) = (1, 5, 2).$$

Damit sind alle Fälle erschöpft, wenn $n > 4$ ist. Für $n = 4$ bleibt der eine Fall noch übrig, der eben das besondere Verhalten bei $n = 4$ herbeiführt, daß $\varkappa$ aus zwei Transpositionen besteht, wodurch die algebraische Auflösung der Gleichung vierten Grades ermöglicht wird[1]).

Es folgt aus diesem Satze weiter, daß die symmetrische Gruppe keine anderen normalen Teiler hat, als sich selbst, die alternierende Gruppe und die Einheitsgruppe. Denn ist S die symmetrische, A die alternierende Gruppe, und Q ein normaler Teiler von S, so ist der größte gemeinschaftliche Teiler Q' von A und Q ein normaler Teiler von A, ist also gleich A oder gleich 1. Denn ist $\varkappa'$ ein Element in Q', also auch in A, π ein beliebiges Element in S, so ist $\pi^{-1}\varkappa'\pi$ zunächst in Q enthalten, weil Q Normalteiler von S ist. Es ist aber zugleich in A, also auch in Q' enthalten.

Ist Q' gleich A, so ist Q entweder auch gleich A oder gleich S. Ist Q' aber $= 1$, so enthält Q außer der Einheit keine Permutation der ersten Art. Sind also $\varkappa$ und λ zwei ver-

[1]) Der erste vollständige Beweis, daß die allgemeine Gleichung von höherem als dem vierten Grade durch Radikale nicht lösbar ist, rührt von Abel her (Crelles Journal, Bd. I, 1826). Über die älteren Beweisversuche von Ruffini (1799 bis 1806) und ihr Verhältnis zum Abelschen Beweis vergleiche man die Abhandlung von Burkhardt, „Die Anfänge der Gruppentheorie und Paolo Ruffini“ (Abhandlungen zur Geschichte der Mathematik VI. Supplement zu Schlömilchs Zeitschrift. Leipzig 1892).

schiedene und von der Einheit verschiedene Permutationen von Q, so müssen $\varkappa^2$ und $\varkappa\lambda$ als von der ersten Art $= 1$ sein, d. h. λ muß $= \varkappa$ sein. Es kann also Q höchstens eine von der identischen verschiedene Permutation $\varkappa$ enthalten. Da aber Q ein Normalteiler von S sein soll, so muß für jede Permutation π aus der symmetrischen Gruppe $\pi^{-1}\varkappa\pi = \varkappa$ sein, d. h. $\varkappa$ darf sich nicht ändern, wenn in seinen Zyklen irgend eine Permutation ausgeführt wird. Dies ist aber nur dann möglich, wenn überhaupt nur zwei Ziffern 1, 2 vorhanden sind und $\varkappa = (1, 2)$ ist. Dann aber ist 1, (1, 2) die ganze Gruppe S.

§ 83.

Nichtmetazyklische Gleichungen im Körper der rationalen Zahlen.

Durch den Satz des vorigen Paragraphen ist der Nachweis geführt, daß eine Gleichung $n$ten Grades, wenn n größer als 4 ist, nicht mehr algebraisch gelöst werden kann, wenn die Koeffizienten als unabhängige Veränderliche betrachtet werden. Von größerem Interesse noch aber ist die Frage, ob es in dem Körper der rationalen Zahlen Gleichungen $n$ten Grades gibt, die nicht algebraisch lösbar sind. Die Frage läßt sich noch etwas allgemeiner stellen, nämlich so, ob es ganzzahlige Gleichungen gibt, deren Gruppe die symmetrische Gruppe ist, die also nach einer früher erklärten Ausdrucksweise keinen Affekt haben.

Bildet man die Galoissche Resolvente $G(t) = 0$ vom Grade $\Pi(n)$ einer allgemeinen Gleichung $n$ten Grades $f(x) = 0$ mit unbestimmten Koeffizienten a, so ist $G(t)$ eine ganze Funktion der Veränderlichen t und a, welche sich nicht in Faktoren zerlegen läßt, die wieder rationale Funktionen von t und a sind. Substituiert man für die Variablen a Größen irgend eines Körpers Ω, und wird dann $G(t)$ in diesem Körper reduzibel, so hat die Gleichung $f(x) = 0$ im Rationalitätsbereich Ω einen Affekt. Unsere Frage kommt also darauf hinaus, ob man in $G(t)$ für die Variablen a solche rationale Zahlen setzen kann, daß $G(t)$ im Körper der rationalen Zahlen irreduzibel bleibt.

Diese Frage hat eine ganz allgemeine Beantwortung gefunden in einer Abhandlung von Hilbert[1]), wo der allgemeine Satz

[1]) Über die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten. Crelle, Bd. 110.

bewiesen ist, daß man in einer irreduziblen Funktion beliebig vieler Variablen für einen beliebigen Teil der Variablen solche rationale Zahlen setzen kann, daß eine irreduzible Funktion der übrigen Variablen entsteht. Auf diesen allgemeinen Satz können wir hier nicht eingehen.

Wir werden aber die gestellte Frage viel einfacher, wenn auch bei weitem nicht so allgemein beantworten, indem wir zeigen, daß sich für jeden Primzahlgrad Gleichungen ohne Affekt finden lassen.

Wir haben in § 50, 4. bewiesen, daß eine transitive Permutationsgruppe von n Ziffern, die nicht die symmetrische Gruppe ist, wenn n eine Primzahl ist, keine einzelne Transposition enthalten kann.

Unter einer Gleichung mit einem Affekt haben wir eine solche verstanden, deren Gruppe nicht die symmetrische ist. Hat also die irreduzible Gleichung $f(x) = 0$ einen Affekt und ist ihr Grad eine Primzahl, so muß ihre Gruppe P transitiv sein, und sie kann keine Transposition zweier Wurzeln enthalten. Wenn wir also irgend $n - 2$ der Wurzeln dem Rationalitätsbereich adjungieren, so muß sie sich auf die Einheitsgruppe reduzieren, da ja außerdem nur noch die Vertauschung der beiden letzten Wurzeln übrig bleiben könnte, die in P nicht vorkommt. Daraus folgt also, daß die beiden letzten Wurzeln in dem erweiterten Körper Ω enthalten sind, oder der Satz:

1. Wenn eine irreduzible Gleichung, deren Grad n eine Primzahl ist, einen Affekt hat, so können zwei beliebige von ihren Wurzeln rational durch die übrigen ausgedrückt werden.

Daraus folgt als Korollar:

2. Wenn der Körper Ω nur reelle Zahlen enthält, so kann eine irreduzible Gleichung von Primzahlgrad n mit einem Affekt in Ω nicht zwei imaginäre und $n - 2$ reelle Wurzeln haben.

Nun gibt es aber unzählige Gleichungen von jedem beliebigen Grade n, mit reellen Koeffizienten, die zwei konjugiert imaginäre und $n - 2$ reelle Wurzeln haben; man kann ja in

$$\begin{aligned} f(x) &= (x - \alpha_1)\,(x - \alpha_2) \dots (x - \alpha_n) \\ &= x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n \end{aligned}$$

α_1, α_2 beliebig konjugiert imaginär und die übrigen α reell annehmen. Die Anzahl der reellen und imaginären Wurzeln von $f(x)$ ändert sich aber nicht, wenn die Koeffizienten innerhalb gewisser endlicher Grenzen stetig verändert werden (vgl. den fünften Abschnitt), und folglich gibt es auch solche Gleichungen, deren Koeffizienten rationale Zahlen sind, da man in beliebiger Nähe von irgend welchen gegebenen reellen Zahlen immer rationale Zahlen finden kann.

Es ist also nur noch nachzuweisen, daß man diese rationalen Koeffizienten so wählen kann, daß $f(x)$ irreduzibel wird. Dies ergibt sich aber aus dem Satz von Schönemann-Eisenstein (§ 21):

3. Ist p eine Primzahl, c_0, c_1, c_2, ... c_n eine Reihe ganzer Zahlen, von denen c_0 durch p nicht teilbar, $c_1, c_2, \ldots c_n$ durch p teilbar, aber c_n nicht durch p^2 teilbar ist, so ist die Funktion

$$\varphi(x) = c_0 x^n + c_1 x^{n-1} + \cdots + c_n$$

irreduzibel.

Denn wenn $\varphi(x)$ in zwei Faktoren mit rationalen Koeffizienten zerfällt, so können die Faktoren auch ganzzahlig angenommen werden. Sei also

$$\varphi(x) = (\alpha_1 x^h + \alpha_2 x^{h-1} + \cdots + \alpha_h)\ (\beta_0 x^k + \beta_1 x^{k-1} + \cdots + \beta_k),$$

h und k beide größer als Null und ihre Summe gleich n.

Da $\alpha_h \beta_k = c_n$ ist, so muß einer der beiden Faktoren α_h, β_k durch p teilbar sein, der andere nicht. Es möge β_k durch p teilbar, α_h nicht teilbar sein, und da nicht alle β durch p teilbar sein können, weil sonst auch c_0 durch p teilbar wäre, so sei β_ν nicht durch p teilbar, $\beta_{\nu-1}$, $\beta_{\nu+2} \ldots \beta_k$ durch p teilbar. Der Koeffizient von $x^{k-\nu}$ in dem Produkt der beiden Faktoren ist dann $\alpha_h \beta_\nu + \alpha_{h-1} \beta_{\nu+1} + \cdots$, also durch p nicht teilbar. Es müßte also $k - \nu = n$ sein, was nicht möglich ist, da schon $k < n$ sein muß.

Setzt man also für die Koeffizienten von $f(x)$:

$$a_1 = \frac{c_1}{c_0}, \quad a_2 = \frac{c_2}{c_0} \cdots a_n = \frac{c_n}{c_0},$$

so ist $f(x)$ irreduzibel, und man hat in der Wahl der ganzen Zahlen c noch Freiheit genug, um die rationalen Brüche $a_1, a_2, \ldots a_n$ einem beliebig gegebenen Wertsystem beliebig nahe zu bringen.

Damit ist aber der Satz bewiesen:

4. Es gibt von jedem beliebigen Primzahlgrade unendlich viele Gleichungen mit rationalen Koeffizienten ohne Affekt.

Der Beweis, der hier geführt ist, zeigt, daß solche affektfreie Gleichungen gefunden werden können, deren Koeffizienten ein endliches Gebiet überall dicht erfüllen, das Gebiet nämlich, in dem die reellen Koeffizienten der Gleichungen mit nur zwei konjugiert imaginären Wurzeln liegen. Der Beweis ist also, abgesehen davon, daß er sich nur auf Primzahlgrade beschränkt, auch insofern nicht erschöpfend, als er uns keinen Aufschluß gibt über die übrigen Gebiete, in denen aller Wahrscheinlichkeit nach die Sache sich ebenso verhält.

§ 84.
Auflösung durch reelle Radikale.

Bei der Auflösung der kubischen Gleichungen mit reellen Koeffizienten hat man von alters her die zwei Fälle unterschieden, in denen die Diskriminante negativ und positiv ist. Im zweiten Falle hat man, obwohl die Gleichung dann gerade drei reelle Wurzeln hat, bei der Anwendung der Cardanischen Formel die dritte Wurzel aus einem imaginären Ausdruck zu ziehen, und der Versuch, die Wurzeln in reeller Form darzustellen, führt immer wieder auf eine kubische Gleichung von derselben Beschaffenheit. Darum hat man diesen Fall den Casus irreducibilis genannt. Die kubischen Gleichungen des Casus irreducibilis sind ein spezieller Fall der zyklischen Gleichungen mit reellen Wurzeln, die wir im § 66 kennen gelernt haben, bei deren Lösung gleichfalls Wurzeln aus imaginären Größen oder, was auf dasselbe hinauskommt, Winkelteilungen vorkommen. Daß diese Wurzeln aus imaginären Größen oder Winkelteilungen auf keine Weise zu vermeiden sind, können wir jetzt beweisen[1]).

Wir setzen einen reellen Rationalitätsbereich Ω voraus, und nehmen in ihm eine Normalgleichung $g(t) = 0$ an, d. h. eine irreduzible Gleichung, deren Wurzeln alle rational durch eine beliebige von ihnen, ϱ, ausdrückbar sind. Wenn eine von den Wurzeln reell ist, so müssen auch alle anderen reell

[1]) Vgl. Hölder, Mathematische Annalen, Bd. 38; Kneser, ebendas., Bd. 41.

sein, und $g(t)$ hat also entweder lauter reelle oder lauter paarweise konjugierte imaginäre Wurzeln. $g(t) = 0$ kann die Galoissche Resolvente irgend einer gegebenen, sei es irreduziblen, sei es reduziblen Gleichung $F(x) = 0$ sein, und wenn also $g(t)$ nur reelle Wurzeln hat, so kann auch $F(x)$ nur reelle Wurzeln haben, weil die Wurzeln von $F(x)$ rational durch eine Wurzel von $g(t)$ darstellbar sind. Wenn unter den Wurzeln von $F(x)$ imaginäre sind, so hat $g(t)$ nur imaginäre Wurzeln. Hat aber $F(x)$ lauter reelle Wurzeln, so sind auch die Wurzeln von $g(t)$ reell, weil sie ja rational durch die Wurzeln von $F(x)$ ausgedrückt werden können.

Wenn nun $g(t)$ durch Adjunktion einer Wurzel ε einer irreduziblen Gleichung $\chi = 0$, deren Grad eine Primzahl ist, reduziert wird, so sind, wie wir im § 80 gesehen haben, die sämtlichen Wurzeln von χ in $\Omega(\varrho)$ enthalten, und wenn also ϱ reell ist, so sind alle Wurzeln von χ reell.

1. Eine Normalgleichung mit reellen Wurzeln kann also nur durch solche irreduzible Gleichungen von Primzahlgrad reduziert werden, die lauter reelle Wurzeln haben.

Wir fragen nun, ob eine Reduktion der Normalgleichung durch Adjunktion eines reellen Radikals $\sqrt[p]{a}$ bewirkt werden kann. Wir dürfen dabei annehmen, daß der Grad p des Radikals eine Primzahl sei, weil jedes Radikal sich auf eine Reihe von Wurzelziehungen von Primzahlgrad reduzieren läßt. Auch können wir voraussetzen, daß a nicht die $p$te Potenz einer rationalen Größe sei, weil sonst die reelle Wurzel $\sqrt[p]{a}$ rational wäre.

Unter dieser Voraussetzung ist, wie wir jetzt beweisen wollen, $x^p - a$ irreduzibel. Denn bezeichnen wir einen, z. B. den reellen Wert $\sqrt[p]{a}$ mit α und eine imaginäre $p$te Einheitswurzel mit ε so sind

$$\alpha, \quad \varepsilon\alpha, \quad \varepsilon^2\alpha \ldots \varepsilon^{p-1}\alpha \tag{1}$$

die Wurzeln der Gleichung

$$x^p - a = 0. \tag{2}$$

Ist diese Gleichung reduzibel, so ist

$$x^p - a = f_1(x) f_2(x), \tag{3}$$

und f_1, f_2 sind Funktionen in Ω von niedrigerem als dem $p$ten Grade. Ein Teil der Wurzeln (1) wird auf $f_1 = 0$, ein anderer

Teil auf $f_2 = 0$ fallen. Ist also μ der Grad von f_1, so muß für irgend einen Exponenten λ

$$\varepsilon^\lambda \alpha^\mu = b$$

eine Größe in Ω sein [das von x unabhängige Glied in $f_1(x)$], also, wenn man in die $p$te Potenz erhebt,

$$a^\mu = b^p. \tag{4}$$

Nun ist $0 < \mu < p$, und daher μ und p relativ prim, so daß sich zwei ganze Zahlen h und k aus der Gleichung

$$\mu h + pk = 1$$

bestimmen lassen. Es ist dann also nach (4)

$$a = a^{\mu h} a^{pk} = (b^h a^k)^p;$$

also wäre, gegen die Voraussetzung, a die $p$te Potenz einer Größe in Ω.

Dieser Satz ist, wie man sieht, unabhängig von der Voraussetzung, daß Ω ein reeller Körper sei. Nehmen wir ihn aber reell an, so hat die Gleichung (2), wenn p nicht gleich 2 ist, imaginäre Wurzeln, und kann nach 1. eine Normalgleichung $g(t) = 0$ nicht reduzieren. Wenn aber $p = 2$ ist, so kann $g(t)$ nur dann durch die Gleichung (2) reduziert werden, wenn a positiv und der Grad von $g(t)$ eine gerade Zahl ist (§ 60); also:

2. Eine Normalgleichung ungeraden Grades mit reellen Koeffizienten kann nicht durch Adjunktion eines reellen Radikals reduziert werden.

Adjungiert man im Falle einer reellen kubischen Gleichung mit positiver Diskriminante die Quadratwurzel aus dieser Diskriminante, so geht die Gleichung in eine Normalgleichung über, und der Rationalitätsbereich bleibt reell; er bleibt auch reell, wenn man noch so viele reelle Radikale adjungiert. Soll die Gleichung also durch reelle Radikale lösbar sein, so muß sie bei solchen Adjunktionen endlich zerfallen, was nach 2. unmöglich ist. Dies ist der Casus irreducibilis der kubischen Gleichungen.

In demselben Falle finden sich die zyklischen und überhaupt alle irreduziblen Abelschen Gleichungen von ungeradem Grade, die also niemals durch reelle Radikale lösbar sind.

Hiernach können wir z. B. die kubischen Gleichungen im Körper der rationalen Zahlen in drei oder vier Arten unterscheiden, die alle in gewissen, von alters her berühmten geometrischen Problemen auftreten. Wir sehen dabei von den

reduziblen Gleichungen ab. Der Grad der Galoisschen Gruppe muß dann immer durch 3 teilbar sein (§ 63) und ist also entweder gleich 3 oder gleich 6. Die Gruppe kann also nach § 60 niemals durch Adjunktion von bloßen Quadratwurzeln auf einen niedrigeren als den dritten Grad reduziert werden; also kann auch die Gleichung nicht durch Quadratwurzeln gelöst werden. Die entsprechenden geometrischen Probleme sind nicht mit Zirkel und Lineal zu lösen.

Ist die Gruppe vom Grade 3, so haben wir eine zyklische Gleichung vom Grade 3. Hierher gehören die aus der Kreisteilung stammenden Gleichungen, z. B. die, von der die Konstruktion des regelmäßigen Siebenecks abhängt. Die drei Wurzeln einer solchen Gleichung sind reell und können nicht durch reelle Radikale ausgedrückt werden.

Unter den kubischen Gleichungen mit einer Gruppe sechsten Grades sind zu unterscheiden die mit positiver und mit negativer Diskriminante. Die ersten gehören zum Casus irreducibilis und lassen sich zurückführen auf die Gleichung, von der die Dreiteilung eines beliebigen Winkels abhängt. Zu den Gleichungen mit negativer Diskriminante gehören als spezielle Fälle auch die reinen kubischen Gleichungen $x^3 = a$, wenn a keine Kubikzahl ist. Für $a = 2$ ergibt sich die Gleichung, von der das Delische Problem der Würfelverdoppelung abhängt.

§ 85.

Metazyklische Gleichungen von Primzahlgrad.

Die allgemeinen Bedingungen, die wir im § 81 gefunden haben, sind noch nicht einfach genug, um eine unmittelbare Anwendung auf die Ermittelung von metazyklischen Gleichungen oder auf die Entscheidung über die Löslichkeit einer vorgelegten Gleichung durch Radikale zu gestatten. Wir leiten also, zunächst unter der Voraussetzung, daß der Grad n der Gleichung eine Primzahl sei, ein anderes von Galois aufgestelltes Kriterium her.

Es sei jetzt $f(x) = 0$ eine irreduzible Gleichung vom Grade n, und n eine Primzahl und größer als 2. Soll $f(x) = 0$ algebraisch lösbar sein, so muß ihre Gruppe P nach § 81 metazyklisch sein, d. h. es muß eine Kette von Gruppen

$$(1) \qquad P,\ P_1,\ P_2 \ldots P_{\mu-1},\ 1$$

geben, deren jede ein Normalteiler der nächst vorangehenden mit Primzahlindex ist. Durch successive Adjunktion von Wurzeln zyklischer Gleichungen von Primzahlgrad wird die Gruppe der Gleichung von P auf P_1, $P_2 \ldots P_{\mu-1}$, 1 reduziert.

Die Funktion $f(x)$ selbst kann, wie schon im § 81 gezeigt ist, nicht reduziert werden, ehe die letzte Adjunktion gemacht ist, worauf sie in n lineare Faktoren zerfällt. Der Grad der letzten zyklischen Gleichung und mithin der Grad der vorletzten Gruppe $P_{\mu-1}$ muß also gleich n sein.

Der Grad eines Elementes einer Gruppe ist immer ein Teiler vom Grade der Gruppe und daher kann $P_{\mu-1}$ außer der identischen nur Permutationen von der Ordnung n enthalten. Sie ist also mit der Periode 1, π, $\pi^2 \ldots \pi^{n-1}$ eines ihrer Elemente identisch. Daraus folgt nach § 49, daß π eine zyklische Permutation von sämtlichen n Ziffern sein muß, und wir können die Bezeichnung der Wurzeln von $f(x)$ so wählen, daß

$$\pi = (0, 1, 2 \ldots n - 1)$$

wird. Bezeichnen wir die Wurzeln mit x_z, und setzen fest, daß $x_z = x_{z'}$ sein soll, wenn $z \equiv z' \pmod{n}$ ist, so geht durch π jedes z in $z + 1$ über, durch π^2 in $z + 2$ usf., und wir können also sagen, daß die Gruppe $P_{\mu-1}$ aus den Substitutionen für z

$$(z, z + b), \quad b = 0, 1, 2, \ldots n - 1 \tag{2}$$

besteht. Jede irreduzible metazyklische Gleichung von Primzahlgrad kann also durch Adjunktion von Radikalen in eine zyklische Gleichung verwandelt werden.

Die Substitutionen $(z, z + b)$ bilden einen speziellen Fall der allgemeineren linearen Substitution

$$(z, az + b), \tag{3}$$

worin a und b feste Zahlen sind, die auch nach dem Modul n reduziert werden können, wo aber für a der Wert 0 natürlich auszuschließen ist, weil sonst ja durch (3) alle verschiedenen z in dasselbe b übergehen, also gar keine Permutation der Größen x_z ausgedrückt wäre. Ist aber a von 0 verschieden (mod n), so wird durch (3) immer eine Permutation dargestellt sein, weil dann nur, wenn $z \equiv z'$ ist, $az + b \equiv az' + b$ sein kann. Die Anzahl der verschiedenen linearen Substitutionen von der Form (3) ist $n(n - 1)$. Ihnen entsprechen ebenso viele verschiedene Permutationen unter den x; denn es kann nur dann für jedes z

$$az + b \equiv a'z + b' \pmod{n}$$

sein, wenn $b \equiv b'$ (aus $z = 0$ zu schließen) und $a \equiv a'$ (aus $z = 1$ zu schließen).

Die Gesamtheit der Permutationen, die durch (3) dargestellt sind, bildet eine Gruppe; denn es seien

$$\lambda = (z,\ az + b), \quad \lambda' = (z,\ a'z + b')$$

zwei von diesen Permutationen, so ist

$$\lambda\lambda' = \begin{pmatrix} x_z \\ x_{az+b} \end{pmatrix} \begin{pmatrix} x_z \\ x_{a'z+b'} \end{pmatrix} = \begin{pmatrix} x_z \\ x_{a'(az+b)+b'} \end{pmatrix},$$

also

$$\lambda\lambda' = \lambda'' = (z,\ aa'z + a'b + b'), \tag{4}$$

was wieder von der Form $(z,\ a''z + b'')$ ist.

Die durch alle Substitutionen von der Form (3) gebildete Gruppe, die eine Verallgemeinerung der zyklischen Gruppe (2) ist, wollen wir eine lineare Gruppe nennen[1]); die in ihr enthaltenen Permutationen sollen lineare Permutationen heißen.

In der linearen Gruppe sind verschiedene Divisoren enthalten. Wir finden darunter eine Gruppe vom Grade $n - 1$, nämlich $(z,\ az)$, die selbst wieder Divisoren haben kann.

Die Kompositionsregel (4) zeigt, wie wir alle Divisoren der linearen Gruppe bilden können. Ist $\lambda = (z,\ az + b)$ eine lineare Substitution, so ergibt sich nach (4) durch Wiederholung für jeden Exponenten h

$$\lambda^h = [z,\ a^h z + (1 + a + \cdots + a^{h-1})b], \tag{5}$$

und für $a = 1$

$$\lambda^h = (z,\ z + hb).$$

Wenn also in einer Gruppe L eine Substitution λ vorkommt, in der $a = 1$ und b von Null verschieden ist, so enthält hiernach L die ganze zyklische Gruppe $(z,\ z + b)$ $(b = 0,\ 1,\ \ldots\ n - 1)$.

Ist a nicht gleich 1, so setzen wir

$$1 + a + \cdots + a^{h-1} = \frac{a^h - 1}{a - 1},$$

und schließen daraus, daß, wenn e der kleinste positive Exponent ist, für den $a^e \equiv 1 \pmod{n}$ ist, λ vom Grade e ist.

Die Periode der Substitution λ gibt, wenn nicht $a = 1$ ist, eine intransitive Permutationsgruppe. Es gibt eine Ziffer z, die durch λ und seine Wiederholungen nicht geändert wird, die aus der Kongruenz $z \equiv az + b \pmod{n}$ bestimmt wird, und mit

[1]) Kronecker nennt nur diese Gruppe metazyklisch.

$$z_0 \equiv -\frac{b}{a-1} \pmod{n}$$

bezeichnet werden kann, wenn unter $\frac{\mu}{\nu}$ oder $\mu : \nu \pmod{n}$ eine ganze Zahl verstanden wird, die durch Multiplikation mit dem Faktor ν nach dem Modul n mit μ kongruent wird.

Die zyklische Gruppe $(z, z+b)$ ist transitiv und vom Grade n. Wir beweisen zunächst den Satz:

I. Jede transitive lineare Gruppe (mod n) enthält die zyklische Gruppe.

Es sei g eine primitive Wurzel von n, und α der Index von a, d. h. $g^\alpha \equiv a \pmod{n}$ (§ 70). Der Kürze wegen nennen wir für den Augenblick α zugleich den Index der Substitution $\lambda = (z, az+b)$. Der Index kann kleiner als $n-1$ und gleich oder größer als Null angenommen werden. Ist er gleich Null, so ist λ entweder die identische Substitution, oder sie gehört der zyklischen Gruppe an. Nach (4) gilt die Regel, daß der Index einer zusammengesetzten Substitution gleich der Summe der Indices der Komponenten ist.

Daraus folgt, daß alle Indices der Substitutionen einer linearen Gruppe L Vielfache des kleinsten positiven unter ihnen sind, und daß folglich, wenn α_0 dieser kleinste positive Index ist, und $g^{\alpha_0} = a_0$ gesetzt wird, alle Substitutionen von L in der Form

$$\lambda = (z, a_0^h z + b)$$

dargestellt werden können, worin a_0 festgehalten wird, und nur h und b gewisse Zahlenreihen durchlaufen. Eine Substitution von der Form

$$\lambda_0 = (z, a_0 z + b_0)$$

muß in der Gruppe L vorkommen.

Wenn nun in der Gruppe L eine Substitution λ vorkommt, in der mit $h = 0$ ein von Null verschiedenes b verbunden ist, so gehört dieses λ der zyklischen Gruppe an, und L enthält die ganze zyklische Gruppe.

Wenn aber h von Null verschieden ist, so bilden wir nach (4) und (5) die Zusammensetzung

$$\lambda \lambda_0^{-h} = \left(z, z + a_0^{-h} b + \frac{a_0^{-h} - 1}{a_0 - 1} b_0\right).$$

Diese gibt dann und nur dann die identische Substitution, wenn

$$\frac{b_0}{a_0 - 1} \equiv \frac{b}{a_0^h - 1} \pmod{n}.$$

Wenn diese Kongruenz für alle von Null verschiedenen h befriedigt ist, und wenn dem Index $h = 0$ nur die identische Substitution entspricht, so bleibt das Element $-b_0 : (a_0 - 1)$ durch die ganze Gruppe ungeändert, und L ist intransitiv. Folglich gibt es in einer transitiven Gruppe L immer ein Element λ, für das $\lambda \lambda_0^{-h}$ nicht die identische Substitution ist, obwohl der Index Null ist, und L enthält also die ganze zyklische Gruppe.

Bildet man die verschiedenen Potenzen von λ_0, so ergibt sich, daß der Exponent h in den Substitutionen λ der Gruppe L jeden Wert annehmen kann. Ist e der kleinste positive Exponent, der der Bedingung $a_0^e \equiv 1 \pmod{n}$ genügt, so kommen also die Werte $h = 0, 1, 2, \ldots e - 1$ vor, und darin ist e ein Teiler von $n - 1$. Aus der zusammengesetzten Substitution

$$(z, a_0^h z + b)\,(z, z + 1) = (z, a_0^h z + b + 1),$$

die nach I. in einer transitiven Gruppe L ja auch vorkommen muß, sieht man weiter, daß mit jedem Werte von h jeder der Werte $b = 0, 1, 2 \ldots n - 1$ verbunden vorkommt, und die ganze Gruppe L besteht daher aus den Substitutionen

$$\lambda = (z, a_0^h z + b), \quad \begin{matrix} h = 0, 1, \ldots e - 1, \\ b = 0, 1, \ldots n - 1, \end{matrix} \tag{6}$$

und ist vom Grade en. Umgekehrt bildet jedes System von Substitutionen dieser Form eine Gruppe.

Von den linearen Gruppen gilt nun der folgende Satz:

II. Ist eine transitive lineare Gruppe L Normalteiler einer anderen Permutationsgruppe P derselben n Ziffern, so ist auch P eine lineare Gruppe.

Der Satz läßt sich so beweisen:

Wenn π eine beliebige Permutation ist:

$$\pi = \begin{pmatrix} 0, & 1, & 2, & \ldots n - 1 \\ a_0, & a_1, & a_2, & \ldots a_{n-1} \end{pmatrix},$$

die auch durch (z, a_z) dargestellt werden kann, so kann man eine ganze Funktion $\varphi(z)$ von z mit rationalen Koeffizienten, deren Grad nicht höher als $n - 1$ ist, so bestimmen, daß $a_z = \varphi(z)$ gesetzt werden kann. Man braucht nur die n Koeffizienten in $\varphi(z)$ aus den n linearen Gleichungen

$$a_0 = \varphi(0), \quad a_1 = \varphi(1), \ldots a_{n-1} = \varphi(n-1)$$

zu bestimmen. Man kann dazu die Lagrangesche Interpolationsformel verwenden, der man aber noch eine einfachere Gestalt geben kann, da es hier nur auf Kongruenzen nach dem Modul n ankommt. Setzt man nämlich

$$\psi(z) = z(z-1)(z-2)\ldots(z-n+1),$$

so gibt die erwähnte Interpolationsformel

$$\varphi(z) = \psi(z)\left(\frac{a_0}{\psi'(0)z} + \frac{a_1}{\psi'(1)(z-1)} + \cdots + \frac{a_{n-1}}{\psi'(n-1)(z-n+1)}\right).$$

Nun ist, wie wir im § 70 gesehen haben, die Kongruenz

$$\psi(z) \equiv z^n - z \pmod{n}$$

identisch; also ist auch

$$\psi'(z) \equiv n z^{n-1} - 1 \equiv -1 \pmod{n},$$

und daher können wir $\varphi(z)$ mit ganzzahligen Koeffizienten so darstellen:

$$\varphi(z) = -a_0 \frac{\psi(z)}{z} - a_1 \frac{\psi(z)}{z-1} - \cdots - a_{n-1} \frac{\psi(z)}{z-n+1}.$$

Wenn, wie in unserer Aufgabe, die Zahlen $a_0, a_1, \ldots a_{n-1}$ von der Ordnung abgesehen, mit den Zahlen $0, 1, 2, \ldots n-1$ übereinstimmen, so ist, wenn $n > 2$ ist, $a_0 + a_1 + \cdots + a_{n-1} \equiv 0 \pmod{n}$ und also $\varphi(z)$ höchstens vom Grade $n-2$, was aber für unseren Beweis nicht von Bedeutung ist.

Hiernach läßt sich also jede beliebige Permutation π durch eine Substitution $[z, \varphi(z)]$ darstellen.

Ist nun L eine transitive lineare Gruppe und zugleich Normalteiler von einer anderen Gruppe P, ist λ eine beliebige Permutation von L, π eine gleichfalls beliebige Permutation von P, so ist $\pi^{-1}\lambda\pi = \lambda'$ in L enthalten, und $\lambda\pi = \pi\lambda'$. Setzen wir nach dem, was eben bewiesen ist, $\pi = [z, \varphi(z)]$ und wählen für λ die zyklische Substitution $(z, z+1)$, die nach der Voraussetzung der Transitivität in L vorkommt, so muß sich a' und a so bestimmen lassen, daß

$$(z, z+1)\,[z, \varphi(z)] = [z, \varphi(z)]\,(z, a'z + a),$$

oder, wenn man die Zusammensetzung ausführt,

$$\varphi(z+1) \equiv a'\varphi(z) + a \pmod{n} \tag{7}$$

für jedes ganzzahlige z. Da $\varphi(z)$ den Grad n nicht erreicht, so müssen (§ 70) in (7) die Koeffizienten der einzelnen Potenzen von z auf beiden Seiten kongruent sein, und aus der Vergleichung

der Koeffizienten der höchsten Potenzen von z ergibt sich $a' \equiv 1 \pmod{n}$, also

$$\varphi(z+1) \equiv \varphi(z) + a \pmod{n}.$$

Setzt man hier $z+1, z+2 \ldots z+h-1$ für z, so folgt

$$\varphi(z+h) \equiv \varphi(z) + ah \pmod{n},$$

wo h mit jeder beliebigen ganzen Zahl nach dem Modul n kongruent sein kann. Darin kann man nun $z = 0$ setzen und erhält, wenn man wieder z für h schreibt und $\varphi(0) = b$ setzt,

$$\varphi(z) = az + b. \tag{8}$$

Es besteht also die Gruppe P aus lauter linearen Permutationen und unser Satz ist bewiesen.

Wir fügen noch die Bemerkung hinzu, die sich unmittelbar aus dem Anblick der Formeln ergibt, daß die zyklische Gruppe $(z, z+b)$ ein normaler Teiler einer jeden linearen Gruppe ist, in der sie überhaupt enthalten ist.

Wenn wir von dem bewiesenen Satze die Anwendung auf die Gruppe der metazyklischen Gleichung machen, so erhalten wir den Satz von Galois:

III. Die Gruppe einer irreduziblen metazyklischen Gleichung von Primzahlgrad ist linear.

Denn kehren wir zu der Kette der Gruppen (1) zurück, so haben wir gesehen, daß $P_{\mu-1}$ die zyklische Gruppe ist. Ist $P_{\mu-1}$ nicht mit P identisch, so ist $P_{\mu-1}$ ein Normalteiler von $P_{\mu-2}$, und also $P_{\mu-2}$ linear. $P_{\mu-2}$ ist wieder Normalteiler von $P_{\mu-3}$, also ist auch $P_{\mu-3}$ linear usf., bis wir zu dem Schluß gelangen, daß auch P selbst linear sein muß. Zugleich ergibt sich, daß $P_{\mu-1}$ Normalteiler aller vorangehenden Gruppen, also auch von P selbst ist.

Es gilt auch der umgekehrte Satz:

IV. Jede irreduzible Gleichung von Primzahlgrad, deren Gruppe linear ist, ist metazyklisch.

Um ihn zu beweisen, genügt es, zu zeigen, daß jede transitive lineare Gruppe L einen Normalteiler L' hat, dessen Index eine Primzahl, und der selbst, wenn er nicht die Einheitsgruppe ist, eine transitive lineare Gruppe ist.

Dies zeigt aber unmittelbar die Darstellung der Substitutionen der Gruppe L durch die Formel (6).

Wenn darin $e = 1$ ist, so ist L die zyklische Gruppe vom Grade n mit dem Normalteiler 1. Ist aber $e > 1$, so sei p eine

in e aufgehende Primzahl, und $e = pe'$. Die lineare Gruppe L' vom Grade ne', die aus den Substitutionen besteht:

$$\lambda' = (z, a_0^h z + b), \quad \begin{matrix} h = 0, 1, 2, \ldots e' - 1, \\ b = 0, 1, 2, \ldots n - 1, \end{matrix}$$

ist gewiß ein Teiler von L vom Index p. Dieser Teiler ist aber auch normal, wie man aus der Kompositionsregel (4) ohne Mühe erkennt.

Die Begriffe der transitiven linearen und der metazyklischen Gruppen decken sich also bei den Gleichungen von Primzahlgrad vollständig, und wir können daher auch in der Folge beide Ausdrücke synonym gebrauchen.

Wenn wir von einer Gruppe P zu einer konjugierten Gruppe $P' = \pi^{-1} P \pi$ übergehen wollen, so geschieht dieser Übergang dadurch, daß bei allen Permutationen von P in den Zyklen die Vertauschung π vorgenommen wird. P' wird also mit P bis auf die Bezeichnung der Wurzeln übereinstimmen, und ist daher, wenn P linear ist, auch als linear zu bezeichnen, wenn auch eine Änderung in der Numerierung der Wurzeln nötig ist, um sie durch lineare Substitutionen darzustellen.

Zur Vereinfachung der Anwendung bemerke man noch, daß man die volle lineare Gruppe durch Wiederholung und Zusammensetzung der beiden Substitutionen

$$(9) \qquad s = (z, z + 1), \quad t = (z, gz),$$

und jeden transitiven Teiler L der vollen linearen Gruppe ebenso aus

$$(10) \qquad s = (z, z + 1), \quad t^{a_0} = (z, a_0 z)$$

ableiten kann. Man nennt daher die beiden Substitutionen (9) oder (10) die erzeugenden Substitutionen dieser Gruppen.

Setzen wir $a_0 = g^2$, so erhalten wir eine häufig vorkommende lineare Gruppe $(z, az + b)$, in der a nur die quadratischen Reste von n durchläuft, die von Kronecker die halbmetazyklische Gruppe genannt worden ist.

Diese Darstellung durch die erzeugenden Substitutionen gestattet einen einfachen Schluß auf die Beziehung der metazyklischen Gruppen zu der symmetrischen und der alternierenden Gruppe.

Die aus s hervorgehende Permutation besteht aus einem einzigen Zyklus mit einer ungeraden Gliederzahl $(0, 1, 2 \ldots n - 1)$, und gehört daher zu der alternierenden Gruppe. Die Substitu-

tion t läßt den Index 0 ungeändert und liefert für die übrigen Ziffern wieder einen einzigen Zyklus. Denn durch t geht 1 in g, g in g^2, g^2 in g^3 usw. über, und da die Potenzen $1, g, g^2 \ldots g^{n-2}$, von der Reihenfolge abgesehen, mit den Ziffern $1, 2, \ldots n-1$ übereinstimmen, so entspricht t dem Zyklus $(1, g, g^2, \ldots g^{n-2})$, der aus einer geraden Gliederzahl besteht und folglich zu den Permutationen zweiter Art gehört, also nicht in der alternierenden Gruppe enthalten ist. Dagegen ist t^2 wieder darin enthalten. Daraus ergibt sich das Resultat:

V. Die volle lineare Gruppe ist kein Teiler der alternierenden Gruppe. Der größte gemeinschaftliche Teiler beider Gruppen ist die halbmetazyklische Gruppe.

Man kann der Bedingung für die metazyklischen Gleichungen verschiedene Formen geben, die sich aus dem Bisherigen ableiten lassen.

Die volle lineare Gruppe ist als Teiler vom Index $\nu = 1.2.3 \ldots (n-2)$ in der symmetrischen Gruppe enthalten. Eine zu der vollen linearen Gruppe gehörige Funktion y der n Variablen $x_0, x_1, \ldots x_{n-1}$, die wir eine metazyklische Funktion nennen können, genügt daher einer Gleichung $F(y) = 0$ vom Grade ν, deren Koeffizienten symmetrische Funktionen der x sind. Diese Gleichung ist irreduzibel im Bereich der symmetrischen Funktionen von x_i (§ 60, 2).

Substituiert man nun für x_i die Wurzeln einer metazyklischen Gleichung $f(x) = 0$, so wird y rational und $F(y) = 0$ wäre bei einer allgemeinen Gleichung fünften Grades eine Resolvente.

Wenn umgekehrt die Funktion y durch die Substitution der Wurzeln einer irreduziblen Gleichung $f(x) = 0$ für die x_i rational wird, während $F'(y)$ von Null verschieden bleibt, so ist $f(x)$ metazyklisch; denn dann ist die Gruppe von $f(x) = 0$ gewiß ein Teiler der vollen linearen Gruppe, und daher selbst linear; und da $f(x)$ irreduzibel ist, so ist die Gleichung $f(x) = 0$ metazyklisch. Es genügt aber auch für die algebraische Auflösbarkeit von $f(x) = 0$, wenn die Resolvente $F(y) = 0$ nur überhaupt eine rationale Wurzel hat, die nicht Doppelwurzel ist. Denn die verschiedenen Wurzeln dieser Gleichung gehören zu konjugierten Gruppen, und wenn daher eine von diesen Wurzeln rational ist, so ist eine der konjugierten Gruppen metazyklisch.

Wir haben also den Satz:

VI. Die notwendige und hinreichende Bedingung für die Lösbarkeit der Gleichung $f(x) = 0$ durch Radikale ist die, daß die Resolvente νten Grades $F(y) = 0$, die man durch passende Wahl der Funktion y so eingerichtet hat, daß sie keine Doppelwurzeln erhält, eine rationale Wurzel hat.

Eine andere Form dieser Bedingung ergibt sich auf folgende Weise.

Unter den Permutationen einer linearen Gruppe ist nur die identische, die irgend zwei Ziffern ungeändert läßt. Denn wenn $(z, az + b)$ zwei Ziffern nicht ändert, so muß die Kongruenz $az + b \equiv z \pmod{n}$ zwei verschiedene Lösungen haben. Das ist aber nur möglich, wenn $a \equiv 1$, $b \equiv 0 \pmod{n}$ ist.

Ist also die Gruppe P von $f(x)$ linear, so reduziert sie sich durch Adjunktion von zwei beliebigen Wurzeln auf die Einheit, und folglich sind alle Wurzeln rational durch zwei beliebige unter ihnen ausdrückbar. Also wie in § 83:

VII. Ist eine irreduzible Gleichung von Primzahlgrad metazyklisch, so sind alle Wurzeln rational durch zwei beliebige unter ihnen ausdrückbar.

Aber dieser Satz gilt auch umgekehrt, was wir folgendermaßen einfach beweisen können.

Es habe eine irreduzible Gleichung vom Primzahlgrad n die Eigenschaft, daß alle Wurzeln rational durch zwei beliebige unter ihnen ausdrückbar sind. Ist etwa $x_\nu = \psi(x_0, x_1)$ eine solche Darstellung, so können auf diese Relation alle Permutationen der Gruppe P unserer Gleichung angewandt werden, und wenn also eine von diesen Permutationen x_0 und x_1 ungeändert läßt, so läßt sie auch alle x_ν ungeändert, d. h. es ist die identische Permutation. Also enthält P außer der identischen Permutation keine, die zwei Ziffern nicht ändert.

Enthält nun eine der Permutationen $\pi = cc_1 \ldots$ von P zwei oder mehr verschiedene Zyklen c, $c_1 \ldots$ und ist c vom Grade h, c_1 aber von einem höheren Grade, so ist $\pi^h = c_1^h \ldots$, und π^h läßt die Ziffern von c ungeändert, während sie doch nicht die identische Substitution ist, weil c_1^h nicht alle Ziffern ungeändert läßt. Dies ist aber nicht möglich, wenn $h > 1$ ist. Es muß also entweder π eine Ziffer ungeändert lassen, oder es muß aus Zyklen

von gleichem Grade bestehen. Da aber n Primzahl ist, so kann π in diesem Falle nur einen Zyklus vom $n$ten Grade bilden.

Es enthält also P außer der identischen nur zyklische Permutationen $n$ten Grades, die wir mit γ bezeichnen wollen, und Permutationen $\varkappa$, die eine Ziffer ungeändert lassen und außerdem aus Zyklen von gleichem Grade bestehen.

Jede der Permutationen $\varkappa$, durch die die Ziffer 0 ungeändert bleibt, wollen wir mit $\varkappa_0$ bezeichnen. Ebenso bedeuten $\varkappa_1, \varkappa_2, \dots \varkappa_{n-1}$ die Permutationen $\varkappa$, die die Ziffern 1, 2, ... $n-1$ ungeändert lassen. Da P transitiv ist, so müssen ebensoviel $\varkappa_0$, wie $\varkappa_1$, wie $\varkappa_2$ usw. vorhanden sein.

Denn ist π eine Permutation, die 0 in 1 überführt, so ist jedes $\pi^{-1}\varkappa_0\pi$ ein $\varkappa_1$, und umgekehrt jedes $\pi\varkappa_1\pi^{-1}$ ein $\varkappa_0$, und ebenso für die übrigen Ziffern. Ist also μ die Anzahl der $\varkappa_0$, ν die Anzahl der γ, m der Grad von P, so ist, da noch die identische Permutation hinzukommt,

$$m = \mu n + \nu + 1. \tag{11}$$

Nun bilden die $\varkappa_0$ mit der identischen Permutation zusammen eine Gruppe Q vom Grade $\mu + 1$, und zwar einen Teiler von P, nämlich den Inbegriff aller Permutationen von P, die 0 ungeändert lassen. Sind also $\pi_1, \pi_2, \dots \pi_{n-1}$ Permutationen in P, die 0 in 1, 2, ... $n-1$ überführen, so ist $Q\pi_1$ der Inbegriff aller der Permutationen von P, die 0 in 1 überführen usf. Wir erhalten also die Zerlegung von P in die Nebengruppen

$$P = Q + Q\pi_1 + Q\pi_2 + \cdots + Q\pi_{n-1},$$

woraus folgt:

$$m = n(\mu + 1), \tag{12}$$

also aus (11) und (12) $\nu = n - 1$.

Es gibt also $n-1$ und nicht mehr zyklische Permutationen $n$ten Grades in P und diese bilden folglich wieder mit der Einheit zusammen eine zyklische Gruppe:

$$C = 1, \gamma, \gamma^2, \dots \gamma^{n-1},$$

da mit γ zugleich alle Potenzen von γ in P vorkommen.

Wenn nun γ eine zyklische Permutation $n$ten Grades ist, so ist auch jedes $\pi^{-1}\gamma^h\pi$ zyklisch und muß also, wenn π zu P gehört, auch in C enthalten sein. C ist also ein Normalteiler von P, und folglich muß nach dem Theorem II die Gruppe P linear sein.

VIII. Eine irreduzible Gleichung von Primzahlgrad, bei der alle Wurzeln rational durch zwei beliebige von ihnen ausdrückbar sind, ist metazyklisch.

Wir schließen aus diesen Sätzen noch auf eine merkwürdige, zuerst von Kronecker bemerkte Eigenschaft der metazyklischen Gleichungen für den Fall eines reellen Rationalitätsbereichs[1]). Wenn bei einer solchen Gleichung zwei Wurzeln reell sind, so folgt aus VII., daß alle ihre Wurzeln reell sind. Eine reelle Wurzel muß aber eine solche Gleichung, da sie ungeraden Grades ist, immer haben. Also folgt:

IX. Eine irreduzible metazyklische Gleichung von ungeradem Primzahlgrad mit reellen Koeffizienten hat entweder lauter reelle Wurzeln oder nur eine.

Sind alle Wurzeln reell, so ist die Diskriminante als Produkt von lauter Quadraten reeller Größen $(x_h - x_k)^2$ positiv. Ist nur eine Wurzel reell, so entspricht jedem komplexen Faktor der Diskriminante $(x_h - x_k)^2$ ein konjugierter, und deren Produkt ist positiv. Nur wenn x_h und x_k konjugiert imaginär, also $x_h - x_k$ rein imaginär ist, so ist $(x_h - x_k)^2$ negativ und die Diskriminante erhält also für jedes Paar konjugiert imaginärer Wurzeln den Faktor -1. Das Vorzeichen der Diskriminante ist also $(-1)^{\frac{n-1}{2}}$. Daraus folgt:

X. Wenn $n \equiv 1 \pmod 4$ ist, so ist die Diskriminante immer positiv, ist aber $n \equiv 3 \pmod 4$, so entscheidet das Vorzeichen der Diskriminante, welcher der beiden Fälle des Theorems IX eintritt.

§ 86.

Anwendung auf die metazyklischen Gleichungen fünften Grades.

Wir machen noch eine Anwendung der allgemeinen Sätze auf die Gleichungen fünften Grades. Ist $n = 5$, so hat die zyklische Gruppe 5, die halbmetazyklische 10 und die volle lineare Gruppe 20 Permutationen. Eine metazyklische Funktion genügt bei einer allgemeinen Gleichung fünften Grades einer

[1]) Über algebraisch auflösbare Gleichungen. Monatsbericht der Berliner Akademie, 14. April 1856.

Resolvente sechsten Grades, und wenn diese Gleichung sechsten Grades eine rationale Wurzel hat, so ist die gegebene Gleichung fünften Grades metazyklisch. Die halbmetazyklischen Funktionen, d. h. die Funktionen, die die Permutationen der halbmetazyklischen Gruppe gestatten, genügen einer Resolvente zwölften Grades, die aber nach Adjunktion der Quadratwurzel aus der Diskriminante in zwei Faktoren sechsten Grades zerfällt.

Wenn man die erzeugenden Substitutionen für den Modul 5

$$s = (z,\ z+1), \quad t = (z,\ 2z), \quad t^2 = (z,\ 4z)$$

auf die Ziffern 0, 1, 2, 3, 4 anwendet, so ergeben sich die Vertauschungen

(1)

		0,	1,	2,	3,	4
	(s)	1,	2,	3,	4,	0
	(t)	0,	2,	4,	1,	3
	(t^2)	0,	4,	3,	2,	1,

und es ist also, durch Zyklen dargestellt,

$$s = (0,\ 1,\ 2,\ 3,\ 4), \quad t = (1,\ 2,\ 4,\ 3), \quad t^2 = (1,\ 4)\ (2,\ 3).$$

Wenden wir die Substitutionen s, t^2, t auf die Paare von Ziffern (0, 1), (1, 2), (2, 3), (3, 4), (4, 0) an, so folgt:

(2)

	(0, 1),	(1, 2),	(2, 3),	(3, 4),	(4, 0)
(s)	(1, 2),	(2, 3),	(3, 4),	(4, 0),	(0, 1)
(t^2)	(0, 4),	(4, 3),	(3, 2),	(2, 1),	(1, 0);

diese Paare bleiben also in ihrer Gesamtheit durch s und durch t^2, und folglich durch die halbmetazyklische Gruppe ungeändert, und eine symmetrische Funktion der entsprechenden Wurzelpaare ist halbmetazyklisch. Die Substitution t bewirkt die Vertauschungen

(3)

	(0, 1),	(1, 2),	(2, 3),	(3, 4),	(4, 0)
(t)	(0, 2),	(2, 4),	(4, 1),	(1, 3),	(3, 0),

führt also die fünf Paare in fünf andere über; und da es nur zehn Paare von fünf Dingen gibt, so kommen in den beiden Reihen (3) alle Paare vor.

Man kann nun auf sehr verschiedene Arten halbmetazyklische Funktionen bilden. Die einfachste ist

$$u = x_0x_1 + x_1x_2 + x_2x_3 + x_3x_4 + x_4x_0, \tag{4}$$

die durch t in

$$u' = x_0x_2 + x_2x_4 + x_4x_1 + x_1x_3 + x_3x_0 \tag{5}$$

übergeht, und u' gehört selbst zu den halbmetazyklischen Funktionen.

Ist

$$f(x) = x^5 - ax^4 + bx^3 - cx^2 + dx - e = 0 \tag{6}$$

die Gleichung, deren Wurzeln x_0, x_1, x_2, x_3, x_4 sind, so ist also

$$u + u' = b. \tag{7}$$

Die Funktion

$$y = u - u' \tag{8}$$

ist gleichfalls halbmetazyklisch, das Quadrat y^2 aber ist vollmetazyklisch und also die Wurzel einer Resolvente sechsten Grades. Die allgemeine Gleichung fünften Grades hat also eine Resolvente sechsten Grades, wie schon Lagrange gefunden hat.

Ist $\sqrt{\Delta}$ das Produkt der zehn Wurzeldifferenzen $(x_0 - x_1)$ $(x_0 - x_2) \ldots$, also Δ die Diskriminante von $f(x)$, so ist

$$Y = \frac{u - u'}{\sqrt{\Delta}} \tag{9}$$

ebenfalls vollmetazyklisch und Wurzel einer Gleichung sechsten Grades [1]).

Um die konjugierten Funktionen zu u zu bilden, bezeichnen wir mit M die vollmetazyklische, mit N die halbmetazyklische, mit A die alternierende Gruppe und zerlegen A in die Nebengruppen:

$$A = N + N\,(1, 2)\,(3, 4) + Nt\,(0, 1) + Nt\,(0, 2) + Nt\,(0, 3) + Nt\,(0, 4). \tag{10}$$

Diese Nebengruppen sind in der Tat alle voneinander verschieden. Denn wäre z. B. $N = N\,(1, 2)\,(3, 4)$, so müßte $(1, 2)\,(3, 4)$ in N, also

$$(1, 2)\,(3, 4)\,t = (1, 2)\,(3, 4)\,(1, 2, 4, 3) = (1, 4)$$

in M vorkommen und dies ist nicht möglich, weil in M keine Permutation auftritt, die zwei Ziffern ungeändert läßt; und in ähnlicher Weise zeigt man, daß keine zwei anderen dieser Nebengruppen einander gleich sind.

Hiernach ergeben sich für u die innerhalb A konjugierten Werte durch Anwendung der Permutationen

---

[1]) Jacobi, „Observatiunculae ad theoriam aequationum pertinentes", Crelles Journ., Bd. 13. Jacobis Werke, Bd. 3. Cayley, Philos. Transactions 1861, Collected math. papers, Vol. IV, p. 309.

(1, 2) (3, 4), $t(0, 1) = (0, 1, 2, 4, 3)$, $t(0, 2) = (0, 2, 4, 3, 1)$
$t(0, 3) = (0, 3, 1, 2, 4)$; $t(0, 4) = (0, 4, 3, 1, 2)$:

$$
\begin{aligned}
u_1 &= x_0x_1 + x_1x_2 + x_2x_3 + x_3x_4 + x_4x_0,\\
u_2 &= x_0x_2 + x_2x_1 + x_1x_4 + x_4x_3 + x_3x_0,\\
u_3 &= x_1x_2 + x_2x_4 + x_4x_0 + x_0x_3 + x_3x_1,\\
u_4 &= x_2x_0 + x_0x_4 + x_4x_1 + x_1x_3 + x_3x_2,\\
u_5 &= x_3x_2 + x_2x_4 + x_4x_1 + x_1x_0 + x_0x_3,\\
u_6 &= x_4x_2 + x_2x_0 + x_0x_1 + x_1x_3 + x_3x_4,
\end{aligned}
\tag{11}
$$

und die entsprechenden u_1', u_2', u_3', u_4', u_5', u_6' findet man, wenn man auf u_1 die Permutation t anwendet, und dann mit u_i' ebenso verfährt, wie eben mit u_1, oder auch einfach, indem man die in jedem u fehlenden Paare zu dem entsprechenden u' vereinigt, so daß für alle diese u die Bedingung (7) befriedigt ist.

$$
\begin{aligned}
u_1' &= x_0x_2 + x_0x_3 + x_1x_3 + x_1x_4 + x_2x_4,\\
u_2' &= x_0x_1 + x_0x_4 + x_1x_3 + x_2x_3 + x_2x_4,\\
u_3' &= x_0x_1 + x_0x_2 + x_2x_3 + x_3x_4 + x_1x_4,\\
u_4' &= x_0x_1 + x_0x_3 + x_1x_2 + x_2x_4 + x_3x_4,\\
u_5' &= x_0x_2 + x_0x_4 + x_1x_2 + x_1x_3 + x_3x_4,\\
u_6' &= x_0x_3 + x_0x_4 + x_1x_2 + x_1x_4 + x_2x_3.
\end{aligned}
\tag{12}
$$

Die sechs Größen (8): $y_1, y_2, y_3, y_4, y_5, y_6$ sind die Wurzeln einer Gleichung sechsten Grades, deren Koeffizienten rational von a, b, c, d, e und $\sqrt{\Delta}$ abhängen. Da die y alle ihre Vorzeichen ändern, wenn $\sqrt{\Delta}$ das Vorzeichen ändert, so hat diese Gleichung die Form

$$y^6 + a_2y^4 + a_4y^2 + a_6 - \sqrt{\Delta}(a_1y^5 + a_3y^3 + a_5y) = 0,$$

worin die a rationale Funktionen der Koeffizienten von $f(x)$ sind. Es müssen aber auch ganze Funktionen dieser Koeffizienten sein; denn a_2, a_4, a_6, $a_1\sqrt{\Delta}$, $a_3\sqrt{\Delta}$, $a_5\sqrt{\Delta}$ sind ganze Funktionen der x, und es müssen also die drei letzteren durch das Differenzenprodukt $\sqrt{\Delta}$ teilbar sein, und dann müssen sich a_2, a_4, a_6, a_1, a_3, a_5 als ganze symmetrische Funktionen der x erweisen. Bestimmt man die Grade in bezug auf die x, so ergeben sich, da die y vom zweiten Grade sind, für

	$a_1\sqrt{\Delta}$,	a_2,	$a_3\sqrt{\Delta}$,	a_4,	$a_5\sqrt{\Delta}$,	a_6
die Grade	2	4	6	8	10	12.

$\sqrt{\Delta}$ ist aber vom zehnten Grade in den x, und folglich muß $a_1 = 0$, $a_3 = 0$ und a_5 eine Zahl sein.

Die Funktionen a_2, a_4, a_6 sind in bezug auf die Variablen x homogen, und wenn wir daher den Koeffizienten a, b, c, d, e die Gewichte 1, 2, 3, 4, 5 beilegen, so sind die a_2, a_4, a_6 isobarische Funktionen dieser Größen mit den Gewichten 4, 8, 12 (§ 28).

Man kann die Koeffizienten a_2, a_4, a_5, a_6 durch wirkliche Bildung der Ausdrücke nach den Vorschriften über die Darstellung der symmetrischen Funktionen berechnen, was von Cayley geschehen ist. Rechnung und Formeln sind aber sehr weitläufig [1]).

Wir wollen uns hier damit begnügen, die Resolvente für einen besonderen Fall zu bilden. Der dabei gefundene Wert für die Zahl a_5 ist dann natürlich allgemein gültig.

Die Gleichung fünften Grades habe die Bring-Jerrardsche Form [2])

$$x^5 + \alpha x + \beta = 0. \tag{13}$$

Dann sind a_2, a_4, a_6 ganze rationale Funktionen von α und β. Von den Koeffizienten der allgemeinen Gleichung (6) ist d vom Gewicht 4, e vom Gewicht 5, und daraus folgt, daß der Koeffizient e in a_2 gar nicht, in a_4, a_6 aber mit wenigstens einem der Koeffizienten a, b, c multipliziert vorkommen kann, und daß also, wenn a, b, c Null gesetzt werden, e aus diesen Ausdrücken herausgehen muß. Für die Gleichung (13) können also a_2, a_4, a_6 nicht von β abhängen, und man erhält mit Rücksicht auf das Gewicht, wenn m, m_1, m_2, m_3 Zahlen bedeuten,

$$a_2 = m_1\alpha, \quad a_4 = m_2\alpha^2, \quad a_6 = m_3\alpha^3, \quad a_5 = m. \tag{14}$$

Die Zahlen m, m_1, m_2, m_3 lassen sich aus einer speziellen Annahme bestimmen. Setzen wir $\beta = 0$, so wird

$$\begin{aligned} x_0 = 0, \quad x_1 &= \sqrt[4]{-\alpha}, \quad x_2 = i\sqrt[4]{-\alpha}, \\ x_3 &= -\sqrt[4]{-\alpha}, \quad x_4 = -i\sqrt[4]{-\alpha}, \end{aligned} \tag{15}$$

und daraus ergibt sich:

$$\begin{aligned} \sqrt{\Delta} &= (x_0 - x_1)(x_0 - x_2)(x_0 - x_3)(x_0 - x_4)(x_1 - x_2) \\ &\quad (x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4) \\ &= 16\, i\sqrt{-\alpha^5} = -16\sqrt{\alpha^5} \\ \Delta &= 256\,\alpha^5. \end{aligned} \tag{16}$$

Da hier $b = 0$ ist, so wird nach (7) und (8) $y = 2u$, und aus (11) und (15) ergeben sich die Werte der y:

[1]) Crelle, Bd. 113 (1894). Coll. math. Papers, Vol. XIII, p. 478.
[2]) Runge, Acta mathematica, Bd. 7.

$$y_1 = y_2 = y_3 = y_6 = -2\sqrt{\alpha}$$
$$y_4 = (4 - 2i)\sqrt{\alpha}, \quad y_5 = (4 + 2i)\sqrt{\alpha}.$$

Danach erhält man für $\beta = 0$ die folgende in y identische Gleichung:

$$(17) \quad \begin{aligned} & y^6 + a_2 y^4 + a_4 y^2 + a_6 - a_5 \sqrt{\Delta}\, y \\ &= (y + 2\sqrt{\alpha})^4 (y^2 - 8y\sqrt{\alpha} + 20\alpha) \\ &= y^6 - 20\alpha y^4 + 240\alpha^2 y^2 + 512\sqrt{\alpha^5}\, y + 320\alpha^3. \end{aligned}$$

Daraus findet man aber die allgemeine Resolvente für die Gleichung (13), wenn man nach (16) für $-16\sqrt{\alpha^5}$ setzt $\sqrt{\Delta}$, also

$$(18) \quad y^6 - 20\alpha y^4 + 240\alpha^2 y^2 - 32\sqrt{\Delta}\, y + 320\alpha^3 = 0.$$

Die Diskriminante ist eine ganze Funktion von α und β und ist in bezug auf die x_i homogen vom Grade 20 und weil α und β vom Grade 4 und 5 sind, so hat Δ die Form

$$\Delta = m\alpha^5 + n\beta^4,$$

worin m und n numerische Koeffizienten sind, die man aus den speziellen Annahmen $\alpha = 0$, $\beta = 0$ findet. Man erhält

$$(19) \quad \Delta = 2^8\alpha^5 + 5^5\beta^4.$$

Einfacher noch wird die Gleichung für u, nämlich

$$(20) \quad u^6 - 5\alpha u^4 + 15\alpha^2 u^2 - \sqrt{\Delta}\, u + 5\alpha^3 = 0,$$

und wenn man $u^2 = v$ setzt, so ergibt sich für v die rationale Gleichung sechsten Grades

$$(21) \quad (v^3 - 5\alpha v^2 + 15\alpha^2 v + 5\alpha^3)^2 = \Delta v.$$

Eine andere Form dieser Gleichung erhält man aus (17), wenn man den zweiten der drei Ausdrücke mit dem multipliziert, den man daraus erhält, wenn man $\sqrt{\alpha}$ in $-\sqrt{\alpha}$ verwandelt. Dadurch bekommt man

$$(v - \alpha)^4 (v^2 - 6\alpha v + 25\alpha^2) = 0,$$

und in diesen muß (21) übergehen, wenn $\beta = 0$ gesetzt wird. Da aber in (21) der von β abhängige Teil durch (19) bestimmt ist, so folgt die gesuchte Form der Resolvente

$$(22) \quad (v - \alpha)^4 (v^2 - 6\alpha v + 25\alpha^2) = 5^5\beta^4 v.$$

Man kann noch die Frage aufwerfen, ob die hier eingeführte Funktion v immer wirklich metazyklisch ist, ob sie nicht vielleicht bei besonderen Gleichungen noch andere Permutationen gestattet.

Wenn dieser Fall eintreten sollte, so müßte die Gleichung (21) oder (22) gleiche Wurzeln oder (20) gleiche oder entgegengesetzte Wurzeln haben.

Der Wert $u = 0$ oder $v = 0$ tritt nur dann ein, wenn $\alpha = 0$ ist. Dann haben wir in der Tat in (13) die wohlbekannte metazyklische Gleichung $x^5 + \beta = 0$; diesen Fall also lassen wir jetzt beiseite. Die Gleichung (20) kann nur dann zwei entgegengesetzte Wurzeln haben, wenn $\Delta = 0$ ist, wenn also die Gleichung (13) gleiche Wurzeln hat. Auch dies ist auszuschließen, da wir (13) als irreduzibel vorausgesetzt haben.

Es bleibt also nur die Frage übrig, ob (20) gleiche Wurzeln haben kann. Es müßte dann mit (20) zugleich die abgeleitete Gleichung

$$6u^5 - 20\alpha u^3 + 30\alpha^2 u - \sqrt{\Delta} = 0$$

befriedigt sein, und wenn wir $\sqrt{\Delta}$ aus dieser und aus (20) eliminieren,

$$5u^6 - 15\alpha u^4 + 15\alpha^2 u^2 - 5\alpha^3 = 5(u^2 - \alpha)^3 = 0.$$

Es müßte also $v = \alpha$ sein und also nach (22) $\beta = 0$. Dann wäre aber die Gleichung fünften Grades wieder reduzibel, da sie den Faktor x hat.

Will man also eine gegebene Gleichung von der Form (13) in bezug auf ihre Auflösbarkeit prüfen, so hat man zuerst die Irreduzibilität festzustellen, und dann zu untersuchen, ob (21) oder (22) eine rationale Wurzel hat.

Sind α und β ganze Zahlen, so muß auch ein rationales v eine ganze Zahl sein, die unter den Faktoren von $25\alpha^6$ zu suchen ist.

Ist z. B. $\alpha = 5$, $\beta = 5t$, so ist, wenn t eine durch 5 nicht teilbare ganze Zahl ist, $x^5 + \alpha x + \beta$ nach § 21, irreduzibel. Für v hat man Potenzen von 5, mit positivem oder negativem Vorzeichen einzusetzen, aber für keine solche Zahl kann (22) erfüllt sein, weil auf der linken Seite die Potenz von 5 nicht hoch genug wird.

Es ist also keine Gleichung von der Form $x^5 + 5x + 5t = 0$ metazyklisch.

Will man metazyklische Gleichungen ermitteln, so setze man in (22)

$$\beta = \alpha\mu, \qquad v = \alpha\lambda,$$

wodurch man erhält:

$$(23)\qquad \alpha = \frac{5^5 \mu^4 \lambda}{(\lambda - 1)^4 (\lambda^2 - 6\lambda + 25)}$$
$$\beta = \frac{5^5 \mu^5 \lambda}{(\lambda - 1)^4 (\lambda^2 - 6\lambda + 25)},$$

und wenn man hierin für λ und μ rationale Größen setzt (aus einem beliebigen Rationalitätsbereich, z. B. auch aus dem Körper der rationalen Funktionen von λ und μ), so ist die Gleichung fünften Grades

$$(24)\qquad x^5 + \alpha x + \beta = 0$$

algebraisch lösbar, da ja der Fall, daß diese Gleichung reduzibel wird, auch auf algebraisch lösbare Gleichungen führt.

Umgekehrt können wir sagen, daß keine irreduzible Gleichung fünften Grades von der Form (24) algebraisch lösbar ist, in der die Koeffizienten α, β nicht in der Form (23) darstellbar sind.

Setzen wir z. B. $\lambda = 3$, $\mu = 1$, so erhalten wir $2^8 \alpha = 3.5^5$, $2^8 \beta = 3.5^5$ und die Substitution $4x = 5\xi$ ergibt aus (24) die metazyklische Gleichung

$$\xi^5 + 15\xi + 12 = 0,$$

die überdies nach dem Schönemannschen Satz irreduzibel ist.

§ 87.

Die Gruppe der Resolvente.

Die Sätze, die wir im vorigen Paragraphen abgeleitet haben, gestatten einen merkwürdigen Schluß über die Gleichungen sechsten Grades.

Wir nehmen jetzt wieder die x_0, x_1, x_2, x_3, x_4 des vorigen Paragraphen als unabhängige Variable an. Dann sind die sechs Größen

$$(1)\qquad \begin{aligned} v_1 &= (u_1 - u_1')^2, \quad v_2 = (u_2 - u_2')^2, \quad v_3 = (u_3 - u_3')^2 \\ v_4 &= (u_4 - u_4')^2, \quad v_5 = (u_5 - u_5')^2, \quad v_6 = (u_6 - u_6')^2 \end{aligned}$$

die Wurzeln einer Gleichung sechsten Grades, $F(v) = 0$, deren Koeffizienten rational durch die symmetrischen Grundfunktionen a, b, c, d, e der x ausdrückbar sind, und die in dem Körper der rationalen Funktionen dieser Größen irreduzibel ist. Diese Gleichung ist eine Resolvente der Gleichung fünften Grades $f(x) = 0$, deren Wurzeln die x sind.

Ist M die volle lineare Gruppe der x, so haben die zu M konjugierten Gruppen $\pi^{-1} M \pi$ keinen anderen gemeinschaftlichen Teiler als die identische Permutation. Denn jeder gemeinsame Teiler aller dieser Gruppen müßte ein Normalteiler der symmetrischen und folglich auch der alternierenden Gruppe sein; da er aber nicht die alternierende Gruppe selbst sein kann, so muß er sich nach § 82 auf die Einheitsgruppe reduzieren.

Die Resolvente $F(v) = 0$ ist also nach der in § 60 eingeführten Bezeichnung eine Totalresolvente von $f(x) = 0$, und ihre Gruppe muß von gleichem Grade sein wie die Gruppe von $f(x) = 0$, deren Grad $1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ ist. Da $F(v)$ auch irreduzibel ist, so ist die Permutationsgruppe unter den Indices der sechs Größen v transitiv. Sie ist vom Grade 120, während der Grad der symmetrischen Gruppe der Permutationen von sechs Ziffern $6 \cdot 120$ ist. Damit haben wir den Satz:

Die symmetrische Permutationsgruppe von sechs Ziffern hat einen transitiven Divisor vom Index 6.

Um diese merkwürdige Gruppe, die wir mit C bezeichnen wollen, zu finden, haben wir nur den Einfluß zu untersuchen, den die sämtlichen 120 Permutationen der x auf die v oder auf die Größen (11) des vorigen Paragraphen ausüben.

Wir haben nun früher gesehen (§ 49, 2.), daß die Transpositionen (0, 1), (0, 2), (0, 3), (0, 4) durch wiederholte Zusammensetzung die ganze symmetrische Gruppe der Permutationen von fünf Ziffern erzeugen. Es genügt also, den Einfluß dieser vier Transpositionen auf die Indices der v zu ermitteln. Durch Anwendung einer Transposition (0, 1) geht jeder der Ausdrücke (11), § 86 in einen der Ausdrücke (12) über und umgekehrt, und wenn man für die vier genannten Transpositionen diese Änderung ermittelt, so erhält man die folgenden vier erzeugenden Permutationen $\pi_1, \pi_2, \pi_3, \pi_4$ der Gruppe C:

$$\begin{aligned}
(0, 1): &\quad \pi_1 = (1, 3)\ (2, 5)\ (4, 6)\\
(0, 2): &\quad \pi_2 = (1, 4)\ (2, 3)\ (5, 6)\\
(0, 3): &\quad \pi_3 = (1, 5)\ (2, 6)\ (3, 4)\\
(0, 4): &\quad \pi_4 = (1, 6)\ (2, 4)\ (3, 5),
\end{aligned}$$

wo sich die in den π vorkommenden Transpositionen (1, 3) ... auf die Indices der u und der v beziehen.

Unter den 120 Permutationen der Gruppe C kommen unter anderen auch vor:

$$\begin{aligned}
\pi_1\pi_2 &= (1, 2, 6)(3, 4, 5)\\
\pi_1\pi_2\pi_3 &= (1, 6, 5, 4)\\
\pi_1\pi_2\pi_3\pi_4 &= (2, 4, 6, 3, 5)\\
(\pi_1\pi_2\pi_3\pi_4)^3 &= (2, 3, 4, 5, 6).
\end{aligned}$$

Hiernach läßt sich leicht eine zur Gruppe C gehörige Funktion der sechs Größen v bilden.

Eine Funktion, wie

$$v_1 v_3 + v_2 v_4 + v_5 v_6,$$

bleibt durch π_1 und durch $\pi_1\pi_2$ ungeändert; wenden wir darauf aber die Potenzen der zyklischen Permutation (2, 3, 4, 5, 6) an, so gehen daraus fünf Funktionen hervor, die wir so bezeichnen wollen:

$$\begin{aligned}
w_0 &= v_1 v_2 + v_4 v_5 + v_3 v_6\\
w_1 &= v_1 v_4 + v_2 v_6 + v_3 v_5\\
w_2 &= v_1 v_6 + v_2 v_5 + v_3 v_4\\
w_3 &= v_1 v_3 + v_2 v_4 + v_5 v_6\\
w_4 &= v_1 v_5 + v_2 v_3 + v_4 v_6.
\end{aligned} \tag{2}$$

Wenn wir auf die Funktionen (2) die Permutationen π_1, π_2, π_3, π_4 anwenden, so gehen diese Funktionen nur ineinander über, und zwar in folgender Weise:

$$\pi_1 = (w_0, w_1), \quad \pi_2 = (w_0, w_2), \quad \pi_3 = (w_0, w_3), \quad \pi_4 = (w_0, w_4),$$

d. h. die π_1, π_2, π_3, π_4 entsprechen den Transpositionen (0, 1), (0, 2), (0, 3), (0, 4) unter den Indices der w, und also entspricht nach § 49, 2. der ganzen Gruppe C die symmetrische Gruppe der Indices der w. Eine symmetrische Funktion der fünf Größen w, die nicht zugleich eine symmetrische Funktion der v ist, wie die Summe der w, ist also eine zu C gehörige Funktion.

Man kann z. B. die Summe der Quadrate dafür nehmen:

$$W = w_0^2 + w_1^2 + w_2^2 + w_3^2 + w_4^2,$$

oder eine Funktion

$$(\lambda - w_0)(\lambda - w_1)(\lambda - w_2)(\lambda - w_3)(\lambda - w_4)$$

für ein beliebiges rationales λ.

Eine solche Größe W ist die Wurzel einer irreduziblen Gleichung sechsten Grades, deren Koeffizienten symmetrische Funktionen der sechs Größen v sind.

§ 88.

Wurzeln metazyklischer Gleichungen.

Wir haben in den letzten Paragraphen allgemeine Kennzeichen gefunden, durch die man entscheiden kann, ob eine vorgelegte Gleichung von Primzahlgrad metazyklisch ist oder nicht. Damit ist aber der Gegenstand noch bei weitem nicht erschöpft. Es handelt sich vielmehr nach der Form der Problemstellung, wie sie von Abel herrührt, darum, ein Verfahren anzugeben, nach dem man alle metazyklischen Gleichungen finden kann. In dieser Form ist das Problem, über das Abel nur kurze Andeutungen ohne Beweise hinterlassen hat, von Kronecker aufgenommen worden, und in der Weise vollständig gelöst, daß zunächst nicht die Gleichungen selbst, sondern ihre Wurzeln gefunden werden[1]). So hat Kronecker für jede Primzahl n einen Ausdruck angegeben, der aus einem gegebenen Körper Ω durch wiederholtes Wurzelziehen abgeleitet ist, und der die doppelte Eigenschaft hat, daß jede Wurzel einer irreduziblen metazyklischen Gleichung $n$ten Grades in Ω in diesem Ausdruck enthalten ist, und daß auch umgekehrt jeder solche Ausdruck einer irreduziblen Gleichung $n$ten Grades in Ω genügt.

Die Grundlage für diese Untersuchung liefern uns die Resolventen von Lagrange.

Es sei jetzt n eine Primzahl, $x_0, x_1, x_2 \ldots x_{n-1}$ sei ein System unabhängiger Variablen, und die Bezeichnung so gewählt, daß $x_n = x_0, x_{n+1} = x_1, \ldots$ ist. Endlich sei ε eine $n$te Einheitswurzel. Wir setzen dann

$$(1)\quad (\varepsilon, x) = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots + \varepsilon^{n-1} x_{n-1} = \sum_{0,n-1}^{h} \varepsilon^h x_h.$$

Diese Ausdrücke haben wir schon früher unter dem Namen der Resolventen von Lagrange kennen gelernt, und wir

[1]) Abel, Sur la résolution algébrique des équations (Oeuvres complètes, ed. Sylow, tome II, p. 217), Brief an Crelle vom 14. März 1826 (Oeuvres, tome II, p. 266). — Kronecker, Über die algebraisch auflösbaren Gleichungen. Monatsberichte der Berliner Akademie 1853, 1856. — Cayley, On a theorem of Abel's etc. (1880, Coll. math. papers, t. XI, p. 132). — H. Weber, Über algebraisch auflösbare Gleichungen von Primzahlgrad. Sitzungsberichte der Gesellschaft zur Beförderung der Naturwissenschaften zu Marburg 1892. Die folgende Darstellung ist gegen die dort gegebene vereinfacht und in einem Punkte berichtigt. Wiman, Die metazyklischen Gleichungen von Primzahlgrad. Acta math., Bd. 27, 1903.

haben dort (§ 64) ihr Verhalten gegenüber den zyklischen Permutationen untersucht.

Es kommt jetzt darauf an, den Einfluß linearer Permutationen von der im § 85 betrachteten Art auf diese Funktionen festzustellen.

Wir schicken einen Hilfssatz voraus, den wir eigentlich nur aus dem Früheren zu reproduzieren brauchen. Die imaginären $n$ten Einheitswurzeln ε sind die Wurzeln einer Gleichung $X = 0$, wenn

$$X = \xi^{n-1} + \xi^{n-2} + \cdots + \xi + 1 \tag{2}$$

gesetzt ist, und ξ eine Variable bedeutet.

Die Funktion X ist, wie wir früher gesehen haben (§ 73), irreduzibel im Körper R der rationalen Zahlen.

Legen wir irgend einen Körper Ω zugrunde, in dem X irreduzibel ist, und leiten daraus den Körper $\Omega(\varepsilon)$ ab, der aus allen rationalen Funktionen von ε mit Koeffizienten in Ω besteht, so ergibt sich nach § 53, daß jede Größe dieses Körpers auf eine und nur auf eine Weise in die Form

$$c_0 + c_1\varepsilon + c_2\varepsilon^2 + \cdots + c_{n-2}\varepsilon^{n-2} \tag{3}$$

gesetzt werden kann, worin die c Größen in Ω sind.

Der Kürze wegen wollen wir diese Form der Größen in $\Omega(\varepsilon)$ die Normalform nennen und wollen also den Satz aussprechen:

1. Ist Ω ein Körper, in dem die Kreisteilungsgleichung $X = 0$ irreduzibel ist, so kann jede Größe des Körpers $\Omega(\varepsilon)$ nur auf eine Weise in die Normalform

$$\Phi(\varepsilon) = c_0 + c_1\varepsilon + c_2\varepsilon^2 + \cdots + c_{n-2}\varepsilon^{n-2}$$

gebracht werden, worin die Koeffizienten c in Ω enthalten sind.

Wir können in diesem Satze für Ω den Körper nehmen, der aus R durch Adjunktion von n unabhängigen Variablen x_0, x_1, $x_2, \ldots x_{n-1}$ entsteht; denn wäre X in diesem Körper reduzibel, also

$$X = X_1 X_2,$$

wo X_1, X_2 ganze Funktionen von ξ und rationale Funktionen der x sind, so könnte man für die Variablen x_0, $x_1, \ldots x_{n-1}$ solche rationale Zahlen setzen, daß X_1, X_2, ohne ihren Grad in

bezug auf ξ zu ändern, in Funktionen in R übergehen (§ 14), und dieses widerspricht der Irreduzibilität von X in R.

Dieser Körper Ω hat verschiedene Divisoren, in denen X gleichfalls irreduzibel ist, und die also alle im Theorem 1 für Ω genommen werden können. Solche Divisoren erhält man wenn man irgend eine Permutationsgruppe P der n Ziffern $0, 1. 2 \ldots n-1$ festsetzt und den Inbegriff aller rationalen Funktionen der $x_0, x_1 \ldots x_{n-1}$ in R betrachtet, die die Permutationen dieser Gruppe gestatten.

Der Inbegriff aller dieser Funktionen bildet offenbar einen Körper, und so bekommt man zu jeder Permutationsgruppe der n Ziffern einen bestimmten zugehörigen Körper Ω. Man kann die Größen eines solchen Körpers rational darstellen durch die symmetrischen Funktionen der x und durch eine zu der Gruppe P gehörige Funktion.

Wir nennen den so bestimmten Körper Ω den zu der Gruppe P gehörigen Körper. Jeder solche Körper kann in 1. die Stelle von Ω vertreten.

Wir können also den zweiten Satz aufstellen:

2. Wenn eine rationale Funktion $\Phi(x_0, x_1 \ldots x_{n-1}, \varepsilon)$ des Körpers $\Omega(\varepsilon)$ ungeändert bleibt, wenn auf die Indices der x die Permutationen der Gruppe P angewandt werden, so gestatten auch die Koeffizienten $c_0, c_1, c_2 \ldots c_{n-2}$ der Normalform von Φ die Permutationen von P.

Ist $\Phi(\varepsilon)$ irgend ein Element des Körpers $\Omega(\varepsilon)$, so erhält man die konjugierten Größen, wenn man für ε irgend eine andere Wurzel von $X = 0$ setzt, also die Substitution $(\varepsilon, \varepsilon^h)$ für $h = 1, 2, \ldots n-1$ ausführt. Alle diese Körper sind wieder in $\Omega(\varepsilon)$ enthalten; sie sind also miteinander identisch, d. h. $\Omega(\varepsilon)$ ist ein Normalkörper über Ω (§ 56). Wenn eine Größe $\Phi(\varepsilon)$ die Eigenschaft hat, ungeändert zu bleiben, wenn ε durch alle ε^h ersetzt wird, so ist sie selbst in Ω enthalten. Dies folgt schon aus den allgemeinen Sätzen des § 55. Wir sehen aber die Richtigkeit sofort ein, wenn wir unter der Voraussetzung $\Phi(\varepsilon^h) = \Phi(\varepsilon)$

$$\Phi(\varepsilon) = \frac{1}{n-1} \sum_{1, n-1}^{h} \Phi(\varepsilon^h) = c_0 - \frac{c_1 + c_2 + \cdots + c_{n-2}}{n-1}$$

setzen, woraus dann nach 1. zu schließen ist, daß $c_1, c_2, \ldots c_{n-2}$ gleich Null sind. Also:

3. Wenn eine Größe in $\Omega(\varepsilon)$ die sämtlichen Substitutionen $(\varepsilon, \varepsilon^h)$ gestattet, so ist sie in Ω enthalten.

§ 89.

Sätze über die Resolventen.

Wir betrachten jetzt die Lagrangeschen Resolventen (ε, x) unter der Voraussetzung, daß ε eine imaginäre $n$te Einheitswurzel und die x unabhängige Variable sind. und wenden darauf die Sätze des vorigen Paragraphen an.

Wir untersuchen zunächst den Einfluß der linearen Permutationen auf (ε, x); dazu genügt es, wenn wir den Einfluß der beiden erzeugenden Substitutionen

$$(1) \qquad s = (h, h+1), \qquad t = (h, gh)$$

feststellen, worin das nach dem Modul n genommene h die Indices von x durchläuft und g eine primitive Wurzel der Primzahl n bedeutet (§ 85).

Wenden wir aber auf

$$(2) \qquad (\varepsilon, x) = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots + \varepsilon^{n-1} x_{n-1}$$

die Substitution s an, so geht diese Funktion über in

$$x_1 + \varepsilon x_2 + \varepsilon^2 x_3 + \cdots + \varepsilon^{n-1} x_0 = \varepsilon^{-1}(\varepsilon, x).$$

Der Einfluß der Substitution s ist also, daß

$$(\varepsilon, x) \quad \text{in} \quad \varepsilon^{-1}(\varepsilon, x)$$

übergeht.

Gehen wir nun zur Substitution t über und setzen in

$$(\varepsilon, x) = x_0 + \sum_{1, n-1}^{h} \varepsilon^h x_h$$

x_{gh} für x_h, so geht diese Funktion über in

$$(3) \qquad x_0 + \sum_{1, n-1}^{h} \varepsilon^h x_{gh}.$$

In dieser Summe kann h ein beliebiges Restsystem nach dem Modul n durchlaufen mit Ausschluß von 0, und wir können also h durch $g^{-1}h$ ersetzen; dadurch wird die Summe (3)

$$x_0 + \sum^{h} \varepsilon^{g^{-1}h} x_h = (\varepsilon^{g^{-1}}, x).$$

Wendet man die Permutationen s und t wiederholt an, so ergibt sich der Einfluß von s^λ, t^λ, der in den Vertauschungen

$$(\varepsilon, x), \quad \varepsilon^{-\lambda}(\varepsilon, x)$$
$$(\varepsilon, x), \quad (\varepsilon^{g^{-\lambda}}, x)$$

besteht, speziell also für $\lambda = -1$

$$(\varepsilon, x), \quad \varepsilon(\varepsilon, x)$$
$$(\varepsilon, x), \quad (\varepsilon^{g}, x),$$

und wir bekommen den folgenden Satz:

4. Die erzeugenden Permutationen s, t der linearen Gruppe bewirken die Vertauschungen

(s): (ε, x), $\varepsilon^{-1}(\varepsilon, x)$; (s^{-1}): (ε, x), $\varepsilon(\varepsilon, x)$

(t): (ε, x), $(\varepsilon^{g^{-1}}, x)$; (t^{-1}): (ε, x), (ε^{g}, x).

Aus 2. und 4. aber ergibt sich, weil s die erzeugende Permutation der zyklischen Gruppe ist, die Folgerung:

5. Stellt man die Funktionen

$$(\varepsilon, x)^n = \Phi(\varepsilon), \quad (\varepsilon^{\lambda}, x)\,(\varepsilon, x)^{-\lambda} = F(\varepsilon) \tag{4}$$

in der Normalform [§ 88, (3)] dar, so sind die Koeffizienten zyklische Funktionen von $x_0, x_1, \ldots x_{n-1}$ (§ 88, 2).

Hierin kann λ jede beliebige positive oder negative ganze Zahl, die nicht durch n teilbar ist, bedeuten.

Aus der Funktion $F(\varepsilon)$ entspringen, wenn wir $\lambda = g$ annehmen und für ε seine $n-1$ verschiedenen Werte setzen, $n-1$ Funktionen, die alle durch s ungeändert bleiben, und die wir folgendermaßen bezeichnen wollen:

$$\begin{aligned}
(\varepsilon^{g}, x)\,(\varepsilon, x)^{-g} &= f_0 \\
(\varepsilon^{g^2}, x)\,(\varepsilon^{g}, x)^{-g} &= f_1 \\
&\cdots\cdots\cdots\cdots \\
(\varepsilon^{g^{n-1}}, x)\,(\varepsilon^{g^{n-2}}, x)^{-g} &= f_{n-2},
\end{aligned} \tag{5}$$

so daß also, wenn $(\varepsilon^{g}, x), (\varepsilon, x)^{-g} = f(\varepsilon)$ gesetzt wird,

$$f_h = (\varepsilon^{g^{h+1}}, x)\,(\varepsilon^{g^h}, x)^{-g} = f(\varepsilon^{g^h}) \tag{6}$$

ist. Wendet man auf (5) die Permutation t^{-1} an, so erleiden, wie nach dem Theorem 4. zu sehen ist, die Funktionen

$$f_0, f_1, f_2 \cdots f_{n-2}$$

eine zyklische Permutation $(0, 1, 2, \ldots n-2)$.

Außerdem sind die Funktionen f_i, solange die x variabel sind, wie man aus (5) ersieht, alle voneinander verschieden, da sie in verschiedene lineare Faktoren zerlegt sind.

Bilden wir also eine zyklische Funktion dieser $n-1$ Größen

$$C(f_0, f_1, f_2 \dots f_{n-2}), \tag{7}$$

so bleibt diese sowohl durch s als durch t^{-1}, also durch t und folglich durch die ganze lineare Gruppe ungeändert, und wir erhalten das Theorem:

6. Stellt man eine zyklische Funktion der Größen $f_0, f_1, \dots f_{n-2}$ in der Normalform dar, so sind die Koeffizienten metazyklische Funktionen der Variablen $x_0, x_1, \dots x_{n-1}$.

Hierbei ist unter einer metazyklischen Funktion jede Funktion zu verstehen, die durch die Permutationen der linearen Gruppe ungeändert bleibt.

Die Funktionen $f_0, f_1, \dots f_{n-2}$ gehen aber auch zyklisch ineinander über, wenn man die x ungeändert läßt, und für ε die Substitutionen $(\varepsilon, \varepsilon^g)$ macht. Wenn also die zyklische Funktion C der Größen f in ihren Koeffizienten ε nicht enthält, so bleibt C durch sämtliche Substitutionen $(\varepsilon, \varepsilon^h)$ ungeändert, und wir erhalten aus 3. das Theorem:

7. Ist $C(f_0, f_1, \dots f_{n-2})$ eine zyklische Funktion der Größen $f_0, f_1, \dots f_{n-2}$ mit rationalen Koeffizienten, so ist C eine metazyklische Funktion der Variablen $x_0, x_1 \dots x_{n-1}$ mit rationalen Koeffizienten.

Wenn wir die Funktionen $f_0, f_1, \dots f_{n-2}$, wie sie durch (5) gegeben sind, der Reihe nach zu den Potenzen $g^{n-2}, g^{n-3}, \dots 1$ erheben und dann alles multiplizieren, und wenn wir noch beachten, daß $\varepsilon^{g^{n-1}} = \varepsilon$ ist, so heben sich im Produkt der linken Seite der Gleichungen (5) alle Resolventen mit Ausnahme von (ε, x) heraus, und es folgt

$$(\varepsilon, x)^{1-g^{n-1}} = f_0^{g^{n-2}} f_1^{g^{n-3}} \dots f_{n-2}. \tag{8}$$

Bezeichnen wir nun mit λ einen noch unbestimmten ganzzahligen Exponenten, so leiten wir aus (8) die folgende Relation her:

$$(\varepsilon^\lambda, x)^n = \left[(\varepsilon, x)^{\frac{g^{n-1}-1}{n}} (\varepsilon^\lambda, x)\right]^n f_0^{g^{n-2}} f_1^{g^{n-3}} \dots f_{n-2}. \tag{9}$$

Verstehen wir unter g_1 irgend eine feste primitive Wurzel von n, so können wir $g = g_1 + ln$ setzen, worin l eine beliebige ganze Zahl bedeutet, da ja die primitiven Wurzeln nur bis auf Vielfache von n definiert sind. Dann ist:

$$g^{n-1} \equiv g_1^{n-1} + n(n-1)\,l\,g_1^{n-2} \pmod{n^2},$$

woraus

$$\frac{g^{n-1}-1}{n} \equiv \frac{g_1^{n-1}-1}{n} - l\,g_1^{n-2} \pmod{n},$$

und l läßt sich so bestimmen, daß

$$-\frac{g^{n-1}-1}{n} \equiv 1 \pmod{n}$$

wird.

In (9) setzen wir nun unter dieser Voraussetzung

$$\lambda = -\frac{g^{n-1}-1}{n} \equiv 1 \pmod{n}, \tag{10}$$

und wie in (4)

$$F(\varepsilon) = (\varepsilon^\lambda, x)\,(\varepsilon, x)^{-\lambda} = (\varepsilon, x)^{nk}, \tag{11}$$

worin $\varepsilon^\lambda = \varepsilon$ und $1 - \lambda$, das nach (10) durch n teilbar ist, $= nk$ gesetzt ist. Daraus erhalten wir aus (9):

$$(\varepsilon, x)^n = [F(\varepsilon)]^n f_0^{g^{n-2}} f_1^{g^{n-3}} \cdots f_{n-2}. \tag{12}$$

Diese Formeln lassen sich verallgemeinern, wenn man beachtet, daß die Funktionen $f_0, f_1, \ldots f_{n-2}$ eine zyklische Permutation erleiden, wenn die Substitution $(\varepsilon, \varepsilon^g)$ ausgeführt wird. Setzen wir also fest, daß $f_h = f_k$ sein soll, wenn $h \equiv k \pmod{n-1}$ ist, und setzen

$$F(\varepsilon^{g^\nu}) = F_\nu = (\varepsilon^{g^\nu}, x)^{nk}, \tag{13}$$

so ergibt die Substitution $(\varepsilon, \varepsilon^{g^\nu})$ in (12):

$$(\varepsilon^{g^\nu}, x)^n = F_\nu^n f_\nu^{g^{n-2}} f_{\nu+1}^{g^{n-3}} \cdots f_{\nu+n-2}. \tag{14}$$

Diese Formeln gelten für jedes ν, wenn wir auch noch in bezug auf die F_ν festsetzen, daß $F_h = F_k$ sein soll, wenn $h \equiv k \pmod{n-1}$ ist. Es ergibt sich dann aus 4., daß die Funktionen $F_0, F_1 \ldots F_{n-2}$ durch die Substitution s ungeändert bleiben, und durch t^{-1} eine zyklische Permutation erfahren.

Wenn wir hier die Exponenten auf ihre kleinsten positiven Reste nach dem Modul n reduzieren wollen, so setzen wir:

$$g^\nu = nq_\nu + r_\nu, \quad 0 < r < n, \quad r_0 = 1, \tag{15}$$

und die Zahlen r_ν fallen in irgend einer Reihenfolge mit den Zahlen $1, 2 \ldots n-1$ zusammen, so daß r_0 immer $= 1$ ist.

Dann wird nach (5):

$$f_{\nu-1} = (\varepsilon^{r_\nu}, x)\,(\varepsilon^{r_\nu - 1}, x)^{-g}, \tag{16}$$

und wir setzen noch:

$$\Phi_\nu = F_\nu f_\nu^{q_n-2} f_{\nu+1}^{q_n-3} \dots f_{\nu+n-3}^{q_1}, \tag{17}$$

und erhalten aus (14):

$$(\varepsilon^{r_\nu}, x)^n = \Phi_\nu^n f_\nu^{r_n-2} f_{\nu+1}^{r_n-3} \dots f_{\nu+n-2}^{r_0}. \tag{18}$$

Die Exponenten r_ν bleiben ungeändert, wenn g durch $g - ln$ ersetzt wird, sind also von der Bedingung (10) unabhängig und können aus einer beliebigen primitiven Wurzel g abgeleitet werden.

Die hier vorkommenden Funktionen f_ν, F_ν, Φ_ν, deren Gesamtheit wir mit ω_ν bezeichnen wollen, enthalten außer den Variablen $x_0, x_1, x_2, \dots x_{n-2}$ und der $n$ten Einheitswurzel ε nur rationale Zahlen. Ihre charakteristischen Eigenschaften sind, um es nochmals zu wiederholen, folgende:

α) Durch die zyklische Permutation $s = (h, h+1)$ unter den Indices von x ändern sich die Funktionen ω_ν nicht.

β) Durch die lineare Permutation $t^{-1} = (h, g^{-1}h)$ unter den Indices der x wird unter den Indices der ω die zyklische Permutation $(0, 1, 2, \dots n-2)$ hervorgerufen.

γ) Durch die Substitution $\sigma = (\varepsilon, \varepsilon^g)$ wird unter den Indices von ω dieselbe zyklische Permutation $(0, 1, 2, \dots n-2)$ bewirkt.

δ) Die zyklischen Funktionen von ω_ν sind metazyklische Funktionen der x und von ε frei.

Nach dem Satze von Lagrange (§ 60, 3) kann man jede Funktion Θ_ν, der die vier Eigenschaften α), β), γ), δ) zukommen, rational im Körper Ω der metazyklischen Funktionen von x durch eine von ihnen, ω_ν, ausdrücken, wenn nur die $\omega_0, \omega_1, \dots \omega_{n-2}$ voneinander verschieden sind; man kann also für diese Funktionen die f_ν wählen.

Denn bedeutet u eine Variable, und ist

$$(u - \omega_0)(u - \omega_1) \dots (u - \omega_{n-2}) = \varphi(u), \tag{19}$$

so gestattet die Funktion $\varphi(u)$ die Permutationen s, t und auch die Substitution $\sigma = (\varepsilon, \varepsilon^g)$, und ist also nach dem Satze 3. des § 88 eine ganze Funktion von u mit in bezug auf x metazyklischen Koeffizienten und unabhängig von ε.

Ist nun also $\Theta_0, \Theta_1, \dots \Theta_{n-2}$ ein Funktionensystem, dem die Eigenschaften α), β), γ), δ) zukommen, so ist die Summe

$$(20) \qquad \sum_{0,n-2}^{\nu} \frac{\Theta_\nu \varphi(u)}{u - \omega_\nu} = \chi(u)$$

eine ganze Funktion $(n-2)$ten Grades von u, die gleichfalls die Substitutionen s, t, σ gestattet und die also Koeffizienten in Ω hat. Setzt man

$$(21) \qquad \frac{\chi(u)}{\varphi'(u)} = \Theta(u)$$

und setzt dann in (20) $u = \omega_\nu$, so folgt

$$\Theta_\nu = \Theta(\omega_\nu),$$

worin Θ eine rationale Funktion bedeutet, deren Koeffizienten metazyklische Funktionen der x sind und ε nicht mehr enthalten.

§ 90.

Wurzeln irreduzibler metazyklischer Gleichungen.

Es sollen jetzt in den Formeln des vorigen Paragraphen für die Variablen $x_0, x_1, \ldots x_{n-1}$ die Wurzeln $\xi_0, \xi_1, \ldots \xi_{n-1}$ einer in irgend einem Körper $\mathfrak{R}$ irreduziblen metazyklischen Gleichung von Primzahlgrade n eingeführt werden, und zwar in der Reihenfolge, daß die metazyklischen Funktionen der ξ rational (in $\mathfrak{R}$) sind, was nach den Sätzen des § 85 immer möglich ist. Wir machen aber dabei zunächst noch zwei beschränkende Vorausetzungen, von denen wir nachträglich das Resultat wieder befreien werden. Diese Voraussetzungen sind:

1. daß durch diese Substitution der ξ für die x keine der Resolventen (ε, ξ), in der ε eine imaginäre $n$te Einheitswurzel ist, verschwindet.

Nach (13), (16), (17), § 89 bekommen dann die Funktionen f_ν, F_ν, Φ_ν bestimmte von Null verschiedene Werte und wir machen die zweite Voraussetzung,

2. daß durch dieselbe Substitution nicht zwei der Funktionen $f_0, f_1, \ldots f_{n-2}$ einander gleich werden.

Wir nehmen an, es gehe durch die Substitution der ξ für die x

$$f_0, f_1, \ldots f_{n-2}$$

in

$$k_0, k_1, \ldots k_{n-2}$$

über, so daß also die $k_0, k_1, \ldots k_{n-2}$ voneinander verschieden sind, und

$$\Phi_0, \Phi_1, \ldots \Phi_{n-2}$$

in

$$K_0, K_1, \ldots K_{n-2},$$

wo aber unter den K auch gleiche vorkommen können.

Wegen der Eigenschaft δ) der Funktionen f_ν sind die Größen $k_0, k_1, \ldots k_{n-2}$ die Wurzeln einer zyklischen Gleichung $(n-1)$ten (also geraden) Grades $\psi(u) = 0$, so daß, wenn u eine Variable bedeutet,

(1) $$\psi(u) = (u - k_0)(u - k_1) \ldots (u - k_{n-2})$$

eine Funktion $(n-1)$ten Grades von u mit rationalen Koeffizienten ist. Die Größen $k_0, k_1, \ldots k_{n-2}$ können rational durcheinander ausgedrückt werden in der Form:

(2) $k_1 = \Theta(k_0)$, $k_2 = \Theta(k_1)$, $\ldots k_{n-2} = \Theta(k_{n-3})$, $k_0 = \Theta(k_{n-2})$ (§ 63).

Ebenso können die Größen K_ν nach dem Satze von Lagrange durch die k_ν ausgedrückt werden, und zwar in der Form:

(3) $$K_0 = \Phi(k_0), \quad K_1 = \Phi(k_1), \ldots K_{n-2} = \Phi(k_{n-2}),$$

worin Φ eine rationale Funktion (in $\mathfrak{R}$) bedeutet.

Danach liefert uns die Formel (18), § 89 folgendes Resultat:

Wir setzen zur Abkürzung

(4) $$\tau_\nu = \sqrt[n]{k_\nu},$$

und erhalten:

(5) $$(\varepsilon^{r_\nu}, \xi) = K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \ldots \tau_{\nu+n-2}^{r_0},$$

und wenn man diese Formeln alle addiert, und noch die rationale Größe $(1, x) = A$ setzt:

(6) $$n\xi_0 = A + \sum_{0,n-2}^{\nu} K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \ldots \tau_{\nu+n-2}^{r_0}.$$

Nach § 89 (16) und (17) sind die Größen K_ν und k_ν alle von Null verschieden.

Durch (6) ist eine der Wurzeln ξ dargestellt.

Setzen wir zur Abkürzung

$$R_\nu = k_\nu^{r_{n-2}} k_{\nu+1}^{r_{n-3}} \ldots k_{\nu+n-2}^{r_0}$$

und schreiben (6) in der Form:

(7) $$n\xi_0 = A + K_0\sqrt[n]{R_0} + K_1\sqrt[n]{R_1} + \cdots + K_{n-2}\sqrt[n]{R_{n-2}},$$

so ist also ξ_0 durch $n-1$ Radikale $n$ten Grades ausgedrückt, von denen jedes an sich n verschiedene Werte haben kann. Diese Werte können aber nicht voneinander unabhängig sein,

weil sonst die Anzahl der Werte von ξ, die sich aus (7) ergeben, zu groß wäre. Man erhält in der Tat aus (5) und (7), wenn man nach § 89, (16)

$$(\varepsilon^{r_\nu}, \xi)\ (\varepsilon^{r_{\nu-1}}, \xi)^{-g} = k_{\nu-1}$$

setzt,

$$K_\nu \sqrt[n]{R_\nu} = K_{\nu-1}^g k_{\nu-1} \left(\sqrt[n]{R_{\nu-1}}\right), \tag{8}$$

und kann also hiernach alle diese Radikale rational durch eines von ihnen und durch k_ν ausdrücken.

Der Ausdruck (6) hat aber vor (7) den großen Vorzug, daß er, wie man auch die $n-1$ Radikale $\sqrt[n]{k_\nu}$ bestimmen mag, doch nur n verschiedene Werte, nämlich die n Wurzeln ξ, darstellt.

Um dies nachzuweisen, bezeichnen wir mit $\varepsilon_0, \varepsilon_1, \ldots \varepsilon_{n-2}$ irgend ein beliebiges System $n$ter Einheitswurzeln, und ersetzen in (5):

$$\sqrt[n]{k_0}, \quad \sqrt[n]{k_1} \ldots \quad \sqrt[n]{k_{n-2}}$$

durch

$$\varepsilon_0 \sqrt[n]{k_0}, \ \varepsilon_1 \sqrt[n]{k_1} \ldots \varepsilon_{n-2} \sqrt[n]{k_{n-2}}.$$

Dann geht (6) in eine andere Form über, die wir so darstellen:

$$n\xi = A + \sum_{0,n-2}^{\nu} E_\nu K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \ldots \tau_{\nu+n-2}^{r_0}, \tag{9}$$

worin E_ν eine $n$te Einheitswurzel ist, die durch ε so ausgedrückt wird:

$$E_\nu = \varepsilon_\nu^{r_{n-2}} \varepsilon_{\nu+1}^{r_{n-3}} \ldots \varepsilon_{\nu+n-2}^{r_0}. \tag{10}$$

Nun ist nach der Definition von r_ν [§ 89, (15)]:

$$r_\nu \equiv g r_{\nu-1} \pmod{n}, \tag{11}$$

und nach (10), da der Index von r nach dem Modul $n-1$ zu nehmen ist,

$$E_{\nu-1} = \varepsilon_{\nu-1}^{r_{n-2}} \varepsilon_\nu^{r_{n-3}} \ldots \varepsilon_{\nu+n-3}^{r_0}$$

$$E_{\nu-1}^g = \varepsilon_\nu^{r_{n-2}} \varepsilon_{\nu+1}^{r_{n-3}} \ldots \varepsilon_{\nu+n-2}^{r_0},$$

also

$$E_\nu = E_{\nu-1}^g = E_0^{r_\nu}. \tag{12}$$

Sind also die ε irgendwie bestimmt, so ergibt sich:

$$E_0 = \varepsilon_0^{r_{n-2}} \varepsilon_1^{r_{n-3}} \ldots \varepsilon_{\nu+n-2}^{r_0}, \tag{13}$$

und dadurch sind nach (12) die übrigen E_ν vollkommen bestimmt. Also ergeben sich in der Tat nur n Werte aus (6), die man

z. B. dadurch erhalten kann, daß man einem der Radikale τ_ν seine verschiedenen Werte beilegt[1]).

Will man die Größen $\xi_0, \xi_1, \ldots \xi_{n-1}$ in der Reihenfolge bestimmen, die der Bildung der zyklischen und metazyklischen Gruppe zugrunde liegt, so muß man in (5) vor der Summation mit ε^{-hr_ν} multiplizieren und findet (§ 64):

$$(14) \qquad n\xi_h = A + \sum{}' \varepsilon^{-hr_\nu} K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \ldots \tau_{\nu+n-2}^{r_0},$$

worin eine veränderte Bestimmung der Radikale nur eine zyklische Vertauschung der ξ_h bedingt.

§ 91.

Befreiung von den beschränkenden Voraussetzungen.

Es wäre eine wesentliche Beschränkung dieser Untersuchungen über auflösbare Gleichungen, wenn die Voraussetzungen 1., 2. des vorigen Paragraphen aufrecht erhalten werden müßten.

Das ist aber nicht notwendig, wie wir jetzt nachweisen wollen.

Es seien jetzt $\eta_0, \eta_1, \ldots \eta_{n-1}$ die n Wurzeln irgend einer irreduziblen metazyklischen Gleichung $n$ten Grades.

Wir führen die neuen Unbekannten $\xi_0, \xi_1, \ldots \xi_{n-1}$ durch ein System von Gleichungen

$$(1) \qquad \xi_0 = \psi(\eta_0), \quad \xi_1 = \psi(\eta_1), \ldots \xi_{n-1} = \psi(\eta_{n-1})$$

ein, worin $\psi(y)$ eine ganze Funktion $(n-1)$ten Grades

$$(2) \qquad \psi(y) = a_0 + a_1 y + a_2 y^2 + \cdots + a_{n-1} y^{n-1}$$

mit unbestimmten Koeffizienten aus dem Körper $\mathfrak{R}$ bedeuten soll. Wir nehmen aber an, daß die n Größen $\xi_0, \xi_1, \ldots \xi_{n-1}$ voneinander verschieden sind; dann sind auch die ξ die Wurzeln einer irreduziblen metazyklischen Gleichung, die man durch die Substitution $\xi = \psi(\eta)$ aus der Gleichung für η ableiten kann; ξ_0 ist ein primitives Element des Körpers $\mathfrak{R}(\eta_0)$; folglich sind die Körper $\mathfrak{R}(\xi_0)$ und $\mathfrak{R}(\eta_0)$ und ebenso die konjugierten Körper $\mathfrak{R}(\xi_r)$ und $\mathfrak{R}(\eta_r)$ miteinander identisch. Es ist also auch, wenn

$$(3) \qquad \chi(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_{n-1} x^{n-1}$$

gesetzt wird, worin die Koeffizienten b gleichfalls Größen in $\mathfrak{R}$ sind,

$$(4) \qquad \eta_0 = \chi(\xi_0), \quad \eta_1 = \chi(\xi_1), \ldots \eta_{n-1} = \chi(\xi_{n-1}).$$

[1]) Man vergleiche hiermit die Cayleysche Auflösung der kubischen Gleichung S. 130.

Es ist nun leicht einzusehen, daß man über die Koeffizienten $a_0, a_1 \dots a_{n-1}$ in (2) so verfügen kann, daß die Voraussetzungen 1. und 2. des vorigen Paragraphen für die ξ erfüllt sind.

Denn betrachten wir die a als unabhängige Variable, so werden die durch (1) bestimmten ξ gleichfalls Variable, die mit den a durch eine lineare Substitution mit nicht verschwindender Determinante zusammenhängen. Die Determinante dieser Substitution ist nämlich:

$$\begin{vmatrix} 1, & \eta_0, & \eta_0^2 & \dots \eta_0^{n-1} \\ 1, & \eta_1, & \eta_1^2 & \dots \eta_1^{n-1} \\ \dots & \dots & \dots & \dots \\ 1, & \eta_{n-1}, & \eta_{n-1}^2 & \dots \eta_{n-1}^{n-1} \end{vmatrix},$$

d. h. gleich dem Differenzprodukt der η, das nach der Voraussetzung von Null verschieden ist.

Die Funktionen (ε, x) und die Differenzen $f_\alpha - f_\beta$, die nach § 89 nicht verschwindende Funktionen der Variablen $x_0, x_1, \dots x_{n-1}$ sind, gehen also, wenn für die x die Substitution $x_h = \psi(\eta_h)$ gemacht wird, in Funktionen der Variablen a über, die auch nicht identisch verschwinden können. Nach dem Satze § 14, kann man also für die Größen $a_0, a_1, \dots a_{n-1}$ solche rationale Zahlen setzen, daß die Funktionen (ε, x) und die Differenzen $f_\alpha - f_\beta$, die dann also in (ε, ξ) und in $k_\alpha - k_\beta$ übergehen, von Null verschieden werden, und daß auch die ξ_h voneinander verschieden bleiben, was zu beweisen war.

Nachdem also dies festgestellt ist, führen wir das System der Variablen $x_0, x_1, \dots x_{n-1}$ und ein davon abhängiges System von Variablen y

$$y_0 = \chi(x_0), \quad y_1 = \chi(x_1), \dots y_{n-1} = \chi(x_{n-1}) \tag{5}$$

ein, worin χ die Funktion (3) bedeutet, und setzen

$$\frac{(\varepsilon^{r_\nu}, y)}{(\varepsilon^{r_\nu}, x)} = \Theta_\nu. \tag{6}$$

Wird auf die Indices der x irgend eine Permutation angewandt, so erleiden die Indices der y die gleiche Permutation. Wenn wir also auf (6) die zyklische Permutation s anwenden so ändert sich Θ_ν nicht (nach § 89, 4.).

Wenn wir also Θ_ν in der Normalform des § 88 darstellen, so sind seine Koeffizienten zyklische Funktionen von $x_0, x_1, \dots x_{n-1}$.

Wendet man auf Θ_ν die Permutation $t^{-1} = (h, g^{-1}h)$ an, so geht Θ_ν nach § 89, 4. in $\Theta_{\nu+1}$ über, und die Θ_ν erleiden also eine zyklische Permutation.

Denselben Erfolg hat aber auch die Substitution $\sigma = (\varepsilon, \varepsilon^g)$, und also hat das System der Funktionen $\Theta_0, \Theta_1, \ldots \Theta_{n-2}$ die Eigenschaften $\alpha)$ $\beta)$ $\gamma)$ $\delta)$ (§ 89).

Nach dem am Ende des § 89 bewiesenen Satze geht also, wenn wir die x durch die ξ und folglich die y durch die η ersetzen, Θ_ν in eine rationale Funktion von k_ν

$$Q_\nu = Q(k_\nu) \tag{7}$$

über, und aus (6) ergibt sich:

$$(\varepsilon^{r_\nu}, \eta) = Q_\nu(\varepsilon^{r_\nu}, \xi). \tag{8}$$

Wenn man hierin für $(\varepsilon^{r_\nu}, \xi)$ den Ausdruck (5), § 90 substituiert, so erhält man für $(\varepsilon^{r_\nu}, \eta)$ einen Ausdruck von ganz derselben Form, nur mit dem Unterschiede, daß an Stelle von K_ν getreten ist $Q_\nu K_\nu$. Die Funktionen $Q_\nu K_\nu$ haben im wesentlichen dieselben Eigenschaften, wie die Funktion K_ν, nur daß sie auch zum Teil Null sein können. Durch dieselbe Veränderung ergibt dann (6) den Wert von η_0.

Damit sind also die beschränkenden Voraussetzungen des § 90 beseitigt, und wir haben das allgemeine Theorem:

I. Jede Wurzel ξ einer metazyklischen Gleichung vom Primzahlgrad n kann in der Form dargestellt werden:

$$\xi = A + \sum_{0,n-2}^{\nu} K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \ldots \tau_{\nu+n-2}^{r_0}, \tag{9}$$

worin A eine rationale Größe, $k_0, k_1, \ldots k_{n-2}$ die von einander und von Null verschiedenen Wurzeln einer zyklischen Gleichung $(n-1)$ten Grades, K_ν eine rationale Funktion von k_ν ist, deren Form für alle ν dieselbe ist. Die Exponenten $r_0, r_1, \ldots r_{n-2}$ sind die kleinsten positiven Reste der Zahlen $1, g, g^2, \ldots g^{n-2}$, wenn g eine primitive Wurzel von n ist. Die n Werte, die man aus (9) erhält, wenn man den Radikalen $\tau_\nu = \sqrt[n]{k_\nu}$ ihre verschiedenen Werte beilegt, sind die n Wurzeln einer und derselben rationalen Gleichung.

Die Formel (9) ergibt sich aus § 90 (6), wenn A und K_ν durch nA und nK_ν ersetzt wird, was zur Vereinfachung geschehen ist.

Das Theorem I läßt sich nun aber auch umkehren.

Um das nachzuweisen, bezeichnen wir mit ε irgend eine imaginäre $n$te Einheitswurzel und setzen:

$$\xi_h = A + \sum_{0,n-2}^{\nu} \varepsilon^{hr_\nu} K_\nu \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \dots \tau_{\nu+n-2}^{r_0} \tag{10}$$

$$h = 0, 1, 2 \dots n-1,$$

oder abgekürzt:

$$\xi_h = A + \sum_{0,n-2}^{\nu} \varepsilon^{hr_\nu} K_\nu \sqrt[n]{R_\nu}, \tag{11}$$

worin wie früher:

$$\sqrt[n]{R_\nu} = \tau_\nu^{r_{n-2}} \tau_{\nu+1}^{r_{n-3}} \dots \tau_{\nu+n-2}^{r_0} \tag{12}$$

gesetzt ist. Wir haben nun die Änderungen zu untersuchen, die sich für ξ_h ergeben, wenn wir

1. eines der Radikale $\tau_\nu = \sqrt[n]{k_\nu}$ anders bestimmen und
2. die $k_0, k_1, \dots k_{n-2}$ zyklisch vertauschen.

Wir wollen einem der Radikale, etwa dem τ_α, ein anderes Vorzeichen geben, also die Vertauschung

$$(\tau_\alpha, \varepsilon^\beta \tau_\alpha) \tag{13}$$

machen, wo β ein beliebiger Exponent sein kann, und α einer der Indices $0, 1 \dots n-2$ ist.

In (12) hat, da der Index von r nach dem Modul $n-1$ zu nehmen ist, das Radikal τ_α den Exponenten $r_{n+\nu-\alpha-2} = r_{\nu-\alpha-1}$, und also entspricht der Vertauschung (13) die Vertauschung

$$\left(\sqrt[n]{R_\nu}, \varepsilon^{\beta r_{\nu-\alpha-1}} \sqrt[n]{R_\nu}\right). \tag{14}$$

Nun ist nach der Bedeutung der Zahlen r allgemein $r_{\alpha+\beta} \equiv r_\alpha r_\beta \pmod{n}$, und wenn man also die Vertauschung (13) in (11) einführt, so geht h in $h + \beta r_{-\alpha-1}$ über.

1. Es ruft also die Vertauschung (13) unter den Indices von ξ die zyklische Permutation

$$(h, h + \beta r_{-\alpha-1})$$

hervor.

Machen wir zweitens die zyklische Permutation

$$(\tau_0, \tau_1, \dots \tau_{n-2}), \tag{15}$$

so entspricht diese den Vertauschungen

(16) $\left(\sqrt[n]{R_\nu}, \sqrt[n]{R_{\nu+1}}\right), \quad (K_\nu, K_{\nu+1}), \quad \nu = 0, 1, 2 \ldots n-2,$

und ξ_h geht über in

$$A + \sum^{\nu} \varepsilon^{hr_\nu} K_{\nu+1} \sqrt[n]{R_{\nu+1}} = A + \sum \varepsilon^{hr_\nu - 1} K_\nu \sqrt[n]{R_\nu}$$
$$= A + \sum \varepsilon^{hg^{-1}r_\nu} \sqrt[n]{R_\nu}.$$

2. Demnach erleiden die Indices von ξ durch die Permutation (15) die Permutation

$$(h, g^{-1}h).$$

Betrachten wir nun irgend eine rationale symmetrische (oder auch nur metazyklische) Funktion der ξ:

$$S(\xi_0, \xi_1, \ldots \xi_{n-1}),$$

so erhalten wir, wenn wir die Werte (10) einführen, daraus eine rationale Funktion der Radikale

$$\tau_0, \tau_1, \ldots \tau_{n-2}.$$

Diese Funktion ändert sich aber nicht, wenn man einem dieser Radikale einen anderen seiner n Werte gibt, und folglich muß die Funktion rational von $k_0, k_1, \ldots k_{n-2}$ abhängen. Wegen 2. ändert sich diese Funktion aber auch nicht, wenn unter den k_ν die zyklische Permutation $(k_0, k_1, \ldots, k_{n-2})$ vorgenommen wird, und weil nun die Größen k_ν die Wurzeln einer zyklischen Gleichung im Körper $\mathfrak{R}$ sind, so ist die Funktion $S(\xi_0, \xi_1 \ldots \xi_{n-1})$ eine Größe in $\mathfrak{R}$, d. h. rational. Wendet man dies auf die Koeffizienten der Gleichung an, deren Wurzeln die Größen (10) sind, so folgt, daß diese Größen die Wurzeln einer Gleichung $n$ten Grades in $\mathfrak{R}$ sind.

Was die Irreduzibilität dieser Gleichung betrifft, so ist darüber folgendes zu bemerken:

Wenn das Radikal $\sqrt[n]{R_0}$ nicht rational durch k_0 ausdrückbar ist, so ist nach § 84 die Funktion

(17) $$x^n - R_0$$

im Körper $\mathfrak{R}(k_0)$, der aus $\mathfrak{R}$ durch Adjunktion von k_0 entsteht, irreduzibel. Nach § 90 (8) kann man jede der Wurzeln ξ, etwa ξ_0, rational durch eine der Wurzeln dieser Funktion, $\sqrt[n]{R_0}$, ausdrücken, und die übrigen Wurzeln ξ erhält man, wenn man für $\sqrt[n]{R_0}$ die verschiedenen Wurzeln von (17) setzt. Wenn also $\Phi(\xi_0) = 0$ eine in $\mathfrak{R}(k_0)$ rationale Gleichung ist, der eine der Wurzeln ξ

genügt, so folgt aus der Irreduzibilität von (17), daß dieser Gleichung auch alle anderen ξ genügen müssen, und daß folglich die Gleichung $n$ten Grades, deren Wurzeln die ξ sind, im Körper $\mathfrak{K}(k_0)$, und um so mehr also im Körper $\mathfrak{K}$ irreduzibel ist.

Wenn aber $\sqrt[n]{R_0}$ in $\mathfrak{K}(k_0)$ enthalten ist, so sind auch die sämtlichen $\sqrt[n]{R_\nu}$ in diesem Körper enthalten, wie aus § 90, (8) hervorgeht. Es sind also zunächst alle ξ in dem Körper $\mathfrak{K}(k_0, \varepsilon)$ enthalten.

Durch die Substitution (k_0, k_1) geht $R_{\nu,-1}$ in R_ν über, und folglich $\sqrt[n]{R_{\nu,-1}}$ in $\varepsilon_\nu \sqrt[n]{R_\nu}$, wenn ε_ν eine $n$te Einheitswurzel ist. Es ergibt sich aber, wenn man auf die Gleichung § 90 (8) die Substitution (k_0, k_1) anwendet,

$$\varepsilon_\nu = \varepsilon_{\nu-1}^{g} = \varepsilon_0^{g^\nu}. \tag{18}$$

Hier muß nun ε_0 dem Körper $\mathfrak{K}(k_0)$ angehören, und die

$$S_\nu = \varepsilon_0^{g^\nu} \sqrt[n]{R_\nu}$$

gehören gleichfalls dem Körper $\mathfrak{K}(k_0)$ an; zugleich geht S_ν durch die Substitution (k_0, k_1) in $S_{\nu+1}$ über.

Die Summe

$$A + K_0 S_0 + K_1 S_1 + \cdots K_{\nu-2} S_{\nu-2}$$

ist daher eine Größe in dem ursprünglichen Rationalitätsbereich $\mathfrak{K}$, und nach (11) ist diese gleich einer der Größen ξ.

Wir können also das Theorem I so umkehren:

II. Jede in der Form (9) enthaltene Größe ξ ist die Wurzel einer Gleichung $n$ten Grades in $\mathfrak{K}$, und diese Gleichung ist irreduzibel, wenn nicht eine der n konjugierten Größen ξ selbst in $\mathfrak{K}$ enthalten ist.

§ 92.

Realitätsverhältnisse.

Wenn der Körper $\mathfrak{K}$ ein reeller ist, so gibt es, wie wir im § 83 gesehen haben, zwei Arten von metazyklischen Gleichungen von Primzahlgrad n, solche mit einer reellen und $n - 1$ imaginären Wurzeln und solche mit lauter reellen Wurzeln. Ebenso haben wir im § 63 gesehen, daß es zwei Arten von zyklischen

Gleichungen eines geraden Grades gibt, nämlich solche mit lauter reellen und solche mit lauter imaginären Wurzeln.

Wenn nun die zyklische Gleichung, deren Wurzeln die Größen $k_0, k_1, \ldots k_{n-2}$ sind, zur ersten Art gehört, wenn also $k_0, k_1, \ldots k_{n-2}$ reell sind, und die Radikale $\tau_0, \tau_1 \ldots \tau_{n-2}$ auch reell genommen werden, so zeigt die Formel (10) § 91, daß ξ_0 reell, ξ_h und ξ_{-h} konjugiert imaginär sind.

Wenn andererseits die k imaginär sind, so ist nach S. 294 k_ν und $k_{\nu+\frac{n-1}{2}}$ konjugiert imaginär, und nach der Formel (8) in § 90 können wir setzen, wenn Φ eine rationale Funktion bedeutet,

$$\sqrt[n]{R_{\nu+\frac{n-1}{2}}} = \Phi(k_\nu)\left(\sqrt[n]{R_\nu}\right)^{g^{\frac{n-1}{2}}},$$

oder auch

$$\sqrt[n]{R_{\nu+\frac{n-1}{2}}}\sqrt[n]{R_\nu} = \Phi(k_\nu)\left(\sqrt[n]{R_\nu}\right)^{g^{\frac{n-1}{2}}+1}.$$

Weil nun g eine primitive Wurzel von n ist, so ist $g^{\frac{n-1}{2}}+1$ durch n teilbar, und die rechte Seite wird daher rational. Bedeutet also Ψ eine rationale Funktion, so ist

$$\sqrt[n]{R_{\nu+\frac{n-1}{2}}}\sqrt[n]{R_\nu} = \Psi(k_\nu),$$

und diese Formel zeigt, daß

$$\Psi(k_\nu) = \Psi\left(k_{\nu+\frac{n-1}{2}}\right),$$

also daß $\Psi(k_\nu)$ reell ist. Daraus folgt, daß

$$\sqrt[n]{R_{\nu+\frac{n-1}{2}}} \quad \text{und} \quad \sqrt[n]{R_\nu}$$

konjugiert imaginär sind; denn erstens sind ihre $n$ten Potenzen konjugiert imaginär, sie selbst also zunächst bis auf eine $n$te Einheitswurzel als Faktor. Da aber ihr Produkt reell ist, so muß diese $n$te Einheitswurzel $= 1$ sein. Da nun außerdem

$$r_\nu \equiv -r_{\nu+\frac{n-1}{2}} \pmod{n}$$

ist, so zeigt die Formel (11), daß in diesem Falle die Wurzeln ξ_ν alle reell sind.

Damit ist also der Satz für einen reellen Körper $\mathfrak{R}$ bewiesen:

Hat die zyklische Gleichung, deren Wurzeln die $k_0, k_1 \ldots k_{n-2}$ sind, reelle Wurzeln, so ist von den Wurzeln ξ eine reell, die übrigen imaginär; hat diese zyklische Gleichung imaginäre Wurzeln, so sind alle ξ reell.

§ 93.

Metazyklische Gleichungen fünften Grades.

Durch die Sätze der vorangehenden Paragraphen sind die Wurzeln einer metazyklischen Gleichung von Primzahlgrad in eine allgemein gültige Form gebracht, die es gestattet, alle diese Größen wirklich zu bilden, wenn man die Kenntnis der Wurzeln zyklischer Gleichungen voraussetzt. Ganze Systeme zyklischer Gleichungen jeden Grades liefert uns z. B. die Kreisteilungstheorie, deren Wurzeln die Kreisteilungsperioden sind. So können wir also beispielsweise für den Körper der rationalen Zahlen beliebig viele metazyklische Gleichungen bilden. Ein Satz von Kronecker, nach dem die Wurzeln nicht nur aller zyklischen, sondern aller Abelschen Gleichungen im absoluten Rationalitätsbereich durch Kreisteilungszahlen darstellbar sind, lehrt uns, daß man auf diesem Wege alle metazyklischen Gleichungen erhalten kann. Hier wollen wir als Anwendung und Veranschaulichung des Vorhergehenden noch die spezielle Aufgabe behandeln, in einem beliebigen Körper $\mathfrak{R}$ die Wurzeln aller metazyklischen Gleichungen fünften Grades zu finden[1]).

Dazu ist erforderlich und hinreichend, daß wir die Wurzeln k_0, k_1, k_2, k_3 einer zyklischen Gleichung vierten Grades allgemein bestimmen, und zwar unter der Voraussetzung, daß k_0, k_1, k_2, k_3 untereinander verschieden und keine von ihnen Null sei.

Als Grundlage dient uns dabei die Funktion

$$w = (k_0 - k_2)(k_1 - k_3), \tag{1}$$

die von Null verschieden sein muß, und die für einen reellen Körper $\mathfrak{R}$ immer reell ist, da entweder k_0, k_1, k_2, k_3 reell sind oder k_0, k_2 und k_1, k_3 zwei konjugiert imaginäre Paare bilden.

Bei der zyklischen Permutation (0, 1, 2, 3) ändert w sein Vorzeichen. Also ist das Quadrat von w eine zyklische Funktion,

---

[1]) Abel, Oeuvres 1881, Vol. II, p. 66; Cayley, Coll. papers, Vol. I, p. 132.

die nach Voraussetzung bekannt sein soll. Wir setzen daher, indem wir durch die kleinen lateinischen Buchstaben a, b, c ... Größen in $\mathfrak{R}$, also rationale Größen bezeichnen,

$$w = 4\sqrt{c}, \tag{2}$$

und bemerken noch, daß jede Funktion der k, die durch die zyklische Permutation (0, 1, 2, 3) ihr Vorzeichen ändert, das Produkt einer rationalen Größe mit $\sqrt{c}$ ist.

Es handelt sich dann nur darum, zyklische Funktionen in genügender Anzahl zu bilden, so daß man die vier Größen k_0, k_1, k_2, k_3 daraus berechnen und durch voneinander unabhängige rationale Größen algebraisch ausdrücken kann.

Nun sind zwei weitere zyklische Funktionen

$$(k_0 - k_2)^2 + (k_1 - k_3)^2, \quad \frac{(k_0 - k_2)^2 - (k_1 - k_3)^2}{(k_0 - k_2)(k_1 - k_3)},$$

und wir bekommen also, wenn a, b rationale Größen sind,

$$\begin{aligned} (k_0 - k_2)^2 + (k_1 - k_3)^2 &= 8b \\ (k_0 - k_2)^2 - (k_1 - k_3)^2 &= 8a\sqrt{c}. \end{aligned} \tag{3}$$

Die drei rationalen Größen a, b, c sind aber nicht voneinander unabhängig, sondern es besteht nach (2) und (3) zwischen ihnen die Relation

$$b^2 = c(1 + a^2). \tag{4}$$

Aus (3) ergibt sich ferner:

$$\begin{aligned} k_0 - k_2 &= 2\sqrt{b + a\sqrt{c}} \\ k_1 - k_3 &= 2\sqrt{b - a\sqrt{c}}, \end{aligned} \tag{5}$$

und zwischen den beiden Quadratwurzeln besteht die Relation

$$\sqrt{c} = \sqrt{b + a\sqrt{c}}\,\sqrt{b - a\sqrt{c}}. \tag{6}$$

Bezeichnen wir ferner mit C und B wieder zwei rationale Größen, so ist

$$\begin{aligned} k_0 + k_1 + k_2 + k_3 &= 4C, \\ k_0 - k_1 + k_2 - k_3 &= 4B\sqrt{c}, \end{aligned} \tag{7}$$

also

$$\begin{aligned} k_0 + k_2 &= 2(C + B\sqrt{c}), \\ k_1 + k_3 &= 2(C - B\sqrt{c}), \end{aligned} \tag{8}$$

wodurch nach (5) folgt:

$$
\begin{aligned}
k_0 &= C + B\sqrt{c} + \sqrt{b + a\sqrt{c}}\\
k_1 &= C - B\sqrt{c} + \sqrt{b - a\sqrt{c}}\\
k_2 &= C + B\sqrt{c} - \sqrt{b - a\sqrt{c}}\\
k_3 &= C - B\sqrt{c} - \sqrt{b - a\sqrt{c}}.
\end{aligned}
\tag{9}
$$

Es ist auch umgekehrt leicht, zu zeigen, daß diese Größen die Wurzeln einer biquadratischen zyklischen Gleichung sind. Denn setzen wir zur Abkürzung

$$r = \sqrt{c}, \quad \varrho = \sqrt{b + a\sqrt{c}}, \quad \varrho' = \sqrt{b - a\sqrt{c}}, \tag{10}$$

also nach (6):

$$r = \varrho\varrho', \quad \varrho^2 = b + ar, \quad \varrho'^2 = b - ar,$$

so können wir jede rationale Funktion der k_0, k_1, k_2, k_3 als lineare Funktion mit rationalen Koeffizienten von den sechs Radikalen

$$1, \quad r, \quad \varrho, \quad \varrho', \quad r\varrho, \quad r\varrho'$$

darstellen. Die zyklische Permutation (k_0, k_1, k_2, k_3) entspricht der Vertauschung

$$\begin{array}{cccc} r, & \varrho, & \varrho', \\ -r, & \varrho', & -\varrho, \end{array}$$

und wenn man diese Vertauschung wiederholt, so sieht man, daß eine zyklische Funktion der k sich nicht ändern kann, wenn in der linearen Darstellung durch diese sechs Größen folgende Vertauschungen gemacht werden:

$$\begin{array}{cccccc} 1, & r, & \varrho, & \varrho' & r\varrho, & r\varrho' \\ 1, & -r, & \varrho', & -\varrho, & -r\varrho', & r\varrho \\ 1, & r, & -\varrho, & -\varrho', & -r\varrho, & -r\varrho' \\ 1, & -r, & -\varrho', & \varrho, & r\varrho', & -r\varrho, \end{array}$$

und wenn man die vier sich so ergebenden Ausdrücke addiert, so erhält man für die zyklische Funktion der k einen rationalen Ausdruck.

Will man die biquadratische Gleichung bilden, deren Wurzeln die k sind, so führt man am besten $k - C = x$ als Unbekannte ein. Man bekommt so durch einfache Rechnung die Gleichung

$$\begin{aligned} x^4 - 2(B^2c + b)x^2 - 4Bacx + B^4c^2 - 2B^2bc \\ + b^2 - a^2c = 0. \end{aligned} \tag{11}$$

Die Darstellung der k durch die Formeln (9) ist aber noch nicht vollständig befriedigend, weil zwischen den darin vorkommenden rationalen Größen a, b, c noch die Relation (4)

stattfindet, und wir suchen eine Darstellung durch unabhängige Größen.

Einen besonderen Fall müssen wir zunächst abmachen, nämlich $b = 0$, was $a = i$ zur Folge hat. Dann geben die Formeln (9):

$$k_0 = C + B\sqrt{c} + \frac{1+i}{\sqrt{2}}\sqrt[4]{c}$$

$$k_1 = C - B\sqrt{c} + \frac{1-i}{\sqrt{2}}\sqrt[4]{c}$$

$$k_2 = C + B\sqrt{c} - \frac{1+i}{\sqrt{2}}\sqrt[4]{c}$$

$$k_3 = C - B\sqrt{c} - \frac{1-i}{\sqrt{2}}\sqrt[4]{c},$$

und die biquadratische Gleichung (11) wird

$$x^4 - 2B^2cx^2 - 4iBcx + B^4c^2 + c = 0.$$

Dieser Fall gehört aber nur dann hierher, wenn i im Körper $\mathfrak{R}$ enthalten ist, also niemals bei reellen Körpern.

Wenn aber b nicht verschwindet, so führen wir eine neue rationale Größe h ein, indem wir

$$b = h(1 + a^2), \quad c = h^2(1 + a^2)$$

setzen, wodurch dann die Relation (4) identisch befriedigt ist, und es geben die Formeln (9), wenn man Bh durch B ersetzt,

$$\begin{aligned} k_0 &= C + B\sqrt{1+a^2} + \sqrt{h(1 + a^2 + a\sqrt{1+a^2})} \\ k_1 &= C - B\sqrt{1+a^2} + \sqrt{h(1 + a^2 - a\sqrt{1+a^2})} \\ k_2 &= C + B\sqrt{1+a^2} - \sqrt{h(1 + a^2 + a\sqrt{1+a^2})} \\ k_3 &= C - B\sqrt{1+a^2} - \sqrt{h(1 + a^2 - a\sqrt{1+a^2})}. \end{aligned} \tag{12}$$

Dieser Ausdruck für k_0 hat noch den Vorzug, daß er, wie man auch die Vorzeichen der darin vorkommenden Quadratwurzeln bestimmen mag, nur vier verschiedene Werte darstellt.

Bei Abel findet sich für k_0 ein etwas anderer Ausdruck, nämlich:

$$k_0 = C + B\sqrt{1+e^2} + \sqrt{h(1 + e^2 + \sqrt{1+e^2})},$$

der aus (12) hervorgeht, wenn man $a = 1:e$ setzt und dann B und h durch Be und he^2 ersetzt. Der Ausdruck (12) ist also insofern allgemeiner, als er auch den besonderen Fall $a = 0$

umfaßt, in dem die biquadratische Gleichung in zwei quadratische Gleichungen zerfällt.

Um nun die Wurzel einer metazyklischen Gleichung fünften Grades darzustellen, sind diese Ausdrücke für k_0, k_1, k_2, k_3 in die Formel (9) des § 91 zu substituieren. Nehmen wir $g = 2$ an, so werden die Exponenten r_0, r_1, r_2, r_3 der Reihe nach kongruent mit 1, 2, 4, 8, also gleich 1, 2, 4, 3, und es ergibt sich, wenn, wie in § 90,

$$\tau_0 = \sqrt[5]{k_0}, \quad \tau_1 = \sqrt[5]{k_1}$$
$$\tau_2 = \sqrt[5]{k_2}, \quad \tau_3 = \sqrt[5]{k_3}$$

gesetzt wird,

$$\xi = A + K_0 \tau_0^3 \tau_1^4 \tau_2^2 \tau_3 + K_1 \tau_1^3 \tau_2^4 \tau_3^2 \tau_0 + K_2 \tau_2^3 \tau_3^4 \tau_0^2 \tau_1 + K_3 \tau_3^3 \tau_0^4 \tau_1^2 \tau_2. \tag{13}$$

Die Koeffizienten in diesem Ausdruck, K_0, K_1, K_2, K_3 sind rationale Funktionen der k_0, k_1, k_2, k_3, die durch zyklische Permutation der k selbst zyklisch permutiert werden. Nach Abel ist

$$K_0 = A_1 + A_2 k_0 + A_3 k_2 + A_4 k_0 k_2$$

zu setzen, worin A_1, A_2, A_3, A_4 rational sind, und dieselben Ausdrücke legt auch Cayley zugrunde. Diese Annahme ist aber nicht allgemein genug, weil zwischen $k_0 k_2$, k_0 und k_2 eine aus (12) leicht abzuleitende lineare Relation besteht. Am einfachsten drückt man die k durch die drei Radikale

$$r = \sqrt{1 + a^2}, \quad \varrho = \sqrt{h(1 + a^2 + a\sqrt{1 + a^2})}$$
$$\varrho' = \sqrt{h(1 + a^2 - a\sqrt{1 + a^2})}, \quad \varrho\varrho' = hr$$

aus.

Man kann das Radikal ϱ' linear ausdrücken durch $r\varrho$ und ϱ, wie man aus der Formel

$$hr^2\varrho' = r\varrho\varrho'^2$$

ersieht, wenn man rechts für ϱ'^2 seinen Ausdruck durch r einsetzt und bedenkt, daß r^2 rational ist. So erhält man, wenn man statt der zyklischen Permutation der k die Vertauschung

$$\begin{pmatrix} r, & \varrho, & \varrho' \\ -r, & \varrho', & -\varrho \end{pmatrix}$$

anwendet, und mit A_1, A_2, A_3, A_4 rationale Größen bezeichnet:

$$(14)\qquad \begin{aligned} K_0 &= A_1 + A_2 r + A_3 \varrho + A_4 r \varrho \\ K_1 &= A_1 - A_2 r + A_3 \varrho' - A_4 r \varrho' \\ K_2 &= A_1 + A_2 r - A_3 \varrho - A_4 r \varrho \\ K_3 &= A_1 - A_2 r - A_3 \varrho' + A_4 r \varrho'. \end{aligned}$$

Diese Größen K_0, K_1, K_2, K_3 sind selbst wieder die Wurzeln einer zyklischen biquadratischen Gleichung. Sie haben nur scheinbar eine allgemeinere Form, als die k; denn setzt man K_0 in die Form

$$K_0 = A_1 + A_2 r + \sqrt{\varrho^2 (A_3 + A_4 r)^2},$$

so erkennt man die Form von k_0 in (9) wieder, natürlich mit veränderten a, b, c.

Vierzehnter Abschnitt.

Zahlen und Funktionale eines algebraischen Körpers.

§ 94.

Ganze algebraische Zahlen.

Eine algebraische Gleichung

$$(1) \quad F(x) = x^m + A_1 x^{m-1} + \cdots + A_{m-1} x + A_m = 0,$$

deren Koeffizienten $A_1, A_2, \ldots A_m$ rationale Zahlen sind, nennen wir eine rationale Gleichung. Sie hat, wie wir früher nachgewiesen haben, immer m oder weniger, aber immer wenigstens eine Wurzel.

In den folgenden Betrachtungen soll es sich nicht um die numerischen Werte der Wurzeln handeln, sondern um die arithmetischen Gesetze, denen diese Zahlen unterworfen sind, die sich aus der Definition selbst und nicht aus den numerischen Werten ableiten lassen. Wir stellen also jetzt folgende Definition an die Spitze:

1. Eine Zahl Θ, die einer rationalen Gleichung

$$F(\Theta) = 0$$

genügt, heißt eine algebraische Zahl.

Jede algebraische Gleichung mit rationalen Koeffizienten liefert uns solche algebraische Zahlen, die sich also in beliebiger Menge angeben lassen.

Eine algebraische Zahl genügt nicht nur einer, sondern unendlich vielen rationalen Gleichungen; denn multipliziert man zwei beliebige Funktionen von der Form $F(x)$ miteinander, so erhält man eine Funktion derselben Form, die für $x = \Theta$ verschwindet, wenn einer der Faktoren diese Eigenschaft hat.

Unter allen rationalen Gleichungen, denen eine algebraische Zahl genügt, ist eine von möglichst niedrigem Grade, $f(\Theta) = 0$, worin $f(x)$ die Form hat:

$$f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n, \tag{2}$$

und es kann auch nur eine solche Gleichung geben, wenn wir, wie bisher immer, den Koeffizienten der höchsten Potenz von x gleich 1 annehmen.

Denn sind $f(x)$, $f_1(x)$ zwei Funktionen von der Form (2) von gleichem Grade n, so ist $f(x) - f_1(x)$ von niedrigerem als dem $n$ten Grade, und wenn sowohl $f(\Theta)$ als $f_1(\Theta)$ verschwindet, so verschwindet auch $f(\Theta) - f_1(\Theta)$; wenn also diese Differenz nicht identisch verschwindet, so genügt Θ einer Gleichung von niedrigerem als dem $n$ten Grade, was gegen die Voraussetzung ist.

2. Die Funktion $f(x)$ ist im Körper der rationalen Zahlen irreduzibel.

Denn zerfällt $f(x)$ in zwei rationale Faktoren $f_1(x)$ und $f_2(x)$, von denen jeder von niedrigerem Grade ist als $f(x)$, so genügt Θ einer der beiden Gleichungen $f_1(\Theta) = 0$, $f_2(\Theta) = 0$, was unserer Voraussetzung widerspricht.

3. Ist n der Grad der rationalen Gleichung niedrigsten Grades, der die Zahl Θ genügt, so nennen wir Θ eine algebraische Zahl $n$ten Grades[1]).

4. Eine algebraische Zahl Θ wird eine **ganze** algebraische Zahl genannt, wenn sie einer rationalen Gleichung

$$\Theta^m + A_1 \Theta^{m-1} + \cdots + A_{m-1}\Theta + A_m = 0$$

genügt, deren Koeffizienten $A_1, A_2, \ldots A_m$ **ganze** rationale Zahlen sind.

Wir bemerken, daß es nach dieser Definition ausreicht, um eine algebraische Zahl Θ als ganz zu charakterisieren, wenn unter den unendlich vielen Gleichungen der Form (1), denen Θ genügt, eine ist, deren Koeffizienten ganze Zahlen sind.

Die ganzen algebraischen Zahlen umfassen als speziellen Fall die gewöhnlichen ganzen Zahlen, die wir zur Unterscheidung

[1]) Eine algebraische Zahl kann hiernach durch ein endliches System gewöhnlicher ganzer Zahlen vollständig definiert werden und ist nichts anderes als ein Name für dieses System in Verbindung mit bestimmten Rechenvorschriften, die aus der Buchstabenrechnung geläufig sind.

ganze rationale Zahlen nennen. Die positiven ganzen rationalen Zahlen nennen wir auch, einem verbreiteten Sprachgebrauche folgend, natürliche Zahlen.

Unter ganzen Zahlen schlechtweg verstehen wir dann ganze algebraische, rationale und irrationale Zahlen.

5. Eine ganze algebraische Zahl, die zugleich rational ist, ist notwendig eine ganze rationale Zahl.

Nehmen wir nämlich an, es sei $\Theta = P:Q$ ein rationaler Bruch, und P, Q ganze rationale Zahlen ohne gemeinsamen Teiler, etwa Q positiv, so ergibt sich aus (1), wenn $x = \Theta$ ist:

$$P^m + A_1 P^{m-1} Q + A_2 P^{m-2} Q^2 + \cdots + A_m Q^m = 0,$$

und daraus ist zu ersehen, daß jeder Primteiler von Q in P enthalten sein müßte. Es muß also $Q = 1$ sein, und $\Theta = P$ ist eine ganze rationale Zahl.

6. Summe, Differenz und Produkt zweier ganzer Zahlen sind wieder ganze Zahlen.

Um diesen Hauptsatz zu beweisen, nehmen wir an, es seien α, β zwei ganze Zahlen, die den Gleichungen

$$\begin{aligned} \alpha^\mu + a_1 \alpha^{\mu-1} + \cdots + a_{\mu-1}\alpha + a_\mu &= 0, \\ \beta^\nu + b_1 \beta^{\nu-1} + \cdots + b_{\nu-1}\beta + b_\nu &= 0 \end{aligned} \tag{3}$$

genügen und machen eine der drei Annahmen

$$\omega = \alpha + \beta, \quad \alpha - \beta, \quad \alpha\beta.$$

Dann setzen wir $\mu\nu = m$ und bezeichnen die m Größen

$$\alpha^r \beta^s, \quad \begin{matrix} r = 0, 1 \ldots \mu - 1, \\ s = 0, 1 \ldots \nu - 1 \end{matrix}$$

in irgend einer Reihenfolge mit $\omega_1, \omega_2, \ldots \omega_m$.

Dann können die Produkte $\omega\omega_1, \omega\omega_2, \ldots \omega\omega_m$ mit Hilfe der Gleichungen (3) in die Form gesetzt werden:

$$\omega\omega_r = c_{r,1}\omega_1 + c_{r,2}\omega_2 + \cdots + c_{r,m}\omega_m, \qquad r = 1, 2, \ldots m,$$

worin die Koeffizienten $c_{s,r}$ ganze rationale Zahlen sind. Wenn man aus diesen Gleichungen aber die $\omega_1, \omega_2 \ldots \omega_m$ eliminiert, so folgt:

$$\begin{vmatrix} c_{1,1} - \omega, & c_{1,2} & \ldots & c_{1,m} \\ c_{2,1}, & c_{2,2} - \omega & \ldots & c_{2,m} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ c_{m,1}, & c_{m,2} & \ldots & c_{m,m} - \omega \end{vmatrix} = 0,$$

was entwickelt die Form erhält:

$$\omega^m + C_1\omega^{m-1} + \cdots + C_m = 0,$$

worin die $C_1, C_2, \ldots C_m$ gleichfalls ganze rationale Zahlen sind. Dies aber zeigt, daß ω eine ganze Zahl ist, wie bewiesen werden sollte.

7. Ist $f(x)$ eine im Körper der rationalen Zahlen irreduzible Funktion, und ist eine Wurzel α von $f(x) = 0$ eine ganze Zahl, so sind alle Wurzeln von $f(x)$ ganze Zahlen.

Denn wenn eine rationale Funktion $F(x)$ für $x = \alpha$ verschwindet, so ist $F(x)$ durch $f(x)$ teilbar, und alle Wurzeln von $f(x)$ sind zugleich Wurzeln von $F(x)$ (§ 53). Wenn nun α eine ganze Zahl ist, so gibt es eine Funktion

$$F(x) = x^m + A_1 x^{m-1} + \cdots + A_m$$

mit ganzzahligen Koeffizienten $A_1 \ldots A_m$, die für $x = \alpha$ verschwindet, und $F(x)$ verschwindet also auch für alle anderen Wurzeln von $f(x)$, die sonach alle ganze Zahlen sind.

Ist

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n,$$

so sind die $a_1, a_2, \ldots a_n$ durch Multiplikation und Addition aus den Wurzeln von $f(x)$ zusammengesetzt und sind also nach 6. ganze rationale Zahlen. Daraus folgt:

8. Ist Θ eine ganze algebraische Zahl, so hat die Gleichung niedrigsten Grades $f(\Theta) = 0$ ganzzahlige Koeffizienten.

Dasselbe ergibt sich auch aus dem Gaußschen Theorem § 20; denn danach kann eine Funktion $f(x)$ mit gebrochenen rationalen Koeffizienten nicht Teiler einer Funktion $F(x)$ mit ganzen Koeffizienten sein.

Wir beweisen noch den Satz:

9. Jede algebraische Zahl Θ läßt sich durch Multiplikation mit einer natürlichen Zahl in eine ganze algebraische Zahl verwandeln.

Denn ist

$$\Theta^m + A_1\Theta^{m-1} + \cdots + A_m = 0,$$

und sind $A_1, \ldots A_m$ rationale Zahlen mit dem gemeinsamen Nenner a, so erhält man durch Multiplikation mit a^m:

$$(a\Theta)^m + A_1 a(a\Theta)^{m-1} + A_2 a^2(a\Theta)^{m-2} + \cdots + A_m a^m = 0,$$

woraus hervorgeht, daß $a\Theta$ eine ganze Zahl ist.

Im § 53 haben wir gesehen, wie man aus jeder Wurzel Θ einer in irgend einem Körper Ω irreduziblen Gleichung $n$ten

Grades $f(\Theta) = 0$ einen algebraischen Körper $\Omega(\Theta)$ über Ω ableitet. Die aus den n Wurzeln dieser Gleichung abgeleiteten n Körper, die auch zum Teil oder alle identisch sein können, heißen konjugierte Körper.

Ein Körper, der mit allen seinen konjugierten Körpern identisch ist, heißt ein Normalkörper oder auch ein Galoisscher Körper, weil es das große Verdienst von Galois ist, die volle Bedeutung dieses Begriffes erkannt zu haben.

Bezeichnen wir mit R den Körper der rationalen Zahlen, so gibt nach unserer Definition jede algebraische Zahl $n$ten Grades, Θ, Anlaß zu einem algebraischen Körper $R(\Theta)$ über R, den wir von jetzt an kurz einen algebraischen Zahlkörper $n$ten Grades nennen. Wir haben auch schon früher nachgewiesen (§ 54), daß man immer einen algebraischen Zahlkörper bestimmen kann, der eine endliche Anzahl beliebig gegebener algebraischer Zahlen enthält.

Dieser Satz wird an dieser Stelle hervorgehoben, um darauf hinzuweisen, daß die Allgemeinheit einer Betrachtung über irgend eine endliche Anzahl algebraischer Zahlen dadurch nicht beeinträchtigt wird, daß man diese Zahlen alle in einem algebraischen Zahlkörper gelegen voraussetzt.

Jede Zahl ω eines Körpers $R(\Theta)$ kann als ganze Funktion $\varphi(\Theta)$ von Θ mit rationalen Koeffizienten dargestellt werden, und jeder Zahl ω entspricht in jedem der n konjugierten Körper eine bestimmte Zahl. Diese konjugierten Zahlen können zum Teil einander gleich sein, und wir haben danach primitive und imprimitive Zahlen des Körpers unterschieden. Jede primitive Zahl kann ebenso wie Θ selbst zur Definition des Körpers verwandt werden. Bei einer imprimitiven Zahl zerfallen die konjugierten Zahlen in Systeme von gleich vielen untereinander gleichen (§ 55).

Eine symmetrische Funktion der konjugierten Zahlen ist eine rationale Zahl. Unter diesen symmetrischen Funktionen sind zwei von besonderer Wichtigkeit, die Summe und das Produkt, von denen die erste die Spur, die zweite die Norm von ω genannt wird. Man bezeichnet diese beiden Zahlen durch $S(\omega)$ und $N(\omega)$.

Hierbei werden, wenn unter den konjugierten Zahlen dieselben Zahlen mehrfach vorkommen, diese gleichen Zahlen so oft in die Summe oder das Produkt aufgenommen, als der Grad ihrer Häufigkeit angibt.

Da sich in jedem solchen Zahlkörper die vier fundamentalen Rechenoperationen ebenso wie im Körper der rationalen Zahlen ausführen lassen, so kann man auch die Frage aufwerfen, inwieweit sich die aus der Theorie der rationalen Zahlen bekannten arithmetischen Grundgesetze in einem beliebigen algebraischen Zahlkörper bewähren. Es handelt sich hierbei in erster Linie um die Zerlegung der ganzen Zahlen in ihre Primfaktoren.

Da diese Zerlegung mit den Zahlen des algebraischen Körpers selbst im allgemeinen nicht gelingt, so ist eine Erweiterung des Rechenmateriales nötig, um die einfachen Gesetze wieder herzustellen, und eine solche Erweiterung ist in verschiedenem Sinne möglich. Es müssen sich aber diese verschiedenen Erweiterungen aufeinander zurückführen oder, genauer gesagt, in eine eindeutige Beziehung zueinander setzen lassen.

Kummer hat zuerst für die aus Einheitswurzeln gebildeten algebraischen Zahlen (die Kreisteilungszahlen) das große Problem durch die Schöpfung der idealen Zahlen[1]) gelöst. Eine andere ganz allgemeine, keiner Ausnahme unterworfene Lösung hat die Aufgabe durch Dedekind gefunden, der als die einfachsten Elemente der Rechnung die von ihm so genannten Ideale[2]) betrachtet. Einen davon verschiedenen Weg hat Kronecker[3]) eingeschlagen.

---

[1]) Kummer, Disq. de numeris complexis etc. Vratisl. 1844. Theorie der idealen Primfaktoren der komplexen Zahlen usw. Crelles Journal, Bd. 35, 1846; Bd. 40, 1850. Abhandlungen der Berliner Akademie 1856. Sur la théorie des nombres complexes composés de racines de l'unité et de nombres entiers. Liouvilles Journal, Bd. 16, 1851.

[2]) Dedekind, in dem letzten Supplement der 2., 3. und 4. Auflage von Dirichlets Vorlesungen über Zahlentheorie (Braunschweig 1871, 1879, 1894). Zu vergleichen ist auch: Sur la théorie des nombres entiers algébriques im Bulletin von Darboux und Houël (1ère sér. XI, 1877). Über den Zusammenhang zwischen der Theorie der Ideale und der höheren Kongruenzen (Abhandlungen d. Ges. d. Wissensch. in Göttingen, Bd. 23, 1878). Über die Diskriminanten endlicher Körper (ebenda Bd. 29, 1882). Über die Anzahl der Idealklassen in den verschiedenen Ordnungen eines endlichen Körpers, Festschrift zur Säkularfeier des Geburtstages von Gauß (Braunschweig 1877). „Zur Theorie der Ideale" und „Über die Begründung der Idealtheorie" (Nachrichten d. Ges. d. Wissensch. in Göttingen 1894, 1895). Hierher gehören auch die Abhandlungen von Hilbert, „Über die Zerlegung der Ideale usw.", Mathem. Annalen, Bd. 44, 1893. „Grundzüge einer Theorie der Galoisschen Zahlkörper" (Göttinger Nachrichten 1894). Die Theorie der algebraischen Zahlkörper (Bericht der Deutschen Mathematiker-Vereinigung 1897). Hurwitz, „Zur Theorie der Ideale". „Über einen Fundamentalsatz usw." (Göttinger Nachrichten 1894, 1895).

[3]) Kronecker, Grundzüge einer arithmetischen Theorie der alge-

Die Theorie von Dedekind ist von ihrem Begründer umfassend und in stets wachsender Einfachheit und Vollkommenheit dargestellt in dem letzten Supplement der drei neuesten Auflagen von Dirichlets Vorlesungen über Zahlentheorie.

Die Theorie Kroneckers ist erst im Jahre 1882 durch die Festschrift zu Kummers Jubiläum dem weiteren Kreise der Mathematiker bekannt geworden.

Auf einem neuen Wege hat Hensel die Theorie der algebraischen Zahlen begründet, der überraschend schnell zu einigen der wichtigsten Sätze, namentlich in bezug auf die Diskriminanten, führt, die sonst nur auf längeren Umwegen zu beweisen waren. Hensel bedient sich einer Art Reihenentwickelungen der algebraischen Zahlen, deren zu jeder natürlichen Primzahl eine gewisse Anzahl gehören. Hierdurch wird die Theorie der algebraischen Zahlen in schöne Übereinstimmung mit der von Riemann und Weierstraß ausgebildeten Theorie der algebraischen Funktionen gesetzt.

§ 95.

Ganze Funktionen in einem algebraischen Körper.

Wir haben schon mehrfach Gelegenheit gehabt, ganze Funktionen einer beliebigen Anzahl von unabhängigen Veränderlichen einzuführen, und haben auch (im § 20) den Fall erörtert, daß die Koeffizienten einem bestimmten Körper Ω angehören. Wir machen jetzt die Annahme, daß dieser Körper ein algebraischer Zahlkörper sei und betrachten also Ausdrücke $\varphi(x, y, z \ldots)$, die als eine Summe von Gliedern der Form

$$\alpha x^r y^s z^t \ldots$$

dargestellt sind, worin die Exponenten r, s, t ... positive oder wenigstens nicht negative ganze Zahlen sind, während die Koeffizienten α Zahlen in Ω bedeuten. Einen solchen Ausdruck

$$\varphi(x, y, z, \ldots) = \Sigma \alpha x^r y^s z^t \ldots \tag{1}$$

nennen wir eine ganze Funktion in Ω. Wir nehmen den Ausdruck immer so geordnet und zusammengefaßt an, daß dieselbe

braischen Größen, Festschrift zu Kummers 50jährigem Doktorjubiläum. Berlin 1882. (Auch in Bd. 92 von Crelles Journal.) Zu erwähnen sind hier noch die Arbeiten von Hensel in den Bänden 101, 103, 105, 111, 113 des Crelleschen Journals und dessen Buch „Theorie der algebraischen Zahlen", Leipzig und Berlin 1908, bei Teubner.

Kombination der Exponenten r, s, t, ... nicht zweimal darin vorkommt, und nennen zwei solche Ausdrücke nur dann einander gleich, wenn sie dieselben Produkte $x^r y^s z^t \ldots$, mit denselben Koeffizienten behaftet, enthalten. Eine ganze Funktion wird dann und nur dann gleich Null gesetzt, wenn alle ihre Koeffizienten Null sind.

Mehrere ganze Funktionen geben durch Addition, Subtraktion und Multiplikation immer wieder ganze Funktionen. Nach § 14 kann man für die Variablen x, y, z, ... solche rationale Zahlwerte setzen, daß eine oder eine beliebige Anzahl von gegebenen, von Null verschiedenen ganzen Funktionen in Ω nicht verschwindende Zahlwerte (in Ω) erhalten. Daraus ergibt sich, daß ein Produkt mehrerer ganzer Funktionen nur dann verschwindet, wenn einer seiner Faktoren verschwindet.

Die Summe $r + s + t + \cdots$ der Exponenten in einem Gliede des Ausdruckes (1) heißt der Grad dieses Gliedes, und der größte Wert, den der Grad eines Gliedes in φ annimmt, heißt der Grad der Funktion φ.

Der Grad eines Produktes aus zweien oder mehreren ganzen Funktionen ist gleich der Summe der Grade der einzelnen Faktoren.

Denn faßt man in jedem der Faktoren die Summe der Glieder höchsten Grades zu einer homogenen Funktion zusammen, so erhält man die Glieder höchsten Grades des Produktes, wenn man alle diese homogenen Funktionen miteinander multipliziert. Das Produkt dieser homogenen Funktionen kann nach dem oben Bemerkten nicht verschwinden, wenn keiner der Faktoren verschwindet, und sein Grad ist gleich der Summe der Grade der einzelnen Faktoren.

Die Zahlen des Körpers Ω sind unter den Funktionen mit enthalten. Man erhält sie, wenn man entweder den Grad oder die Anzahl der Variablen auf Null heruntersinken läßt.

Als spezielle Fälle sind unter den ganzen Funktionen in Ω auch die ganzen Funktionen im Körper der rationalen Zahlen R enthalten:

$$\Phi(x, y, z \ldots) = \Sigma a x^r y^s z^t \ldots, \tag{2}$$

worin die Koeffizienten a rationale Zahlen sind.

Wenn diese Koeffizienten ganze Zahlen ohne gemeinsamen Teiler sind, so heißt die Funktion eine ursprüngliche

oder primitive. Ist m der größte gemeinschaftliche Teiler der Zahlen a, so heißt m auch der Teiler der Funktion Φ. Nach § 20 gilt der Satz:

> Der Teiler eines Produktes von zwei oder mehr ganzen Funktionen Φ ist gleich dem Produkt der Teiler der einzelnen Faktoren und das Produkt von primitiven Funktionen ist wieder eine primitive Funktion.

Die ganzen Funktionen in einem algebraischen Zahlkörper Ω hängen außer von den Variablen von einer algebraischen Zahl Θ ab, enthalten aber sonst nur rationale Zahlenkoeffizienten. Bezeichnen wir eine solche Funktion mit $\varphi(\Theta, x, y, z \ldots)$, so erhalten wir die konjugierten Funktionen φ, $\varphi_1, \varphi_2 \ldots$ oder

(3) $\varphi(\Theta, x, y, z \ldots)$, $\varphi(\Theta_1, x, y, z \ldots)$, $\varphi(\Theta_2, x, y, z \ldots)$, $\ldots$

wenn wir für Θ die sämtlichen Wurzeln der irreduziblen Gleichung $f(x) = 0$ [§ 94, (2)] einsetzen. Diese konjugierten Funktionen können auch zum Teil einander gleich sein.

Sie sind alle einander gleich, wenn φ eine Funktion in R ist, und es ist umgekehrt φ eine Funktion in R, wenn die konjugierten Funktionen alle einander gleich sind; denn es ist dann, wenn wir unter $S(\varphi)$ die Summe der konjugierten Funktionen (die Spur) verstehen,

$$n\varphi = S(\varphi),$$

und $S(\varphi)$ ist eine ganze Funktion, deren Koeffizienten symmetrische Funktionen der n Wurzeln Θ, d. h. rationale Zahlen, sind.

Zu den ganzen Funktionen in R gehört auch die Norm von φ, d. h. das Produkt

(4) $$N(\varphi) = \varphi\varphi_1\varphi_2 \ldots,$$

denn alle Koeffizienten dieser Funktion sind symmetrische Funktionen der Θ.

§ 96.

Funktionale.

Die Variablen, die in der Theorie der algebraischen Zahlen verwendet werden, haben nicht die Bedeutung von Zeichen für veränderliche Zahlenreihen, wie man es aus der Funktionentheorie gewöhnt ist, sondern sie sind lediglich Rechnungssymbole ohne eine selbständige Bedeutung. Bei den Funktionen dieser Variablen kommt es eigentlich nur auf die Koeffizientensysteme

an, und die Variablen werden nur dazu benutzt, um die bekannten und geläufigen Regeln der Buchstabenrechnung auf diese Koeffizientensysteme anzuwenden. Es ist damit freilich nicht ausgeschlossen, daß gelegentlich auch die Zahlen betrachtet werden, die man erhält, wenn man die Variablen durch gewisse Zahlen, z. B. durch rationale Zahlen, ersetzt.

Demnach führen wir in unsere Betrachtungen sowohl ganze als gebrochene Funktionen von beliebig vielen Veränderlichen ein, deren Koeffizienten Zahlen eines algebraischen Körpers Ω sind, und setzen fest, daß mit diesen Funktionen so gerechnet wird, wie es die Buchstabenrechnung vorschreibt.

Jede solche Funktion ω kann als Quotient zweier ganzer Funktionen in Ω,

$$\omega = \frac{\varphi}{\psi}, \tag{1}$$

dargestellt werden, wobei ψ immer von Null verschieden angenommen werden muß. Zwei solche Funktionen heißen nur dann einander gleich:

$$\frac{\varphi}{\psi} = \frac{\varphi_1}{\psi_1},$$

wenn $\varphi\psi_1 = \varphi_1\psi$ ist. Haben die beiden Funktionen φ, ψ keinen gemeinsamen Teiler, so heißt $\varphi : \psi$ ein irreduzibler Bruch. Unter den verschiedenen Darstellungen einer Funktion ω gibt es eine durch einen irreduziblen Bruch, und diese wollen wir die einfachste Darstellung nennen. In ihr sind Zähler und Nenner bis auf einen gemeinsamen Faktor, der eine beliebige Zahl in Ω sein kann, durch ω selbst völlig bestimmt.

Eine solche Funktion ω wollen wir ein Funktional des Körpers Ω nennen. Als spezielle Fälle sind darunter die ganzen Funktionen und die Zahlen selbst enthalten. Bei den Zahlen ist jede Darstellung als Quotient zweier Zahlen in Ω als einfachste Darstellung zu betrachten.

Auf die Funktionale lassen sich die vier Grundrechnungsarten in demselben Umfange anwenden, wie auf die Zahlen, und der Inbegriff aller Funktionale des Körpers Ω ist daher gleichfalls ein Körper, den wir mit $\overline{\Omega}$ bezeichnen und den Funktionalkörper von Ω nennen wollen.

Der Funktionalkörper $\overline{\Omega}$ enthält den Zahlkörper Ω als Teiler.

Ein Funktional, dessen Koeffizienten rationale Zahlen sind, heißt ein rationales Funktional. Der Inbegriff aller rationalen Funktionale ist der Funktionalkörper $\overline{R}$ des Körpers R der rationalen Zahlen, und der Körper $\overline{R}$ ist in jedem algebraischen Funktionalkörper $\overline{\Omega}$ enthalten.

Jedes rationale Funktional A kann als Quotient zweier ganzer Funktionen in R dargestellt werden, denen man auch ganzzahlige Koeffizienten geben kann, indem man Zähler und Nenner mit dem Hauptnenner aller Koeffizienten multipliziert. Sind in dieser Darstellung Zähler und Nenner imprimitiv, so kann man den Teiler herausnehmen und erhält eine Darstellung in der Form

$$A = a\,\frac{E_1(x,y,z\ldots)}{E_2(x,y,z\ldots)} = a\,\frac{E_1}{E_2} = a\,E, \tag{2}$$

in der a eine positive, ganze oder gebrochene, rationale Zahl ist, während E_1, E_2 primitive Funktionen in R sind. Hieraus folgt:

1. Man kann jedes rationale Funktional durch Multiplikation mit einer primitiven Funktion in R in eine ganze Funktion in R verwandeln, und mehrere rationale Funktionale lassen sich als Brüche darstellen, deren gemeinsamer Nenner eine primitive Funktion ist.

Die positive Zahl a und der Quotient $E_1 : E_2$ sind durch A vollständig bestimmt. Denn setzen wir a in die Form eines rationalen Bruches $q_1 : q_2$ und nehmen an, es sei

$$\frac{q_1}{q_2}\,\frac{E_1}{E_2} = \frac{q_1'}{q_2'}\,\frac{E_1'}{E_2'},$$

so folgt:

$$q_1 q_2' E_1 E_2' = q_2 q_1' E_2 E_1'.$$

Hier haben wir also zwei ganze Funktionen in R mit ganzzahligen Koeffizienten, die einander gleich sind und deren Teiler, da $E_1 E_2'$ und $E_2 E_1'$ primitive Funktionen sind, $q_1 q_2'$ oder $q_2 q_1'$ ist. Folglich ist $q_1 q_2' = q_2 q_1'$, und daher auch

$$\frac{q_1}{q_2} = \frac{q_1'}{q_2'}, \quad \frac{E_1}{E_2} = \frac{E_1'}{E_2'}.$$

2. Wir nennen die positive rationale Zahl a den absoluten Wert des Funktionals A.

In dem Falle, daß A selbst eine Zahl ist, ist a der absolute Wert von A in dem gewöhnlichen Sinne dieses Wortes, und E ist $= +1$ oder $= -1$ zu setzen, je nachdem A positiv oder negativ ist. Es ist also E in diesem Falle nichts weiter als das Vorzeichen von A. In der weitgehenden Verallgemeinerung dieser elementaren Begriffe liegt das Befremdende, was unsere Definition dem ersten Blick bietet. Sie wird sich aber in der Folge als durchaus sachgemäß und nützlich erweisen.

Aus § 95 ergibt sich der Satz:

3. Der absolute Wert eines Produktes zweier oder mehrerer rationaler Funktionale ist gleich dem Produkte der absoluten Werte der Faktoren.

Ist ω ein Funktional des Körpers Ω, so ist $N(\omega)$ ein rationales Funktional, und die Norm eines Produktes oder eines Quotienten zweier Funktionale ist gleich dem Produkte oder dem Quotienten der Normen der Bestandteile.

4. Wir nennen den absoluten Wert des rationalen Funktionals $N(\omega)$ die absolute Norm $N_a(\omega)$ von ω und setzen

$$N(\omega) = N_a(\omega)\, E(\omega). \tag{3}$$

$N_a(\omega)$ ist immer eine positive rationale Zahl, und $E(\omega)$ ist der Quotient zweier primitiver ganzer Funktionen. Aus 3. folgt, wenn α, β zwei Funktionale in Ω sind, die Formel

$$N_a(\alpha\beta) = N_a(\alpha)\, N_a(\beta), \tag{4}$$

oder der Satz:

5. Die absolute Norm eines Produktes von Funktionalen in Ω ist gleich dem Produkte der absoluten Normen der Faktoren.

Wenn wir alle Koeffizienten eines Funktionals in Ω durch die entsprechenden Zahlen eines zu Ω konjugierten Körpers Ω_1 ersetzen, so erhalten wir ein konjugiertes Funktional. Der gesamte Funktionalkörper $\overline{\Omega}$ geht dadurch in einen konjugierten Funktionalkörper $\overline{\Omega_1}$ über. Da jede Gleichung zwischen Funktionalen des Körpers Ω im Grunde nur die Zusammenfassung einer Reihe von Gleichungen zwischen Zahlen des Körpers Ω ist, so haben wir den Satz:

6. Jede Gleichung zwischen Funktionalen des Körpers $\overline{\Omega}$ bleibt richtig, wenn für alle Funktionale

die entsprechenden Elemente eines konjugierten Körpers $\overline{\Omega_1}$ gesetzt werden.

Bedeutet t eine in ω nicht vorkommende Variable, so hat die Norm $N(t - \omega)$ die Form

$$\Phi(t) = t^n + A_1 t^{n-1} + A_2 t^{n-2} + \cdots + A_n, \tag{5}$$

worin die Koeffizienten A_i rationale Funktionale sind, und diese Funktion $\Phi(t)$ verschwindet, wenn ω für t gesetzt wird. Es gibt aber nicht bloß eine solche Funktion $\Phi(t)$, die für $t = \omega$ verschwindet, sondern beliebig viele, da man jedes $\Phi(t)$ mit einer beliebigen Funktion der gleichen Form multiplizieren kann. Auch kann es vorkommen, daß schon ein Produkt von weniger als n Faktoren von $N(t - \omega)$ rationale Koeffizienten erhält, woraus Funktionen $\Phi(t)$ von niedrigerem als dem $n$ten Grade entspringen, die für $t = \omega$ verschwinden. Daraus folgt:

7. Es gibt für jedes Funktional ω des Körpers Ω unendlich viele Funktionen $\Phi(t)$, die für $t = \omega$ verschwinden, in denen die höchste Potenz von t den Koeffizienten 1 hat, und deren übrige Koeffizienten in $\overline{R}$ enthalten sind.

Wir sagen dann, ω ist eine Wurzel der Gleichung

$$\Phi(t) = 0.$$

Unter den verschiedenen Funktionen $\Phi(t)$, deren Existenz im Satze 7. ausgesprochen ist, gibt es eine und nur eine $F(t)$ von möglichst niedrigem Grade. Denn existieren zwei solche Funktionen $F(t)$ und $F_1(t)$ von gleichem Grade m, so ist die Differenz $F(t) - F_1(t)$ in bezug auf t höchstens vom Grade $m - 1$ und verschwindet für $t = \omega$. Dividiert man durch den Koeffizienten der höchsten Potenz von t, so erhält man eine Funktion $\Phi(t)$ von niedrigerem Grade als $F(t)$, die nach der Voraussetzung über $F(t)$ nicht existiert.

Die Funktion $F(t)$ kann im Körper $\overline{R}$ nicht in Faktoren zerlegt werden, die ganze Funktionen von t sind. Denn zerfiele sie in mehrere Faktoren derselben Form $F_1(t)$, $F_2(t)$, so müßte einer dieser Faktoren, die doch alle von niedrigerem Grade als $F(t)$ selbst sind, für $t = \omega$ verschwinden, entgegen unserer Voraussetzung.

Jede Funktion $\Phi(t)$ des Satzes 7. ist durch diese irreduzible Funktion $F(t)$ teilbar, so daß der Quotient $\Phi(t) : F(t)$

eine ganze Funktion von t in $\overline{R}$ ist, in der die höchste Potenz von t den Koeffizienten 1 hat.

Unter den Funktionen $\Phi(t)$ findet sich, wie schon bemerkt, auch die Norm $N(t - \omega)$, und folglich ist $N(t - \omega)$ durch $F(t)$ teilbar. Sind beide Funktionen von gleichem Grade, so ist $N(t - \omega) = F(t)$. Ist aber der Grad von $F(t)$ niedriger als n, so verschwindet der Quotient $N(t - \omega) : F(t)$ wenigstens noch für eines der mit ω konjugierten Funktionale, und folglich für alle; folglich auch für $t = \omega$, und daher ist dieser Quotient nochmals durch $F(t)$, d. h. $N(t - \omega)$ ist durch $F(t)^2$ teilbar. Durch Fortsetzung dieses Schlußverfahrens erkennt man, daß $N(t - \omega)$ eine Potenz von $F(t)$ sein muß, und daß folglich der Grad von $F(t)$ ein Teiler von n ist.

Wir fassen dies noch in dem Satze zusammen:

8. Jedes Funktional in Ω ist die Wurzel einer und nur einer irreduziblen Gleichung $F(t) = 0$ in $\overline{R}$, und $N(t - \omega)$ ist eine Potenz von $F(t)$.

§ 97.

Ganze Funktionale.

1. **Definition**: Ein rationales Funktional soll ganz genannt werden, wenn sein absoluter Wert eine ganze Zahl ist.

Ein ganzes rationales Funktional ist also nach dieser Definition keineswegs notwendig eine ganze Funktion der Variablen. Aus der Definition ergibt sich zunächst, daß die Summe, die Differenz und das Produkt von zwei und folglich auch von beliebig vielen ganzen rationalen Funktionalen wieder ganz sind. Für das Produkt ist dies eine unmittelbare Folge des Satzes § 96, 3. Um aber den Beweis für die Summe und die Differenz zu führen, stellen wir zwei ganze rationale Funktionale A_1, A_2 so durch gebrochene Funktionen dar, daß sie eine primitive Funktion als gemeinschaftlichen Nenner erhalten:

$$A_1 = a_1 \frac{E_1}{E}, \quad A_2 = a_2 \frac{E_2}{E}.$$

Dann ist

$$A_1 \pm A_2 = \frac{a_1 E_1 \pm a_2 E_2}{E};$$

da nun a_1, a_2 ganze Zahlen sind, so ist der Teiler der ganzen Funktion $a_1 E_1 \pm a_2 E_2$ der absolute Wert von $A_1 \pm A_2$. Dieser ist eine ganze Zahl, also $A_1 \pm A_2$ ein ganzes Funktional.

Für konstante rationale Funktionale, d. h. für rationale Zahlen, gibt die Definition 1. die ganzen rationalen Zahlen.

Wir stellen ferner, ebenso wie im § 94 für die Zahlen, folgende Definition der ganzen Funktionale des Körpers Ω auf:

2. **Definition:** Ein Funktional ω aus $\overline{\Omega}$ heißt ganz, wenn es die Wurzel einer Gleichung

$$\Phi(t) = t^m + A_1 t^{m-1} + A_2 t^{m-2} + \cdots + A_m = 0$$

ist, in der die Koeffizienten A_1, A_2, ... A_m ganze rationale Funktionale sind.

Funktionale, die nicht zu den ganzen gehören, werden wir gelegentlich auch der Kürze wegen als gebrochene Funktionale bezeichnen.

Zur Rechtfertigung dieser Definition beweisen wir zunächst den Satz:

3. Ist ω ein ganzes Funktional nach der Definition 2. und zugleich rational, so ist es ein ganzes rationales Funktional (nach der Definition 1.).

Angenommen, es sei ω ein gebrochenes rationales Funktional, und daher der absolute Wert von ω ein rationaler Bruch $p:q$, worin p und q ganze rationale Zahlen ohne gemeinsamen Teiler sind, dann ist $q\omega$ ein ganzes rationales Funktional, dessen absoluter Wert $= p$, also relativ prim zu q ist. Wenn aber andererseits ω zugleich ganz im Sinne der Definition 2. ist, so können wir

$$\omega^m = -(A_1 \omega^{m-1} + A_2 \omega^{m-2} + \cdots)$$

setzen, woraus

$$(q\omega)^m = -q[A_1 (q\omega)^{m-1} + A_2 q (q\omega)^{m-2} + \cdots]$$

folgt.

Hieraus aber ergibt sich, daß der absolute Wert p^m von $(q\omega)^m$ durch q teilbar sein muß, was nur möglich ist, wenn $q = 1$ ist. Die Definition 1. ist also in der Definition 2. als Spezialfall enthalten.

Ist

$$\Phi(t) = t^m + A_1 t^{m-1} + A_2 t^{m-2} + \cdots + A_m$$

eine Funktion von t, in der die Koeffizienten A_1, A_2, ... A_m gebrochene rationale Funktionale mit den Variablen x, y ..., aber

von der Variablen t frei sind, so kann man eine Funktion $aE(x, y, \ldots)$ bestimmen, in der a den Hauptnenner der absoluten Werte von $A_1, A_2 \ldots$ und $E(x, y \ldots)$ eine primitive ganze Funktion bedeutet, so daß

$$aE(x, y \ldots)\, \Phi(t) = P(t, x, y \ldots) \tag{1}$$

selbst eine primitive ganze Funktion ist (§ 96, 1.).

Zerfällt $\Phi(t)$ in zwei Faktoren $\Phi_1(t)$, $\Phi_2(t)$ in R, ist also

$$\Phi(t) = \Phi_1(t)\, \Phi_2(t),$$

so bestimme man hiernach für die beiden Funktionen Φ_1, Φ_2 die Faktoren $a_1 E_1$, $a_2 E_2$, so daß

$$a_1 E_1 \Phi_1 = P_1, \quad a_2 E_2 \Phi_2 = P_2$$

primitive Funktionen werden, und dann ist auch

$$a_1 a_2 E_1 E_2 \Phi = P_1 P_2 \tag{2}$$

eine primitive Funktion. Aus (1) und (2) folgt:

$$a_1 a_2 E_1 E_2 P = aE P_1 P_2, \tag{3}$$

und mithin

$$a = a_1 a_2. \tag{4}$$

Wenn also $a = 1$ ist, so müssen die natürlichen Zahlen a_1, a_2 auch $= 1$ sein, und wir haben den Satz:

4. Ist ω ein ganzes Funktional in Ω und $\Phi(t)$ eine Funktion von t mit ganzen Koeffizienten in $\overline{R}$, die für $t = \omega$ verschwindet, so hat auch jeder rationale Teiler von $\Phi(t)$ ganze Koeffizienten in $\overline{R}$; insbesondere hat die irreduzible Funktion $F(t)$, von der ω nach § 96, 8. eine Wurzel ist, ganze rationale Funktionale zu Koeffizienten.

Wenn ein Funktional ω nicht ganz ist, so genügt es einer Gleichung

$$\Phi(\omega) = \omega^m + A_1 \omega^{m-1} + A_2 \omega^{m-2} + \cdots + A_m = 0,$$

worin die Koeffizienten A_1, A_2, $\ldots$ A_m zwar rationale, aber nicht ganze Funktionale sind. Ist a der Hauptnenner der absoluten Werte von A_1, A_2, $\ldots$ A_m, so sind aA_1, aA_2, $\ldots$ aA_m ganze Funktionale. Nun ist

$$a^m \Phi(\omega) = (a\omega)^m + aA_1 (a\omega)^{m-1} + a^2 A_2 (a\omega)^{m-2} + \cdots + a^m A_m = 0,$$

und $a\omega$ ist daher ein ganzes Funktional. Daraus folgt:

5. Jedes Funktional ω des Körpers Ω läßt sich durch Multiplikation mit einer ganzen rationalen Zahl in ein ganzes Funktional verwandeln, und daher kann man jedes Funktional ω als Quotienten zweier ganzer Funktionale darstellen. Diese Darstellung ist auf unendlich viele verschiedene Arten möglich, unter anderem so, daß der Nenner eine natürliche Zahl ist.

6. Ist ω ein Funktional in Ω und gibt es m Größen $\omega_1, \omega_2, \ldots \omega_m$, die nicht alle verschwinden, von der Art, daß die m Produkte $\omega\omega_i$ für $i = 1, 2 \ldots m$ in die Form gesetzt werden können:

$$\omega\omega_i = A_{1,i}\omega_1 + A_{2,i}\omega_2 + \cdots + A_{m,i}\omega_m, \tag{5}$$

worin die m^2 Symbole $A_{k,i}$ ganze rationale Funktionale sind, so ist ω ein ganzes Funktional.

Hierin ist m irgend eine ganze natürliche Zahl. Die ω_i sind in der Anwendung immer Zahlen oder Funktionale in Ω, jedoch ist für die Gültigkeit des Satzes nur die Ausführbarkeit der in (5) angedeuteten Multiplikation wesentlich.

Der Satz ist eine einfache Folge der Definition der ganzen Funktionale; denn da die ω_i nicht alle verschwinden, so muß nach dem Determinantensatze

$$\begin{vmatrix} A_{1,1} - \omega, & A_{2,1} & \ldots & A_{m,1} \\ A_{1,2}, & A_{2,2} - \omega & \ldots & A_{m,2} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ A_{1,m}, & A_{2,m} & \ldots & A_{m,m} - \omega \end{vmatrix} = 0 \tag{6}$$

sein. Durch Ordnen nach Potenzen von ω ergibt sich hieraus eine Gleichung:

$$\omega^m + A_1\omega^{m-1} + \cdots + A_m = 0, \tag{7}$$

worin die Koeffizienten A_1, A_2 ... durch Addition und Multiplikation aus den $A_{i,k}$ zusammengesetzt sind und daher selbst ganze rationale Funktionale sind; demnach ist auch ω ein ganzes Funktional.

Für den letzten Koeffizienten A_m in der Gleichung (7) erhalten wir den Ausdruck:

$$(-1)^m A_m = \Sigma \pm A_{1,1} A_{2,2} \ldots A_{m,m}, \tag{8}$$

und diese Determinante ist also gleich dem Produkte der sämtlichen Wurzeln der Gleichung (6) oder (7).

Der Satz 6. ist wichtig als Kennzeichen für ganze Funktionale; er dient uns hier zum Beweise des folgenden Satzes:

7. Ist $\Psi(x, y \ldots)$ eine ganze Funktion, deren Koeffizienten ganze rationale Zahlen oder Funktionale, aber frei von den Variablen x, y ... sind, sind ferner α, β ... ganze Funktionale in Ω, so ist

$$\omega = \Psi(\alpha, \beta \ldots) \tag{9}$$

auch ein ganzes Funktional.

Es seien nämlich μ, ν ... die Grade der ganzzahligen Gleichungen, denen (nach 2.) die Funktionale α, β ... genügen, und $m = \mu\nu \ldots$ Wir verstehen unter $\omega_1, \omega_2, \ldots \omega_m$ die m Größen

$$\alpha^r \beta^s \ldots; \quad r = 0,1 \ldots \mu - 1; \quad s = 0,1 \ldots \nu - 1; \quad \ldots$$

Dann können mit Hilfe der Gleichungen μten, νten ... Grades, denen die Zahlen α, β ... genügen, alle Produkte $\alpha^r \beta^s \ldots$, in denen einer der Exponenten r, s ... größer als $\mu - 1$, $\nu - 1$ ... ist, linear in der Form (5) durch $\omega_1, \omega_2, \ldots \omega_m$ ausgedrückt werden, und dasselbe gilt daher auch von jedem Funktional ω der Form (9), und folglich auch von den Produkten $\omega\omega_1, \omega\omega_2, \ldots \omega\omega_m$. Daraus folgt nach 6., daß ω ganz ist, wie wir beweisen wollten.

Als speziellen Fall des Satzes 7., aus dem übrigens der allgemeine Satz leicht wieder gefolgert werden kann, heben wir hervor:

8. Durch Addition, Subtraktion und Multiplikation ganzer Funktionale entstehen immer wieder ganze Funktionale.

Wir haben schon früher (§ 94) bemerkt, daß man immer einen algebraischen Körper bestimmen kann, der beliebig gegebene algebraische Zahlen enthält. Daraus folgt, daß in dem Satze 7. die α, β ... beliebige ganze algebraische Zahlen oder Funktionale sein können, und Entsprechendes gilt von dem Satze 8.

Aus der Definition 2 folgt noch nach dem Satze § 96, 6.:

9. Ist ω ein ganzes Funktional des Körpers Ω, so sind auch alle mit ω konjugierte Funktionale ganz.

Als besondere Folgerung dieser Sätze sei noch erwähnt:

10. Die absolute Norm eines ganzen Funktionals ist eine natürliche ganze Zahl.

Wir beweisen endlich noch den folgenden Satz:

11. Wenn ein Funktional ω einer Gleichung von der Form

$$\omega^m + \alpha_1 \omega^{m-1} + \cdots + \alpha_m = 0 \tag{10}$$

genügt, in der $\alpha_1, \alpha_2 \ldots \alpha_m$ ganze Funktionale in Ω sind, so ist auch ω ein ganzes Funktional.

Um ihn zu beweisen, bezeichnen wir mit t eine in den α und in ω nicht vorkommende Variable und setzen:

$$\varphi(t) = t^m + \alpha_1 t^{m-1} + \cdots + \alpha_m. \tag{11}$$

Dann ist, wenn n der Grad des Körpers Ω ist,

$$\Phi(t) = N[\varphi(t)]$$

eine ganze Funktion mnten Grades von t, deren Koeffizienten ganze rationale Funktionale sind. Zugleich ist $\Phi(\omega) = 0$, und folglich ω ein ganzes Funktional (nach 2.).

Daraus folgt noch, daß der Satz 6. richtig bleibt, wenn die Koeffizienten $A_{k,i}$ nicht ganze Funktionale in R, sondern in Ω sind.

Ist ein ganzes Funktional ω zugleich eine Zahl, so ist es eine ganze Zahl, weil in diesem Falle $N(t - \omega)$ ganze rationale Zahlen zu Koeffizienten hat. Die ganzen Zahlen, die wir im § 94 betrachtet haben, sind demnach als spezielle Fälle unter den ganzen Funktionalen enthalten.

§ 98.

Teilbarkeit. Einheiten.

Die ganzen Funktionale unterliegen ähnlichen Gesetzen der Teilbarkeit, wie die ganzen rationalen Zahlen. Um sie zu erkennen, stellen wir folgende Definition an die Spitze:

1. Ein ganzes Funktional α heißt durch ein anderes, von Null verschiedenes ganzes Funktional β teilbar, wenn der Quotient $\alpha : \beta = \gamma$ ein ganzes Funktional ist.

Es ist dann $\alpha = \beta\gamma$, und β, γ heißen Teiler von α, und man sagt auch, β und γ gehen in α auf. Die Zahl 0 ist durch jedes ganze Funktional teilbar.

Aus dieser Definition ergeben sich ohne Schwierigkeit die folgenden fundamentalen Sätze über Teilbarkeit:

2. Sind α und α_1 teilbar durch β, so ist auch $\alpha \pm \alpha_1$ teilbar durch β.

Denn ist $\alpha = \beta\gamma$, $\alpha_1 = \beta\gamma_1$, so ist $\alpha \pm \alpha_1 = \beta(\gamma \pm \gamma_1)$, und wenn γ und γ_1 ganz sind, so ist auch $\gamma \pm \gamma_1$ ganz.

3. Ist α teilbar durch β, und β teilbar durch γ, so ist auch α teilbar durch γ.

Denn nach der Voraussetzung gibt es zwei ganze Funktionale $\varkappa$, λ, die den Bedingungen $\alpha = \varkappa\beta$, $\beta = \lambda\gamma$ genügen. Demnach ist auch $\alpha = \varkappa\lambda\gamma$, und da $\varkappa\lambda$ ganz ist, so ist α durch γ teilbar.

Ein Produkt $\beta\gamma$ zweier ganzer Funktionale ist sowohl durch β als durch γ teilbar, und folglich ist, wenn β durch α teilbar ist, auch $\beta\gamma$ durch α teilbar. Aus diesen Sätzen ergibt sich:

4. Sind α, β ... durch δ teilbar, und sind ξ, η ... beliebige ganze Funktionale, so ist $\xi\alpha + \eta\beta + \cdots$ durch δ teilbar.

Selbstverständlich erstrecken sich diese Definitionen und Sätze auch auf den Fall, daß Zahlen an die Stelle von Funktionalen treten, und so erhalten wir die Teilbarkeit ganzer Zahlen.

5. Zwei ganze Funktionale α, β, die gegenseitig durcheinander teilbar sind, heißen assoziiert.

6. Ein mit der natürlichen Zahl 1 assoziiertes ganzes Funktional ε, d. h. jeder Teiler der Zahl 1, heißt eine Einheit.

Je nachdem ε ein Funktional oder eine Zahl ist, ist ε eine funktionale oder eine numerische Einheit.

Im Körper R der rationalen Zahlen sind als funktionale Einheiten die primitiven ganzen Funktionen und die Quotienten von zweien unter ihnen anzusehen. Numerische Einheiten gibt es in R nur zwei, nämlich $+1$ und -1.

Über die hierdurch eingeführten Begriffe, die in enger gegenseitiger Beziehung stehen, leiten wir eine Reihe von Sätzen ab.

7. Sind α, β assoziierte Funktionale, so sind die Quotienten $\beta:\alpha$ und $\alpha:\beta$ Einheiten.

Denn setzen wir $\beta = \alpha\varepsilon$, so ist ε ein ganzes Funktional und $1:\varepsilon = \alpha:\beta$ ist gleichfalls ganz; also ist ε ein Teiler der Zahl 1, d. h. ε ist eine Einheit.

8. Ist α ein ganzes Funktional und ε eine Einheit, so sind α und $\alpha\varepsilon$ assoziiert.

Dies folgt unmittelbar aus der Definition; denn $\alpha:\alpha\varepsilon = 1:\varepsilon$ und $\alpha\varepsilon:\alpha = \varepsilon$ sind beides ganze Funktionale.

Eine Einheit ist ein ganzes Funktional, dessen reziprokes gleichfalls ganz ist, und dieses reziproke Funktional ist selbst eine Einheit. Durch eine Einheit ist jedes beliebige ganze Funktional teilbar. Überhaupt gilt der Satz:

9. Das Produkt und der Quotient zweier Einheiten sind wieder Einheiten.

Denn sind ε_1, ε_2 zwei Einheiten, so sind $\varepsilon_1\varepsilon_2$ und $1:\varepsilon_1\varepsilon_2$ ganze Funktionale, also $\varepsilon_1\varepsilon_2$ eine Einheit, und ebenso sind $\varepsilon_1:\varepsilon_2$ und $\varepsilon_2:\varepsilon_1$ ganz.

10. Ist ein ganzes Funktional μ teilbar durch ein anderes, α, so ist jedes mit μ assoziierte Funktional μ' auch durch jedes mit α assoziierte Funktional α' teilbar.

Denn wenn $\mu:\alpha$ ganz und ε, ε_1 Einheiten sind, so ist auch $\mu\varepsilon:\alpha\varepsilon_1$ ganz.

11. Ist α assoziiert mit β und mit γ, so sind auch β und γ untereinander assoziiert.

Denn ist $\beta = \alpha\varepsilon$, $\gamma = \alpha\varepsilon_1$, so ist $\beta = \gamma\varepsilon:\varepsilon_1$, und $\varepsilon:\varepsilon_1$ ist nach 9. eine Einheit.

Ist α teilbar durch β, so ist die absolute Norm von α teilbar durch die absolute Norm von β, denn aus $\alpha = \beta\gamma$ folgt nach § 96, 5.:

$$N_a(\alpha) = N_a(\beta)\,N_a(\gamma). \tag{1}$$

Daraus ergibt sich weiter, daß die absolute Norm einer Einheit $= 1$ sein muß. Es gilt aber auch das Umgekehrte:

12. Ein ganzes Funktional, dessen absolute Norm $= 1$ ist, ist eine Einheit.

Denn ist ε ein ganzes Funktional mit der absoluten Norm 1, so ist $N(\varepsilon)$ ein ganzes rationales Funktional mit dem absoluten Werte 1, und daher ist auch $1:N(\varepsilon)$ ein ganzes Funktional. Setzen wir dann

$$N(\varepsilon) = \varepsilon\varepsilon',$$

so ist ε' als Produkt von ganzen Funktionalen (den Konjugierten zu ε) selbst ganz, und folglich ist auch

$$\frac{1}{\varepsilon} = \frac{\varepsilon'}{N(\varepsilon)}$$

ein ganzes Funktional, also ε eine Einheit.

Daraus schließen wir noch nach der Formel (1), daß assoziierte Funktionale dieselbe absolute Norm haben.

Ein ganzes rationales Funktional ist hiernach immer mit seinem absoluten Werte assoziiert. Ist also α irgend ein ganzes Funktional des Körpers $\overline{\Omega}$, so sind auch $N(\alpha)$ und $N_a(\alpha)$ assoziiert. Da nun in $N(\alpha) = \alpha\alpha'$ der Faktor α' als Produkt von ganzen Funktionalen selbst ganz ist, so ist $N(\alpha)$ und folglich auch die natürliche Zahl $N_a(\alpha)$ durch α teilbar. Wir formulieren also noch den Satz:

13. Es gibt natürliche ganze Zahlen (in unendlicher Menge), die durch ein beliebiges ganzes Funktional α teilbar sind; darunter ist die absolute Norm von α. Ist a die kleinste unter den durch α teilbaren natürlichen Zahlen, so ist jede durch α teilbare ganze rationale Zahl durch a teilbar.

Denn ist m eine durch α teilbare ganze rationale Zahl, so ist auch der Rest der Division von m durch a eine durch α teilbare, ganze rationale Zahl. Da diese kleiner als a ist, so muß sie $= 0$ sein.

§ 99.

Größter gemeinschaftlicher Teiler.

Es mögen $\alpha, \beta \ldots$ von Null verschiedene ganze Funktionale in Ω in beliebiger Anzahl bedeuten und $x, y \ldots$ Variable, die in $\alpha, \beta \ldots$ nicht vorkommen. Dann ist

$$\delta = \alpha x + \beta y + \cdots \tag{1}$$

gleichfalls ein ganzes Funktional in Ω. Die Norm von δ ist eine ganze homogene Funktion $n$ten Grades der Variablen $x, y \ldots$, deren Koeffizienten ganze rationale Funktionale sind. Bezeichnen wir die absolute Norm von δ mit D, so ist

$$N(\delta) = DE(x, y \ldots) = DE, \tag{2}$$

und darin ist $E(x, y \ldots) = E$ eine ganze rationale Funktion

der Variablen $x, y \ldots$ und im allgemeinen eine gebrochene Funktion der in $\alpha, \beta \ldots$ vorkommenden Variablen, jedenfalls aber eine funktionale Einheit in R [§ 96, (2)].

Wir wollen jetzt unter $x_0, y_0 \ldots$ irgend welche ganze rationale Zahlen verstehen. Dann ist auch

$$\delta_0 = \alpha x_0 + \beta y_0 + \cdots$$

ein ganzes Funktional, und wir wollen beweisen, daß δ_0 durch δ teilbar ist.

Wir brauchen zu diesem Zwecke nur das ganze Funktional

$$\delta t - \delta_0 = \alpha(xt - x_0) + \beta(yt - y_0) + \cdots$$

zu betrachten, worin t eine neue Variable bedeutet. Dann ist [nach (2)]:

$$N(\delta t - \delta_0) = DE(xt - x_0, yt - y_0 \ldots), \tag{3}$$

oder, indem wir nach absteigenden Potenzen von t ordnen:

$$N(\delta t - \delta_0) = D(t^n E + t^{n-1} E_1 + t^{n-2} E_2 + \cdots), \tag{4}$$

worin $E_1, E_2 \ldots$ nach § 97, 7., 8. ganze rationale Funktionale sind. Setzen wir nun

$$C_1 = \frac{E_1}{E}, \qquad C_2 = \frac{E_2}{E} \cdots,$$

so sind, da E eine Einheit ist, auch $C_1, C_2 \ldots$ ganze rationale Funktionale, und es folgt, wenn wir noch

$$\frac{\delta_0}{\delta} = \eta$$

setzen, aus (4):

$$N(\delta)\, N(t - \eta) = DE(t^n + C_1 t^{n-1} + C_2 t^{n-2} + \cdots),$$

oder wegen (2):

$$N(t - \eta) = t^n + C_1 t^{n-1} + C_2 t^{n-2} + \cdots \tag{5}$$

Da diese Funktion nun verschwindet, wenn $t = \eta$ gesetzt wird, so folgt, daß η ein ganzes Funktional ist (§ 97, 2.), und damit ist bewiesen, daß δ_0 durch δ teilbar ist.

Da $x_0, y_0 \ldots$ beliebige ganze rationale Zahlen bedeuten können, so schließen wir daraus, daß die Funktionale $\alpha, \beta \ldots$ selbst durch δ teilbar sind, daß also δ ein gemeinsamer Teiler der Funktionale $\alpha, \beta, \ldots$ ist.

Andererseits ist aber auch (nach § 98, 4.) δ durch jeden gemeinsamen Teiler von $\alpha, \beta \ldots$ teilbar, und δ hat also die charakteristischen Eigenschaften des größten gemeinschaft-

lichen Teilers von $\alpha, \beta \ldots$ Dieselben Eigenschaften kommen aber nach § 98, 10. jedem mit δ assoziierten Funktional zu, und ebenso sind auch zwei Funktionale δ, δ' mit der doppelten Eigenschaft, daß δ und δ' durch jeden Teiler von $\alpha, \beta \ldots$ teilbar sind, und daß δ und δ' Teiler von $\alpha, \beta \ldots$ sind, durcheinander teilbar, also assoziiert, und wir stellen also die Definition auf:

1. Das Funktional $\delta = \alpha x + \beta y + \cdots$ und jedes damit assoziierte Funktional heißt größter gemeinschaftlicher Teiler von $\alpha, \beta \ldots$

Wenn δ eine Einheit ist, so sagen wir auch, $\alpha, \beta \ldots$ seien ohne gemeinsamen Teiler; denn dann gibt es außer den Einheiten kein ganzes Funktional, das in allen $\alpha, \beta \ldots$ aufgeht.

2. Zwei Funktionale α, β, die keinen gemeinsamen Teiler haben, für die also $\alpha x + \beta y$ eine Einheit ist, heißen relativ prim oder teilerfremd.

Aus diesen Definitionen ergeben sich sehr einfach folgende Sätze:

3. Wenn ganze Funktionale $\xi, \eta \ldots$ in Ω existieren, derart, daß
$$\alpha\xi + \beta\eta + \cdots = \varepsilon$$
eine Einheit ist, so sind die ganzen Funktionale $\alpha, \beta \ldots$ ohne gemeinsamen Teiler.

Denn haben $\alpha, \beta \ldots$ einen gemeinsamen Teiler δ, so ist (§ 98, 4.) δ auch Teiler von $\alpha\xi + \beta\eta + \cdots$, und δ muß also auch eine Einheit sein.

4. Sind α, β, γ drei ganze Funktionale und ist α teilerfremd zu β und zu γ, so ist α auch teilerfremd zu $\beta\gamma$.

Denn nach Voraussetzung sind, wenn x, y, u, v vier Variable sind, die in α, β, γ nicht vorkommen,
$$\varepsilon = \alpha x + \beta y, \qquad \varepsilon_1 = \alpha u + \gamma v$$
Einheiten. Demnach ist auch
$$\alpha(\alpha u x + \gamma v x + \beta u y) + \beta\gamma v y = \varepsilon\varepsilon_1$$
eine Einheit. Da aber $\alpha u x + \gamma v x + \beta u y$ und $v y$ ganz sind, so folgt nach dem Satze 3., daß α teilerfremd zu $\beta\gamma$ ist.

Hieran schließt sich der Beweis des folgenden sehr wichtigen Satzes:

5. Sind α, β, μ ganze Funktionale, α teilerfremd zu β und $\alpha\mu$ durch β teilbar, so ist μ durch β teilbar.

Denn nach der Voraussetzung über α, β ist

$$\alpha x + \beta y = \varepsilon$$

eine Einheit. Durch Multiplikation mit μ folgt daraus:

$$\alpha\mu x + \beta\mu y = \varepsilon\mu,$$

und da $\alpha\mu$ und $\beta\mu$ nach Voraussetzung durch β teilbar sind, so ist auch $\varepsilon\mu$ und folglich auch das mit $\varepsilon\mu$ assoziierte μ durch β teilbar.

§ 100.

Primfunktionale im Körper Ω.

Durch die Sätze des vorigen Paragraphen haben wir die Hilfsmittel gewonnen, um die Gesetze der Teilbarkeit der ganzen Funktionale im Körper Ω genau auf demselben Wege abzuleiten, den man in den Elementen der Arithmetik auf die natürlichen ganzen Zahlen anwendet. Wir definieren folgendermaßen:

1. Ein ganzes Funktional π des Körpers Ω, welches keine Einheit ist, heißt ein Primfunktional, wenn es außer durch die Einheiten nur noch durch die mit ihm selbst assoziierten Funktionale teilbar ist. Jedes ganze Funktional in Ω, das außer diesen noch andere Teiler hat, heißt zusammengesetzt.

Der Begriff des Primfunktionals ist hiernach wesentlich von dem Körper Ω abhängig. Es können sehr wohl die Primfunktionale eines Körpers in einem anderen erweiterten Körper zusammengesetzt sein; daß es in jedem Körper überhaupt Primfunktionale gibt, wird erst weiter unten (unter 4.) bewiesen.

Wenn ω ein beliebiges ganzes und π ein Primfunktional des Körpers Ω ist, so sind nur zwei Fälle möglich: ω ist entweder teilerfremd zu π oder durch π teilbar; denn ein gemeinschaftlicher Teiler von ω und π kann nur entweder eine Einheit oder mit π assoziiert sein, und im letzteren Falle ist ω durch π teilbar. Daraus ergibt sich der Satz:

2. Wenn das Produkt $\alpha\beta$ zweier ganzer Funktionale α, β in Ω durch ein Primfunktional π teilbar ist, so muß einer der beiden Faktoren durch π teilbar sein.

Denn wenn α und β beide nicht durch π teilbar, also beide teilerfremd zu π sind, so ist nach § 99, 4. auch $\alpha\beta$ teilerfremd zu π.

Es ergibt sich daraus durch wiederholte Anwendung, daß, wenn ein Produkt aus mehreren Faktoren durch π teilbar ist, mindestens einer der Faktoren durch π teilbar sein muß.

3. Die kleinste natürliche ganze Zahl p, die durch ein Primfunktional π in Ω teilbar ist, ist eine natürliche Primzahl, und die absolute Norm von π ist eine Potenz von p. Jede durch π teilbare ganze rationale Zahl ist auch durch p teilbar.

Nach § 98, 13. gibt es natürliche Zahlen, die durch π teilbar sind. Die kleinste unter ihnen, p, kann nicht in zwei natürliche Faktoren, die größer als 1 sind, zerlegbar sein; denn ist $p = p_1 p_2$, so muß entweder p_1 oder p_2 durch π teilbar sein (nach 2.). Ist aber keiner der Faktoren p_1, p_2 gleich 1, so sind sie beide kleiner als p, was der Voraussetzung über p widerspricht. Folglich ist p eine natürliche Primzahl. Daß jede durch π teilbare natürliche Zahl m durch p teilbar ist, haben wir schon in § 98, 13. bewiesen.

Setzen wir nun

$$p = \pi\omega, \tag{1}$$

so ist ω ein ganzes Funktional, und wenn wir beiderseits die absoluten Normen nehmen, so folgt, da die absolute Norm einer natürlichen Zahl die $n$te Potenz dieser Zahl ist:

$$p^n = N_a(\pi)\, N_a(\omega). \tag{2}$$

Hieraus folgt der zweite Teil unseres Satzes, daß die natürliche Zahl $N_a(\pi)$ eine Potenz von p ist.

Setzen wir demnach

$$N_a(\pi) = p^f, \tag{3}$$

so ist f eine Zahl, die nur einen der Werte 1, 2, 3 ... n haben kann, wenn n den Grad des Körpers Ω bedeutet.

Die Zahl f heißt der Grad des Primfunktionals π.

4. Jedes von Null verschiedene ganze Funktional ω des Körpers Ω, das keine Einheit ist, ist durch ein Primfunktional teilbar.

Wenn ω prim ist, so ist der Satz evident, weil ω durch sich selbst teilbar ist. Wenn aber ω nicht prim ist, so ist es durch

ein ganzes Funktional ω_1 teilbar, das weder eine Einheit, noch mit ω assoziiert ist. Ist also

$$\omega = \omega_1 \alpha, \tag{4}$$

so ist weder ω_1 noch α eine Einheit. Daraus folgt aber:

$$N_a(\omega) = N_a(\omega_1)\, N_a(\alpha),$$

und da $N_a(\alpha)$ größer als 1 ist, so ist

$$N_a(\omega_1) < N_a(\omega).$$

Wenn ω_1 noch nicht prim ist, so kann man denselben Schluß auf ω_1 anwenden und findet einen Teiler ω_2 von ω_1 derart, daß

$$N_a(\omega_2) < N_a(\omega_1) < N_a(\omega)$$

ist. Da es aber nur eine endliche Anzahl natürlicher Zahlen gibt, die kleiner sind als $N_a(\omega)$, so muß die Reihe der Funktionale ω, ω_1, ω_2 ... abbrechen, und dies ist nur möglich, wenn das letzte von ihnen ein Primfunktional ist, wodurch der Satz 4. bewiesen ist.

5. Jedes ganze von Null und von den Einheiten verschiedene Funktional ω im Körper Ω kann in eine endliche Anzahl von Primfaktoren zerlegt werden.

Ist nämlich π_1 ein Primfaktor von ω und

$$\omega = \pi_1 \omega_1,$$

so ist, da $N_a(\pi_1) > 1$ ist,

$$N_a(\omega) > N_a(\omega_1).$$

Ist ω_1 keine Einheit, so ist es durch ein Primfunktional π_2 teilbar, und aus

$$\omega_1 = \pi_2 \omega_2$$

folgt:

$$N_a(\omega_1) > N_a(\omega_2).$$

Fährt man so fort, so erhält man eine Reihe ganzer Funktionale ω_1, ω_2 ..., deren absolute Normen fortwährend abnehmen, und diese Reihe bricht also mit einer Einheit ab. Ist π_ν die letzte von ihnen, die keine Einheit ist, so ist π_ν selbst ein Primfunktional, und es folgt

$$\omega = \pi_1 \pi_2 \ldots \pi_\nu, \tag{5}$$

was zu beweisen war.

6. Ein ganzes Funktional ω im Körper Ω ist nur auf eine Weise in Primfaktoren zerlegbar, wenn

assoziierte Primfaktoren als nicht verschieden betrachtet werden. Assoziierte Funktionale enthalten dieselben Primfaktoren.

Nehmen wir nämlich an, es seien die beiden Produkte von Primfaktoren

$$\pi_1 \pi_2 \dots \pi_\nu, \qquad \varkappa_1 \varkappa_2 \dots \varkappa_\mu \tag{6}$$

miteinander assoziiert, so ist das Produkt $\pi_1 \pi_2 \dots \pi_\nu$ durch den Primfaktor $\varkappa_1$ teilbar, und es muß daher, nach 2., einer der Faktoren, etwa π_1, durch $\varkappa_1$ teilbar und folglich mit $\varkappa_1$ assoziiert sein. Dann sind auch die Produkte

$$\pi_2 \dots \pi_\nu, \qquad \varkappa_2 \dots \varkappa_\mu$$

assoziiert, und folglich ist einer der Faktoren des ersten Produktes, etwa π_2, durch $\varkappa_2$ teilbar und daher mit $\varkappa_2$ assoziiert. So kann man weiter schließen, und es ergibt sich, daß nicht nur die Anzahl der $\varkappa$ mit der Anzahl der π übereinstimmen muß sondern daß auch die $\varkappa$ einzeln den π assoziiert sind.

Unter den Primfaktoren eines ganzen Funktionals ω kann derselbe mehrmals vorkommen, und diese einander gleichen (oder assoziierten) Faktoren können zu einer Potenz zusammengefaßt werden. Ist ω durch π^h teilbar, so sagen wir, das Primfunktional π geht $h$mal in ω auf.

Hiernach hat es einen ganz bestimmten Sinn, wenn von den Primfaktoren eines ganzen Funktionals gesprochen wird. Wir folgern noch aus dem Bewiesenen:

7. Ein ganzes Funktional α ist dann und nur dann durch ein anderes β teilbar, wenn alle Primfaktoren von β unter den Primfaktoren von α vorkommen, und jeder von ihnen mindestens so oft in α aufgeht als in β.

Denn ist α durch β und β durch π^h teilbar, so ist auch α durch π^h teilbar.

Sind $\alpha, \beta \dots$ ganze Funktionale, $\pi_1, \pi_2 \dots$ verschiedene Primfunktionale, $\varepsilon_1, \varepsilon_2 \dots$ Einheiten, so können wir Exponenten $a_1, a_2 \dots, b_1, b_2 \dots$ so bestimmen, daß

$$\begin{aligned} \varepsilon_1 \alpha &= \pi_1^{a_1} \pi_2^{a_2} \dots \\ \varepsilon_2 \beta &= \pi_1^{b_1} \pi_2^{b_2} \dots \\ &\dots\dots\dots\dots \end{aligned} \tag{7}$$

wird, falls wir den Exponenten gleich Null setzen, wenn einer der Primfaktoren in dem betreffenden Funktional nicht aufgeht.

Ein gemeinsamer Teiler der Zahlen α, β ... kann keine anderen Primfaktoren als π_1, π_2 ... enthalten, und jeder gemeinsame Teiler von α, β ... hat die Form

$$\varepsilon\delta = \pi_1^a \pi_2^b \ldots, \tag{8}$$

worin a nicht größer als die kleinste der Zahlen a_1, b_1 ... sein darf, b nicht größer als die kleinste der Zahlen a_2, b_2 .... usf.

Ist a die kleinste unter den Zahlen a_1, b_1 ..., b die kleinste unter den Zahlen a_2, b_2 ..., so ist die in (8) dargestellte Zahl δ der größte gemeinschaftliche Teiler der Zahlen α, β ... In Worten ausgedrückt:

Man erhält den größten gemeinschaftlichen Teiler mehrerer ganzer Funktionale α, β ..., wenn man ein Produkt aus Primfaktoren bildet, in das man jeden Primfaktor so oft aufnimmt, als er in jeder der Zahlen α, β ... aufgeht. Zwei Zahlen oder Funktionale sind teilerfremd, wenn sie keinen gemeinschaftlichen Primfaktor enthalten.

Dementsprechend definieren wir als das kleinste gemeinschaftliche Multiplum μ der Funktionale α, β ... ein Produkt aus Primfaktoren, in das wir einen Faktor π nur so oft aufnehmen, daß er in keinem der Funktionale α, β ... öfter als in μ aufgeht. Dieses Funktional μ hat dann, ebenso wie jedes mit μ assoziierte Funktional, die doppelte und, wie wir hinzufügen können, charakteristische Eigenschaft, daß es durch jedes der Funktionale α, β ... teilbar ist, und daß jedes andere Funktional, das durch α, β ... teilbar ist, auch durch μ teilbar ist.

Das kleinste gemeinschaftliche Multiplum zweier teilerfremder Funktionale ist ihr Produkt.

Man kann diese Sätze anwenden, um gebrochene Funktionale in der einfachsten Gestalt oder als reduzierte Brüche darzustellen, indem man Zähler und Nenner in ihre Primfaktoren zerlegt und den größten gemeinschaftlichen Teiler weghebt. Zähler und Nenner eines reduzierten Bruches sind durch den Bruch selbst völlig bestimmt, abgesehen von einem gemeinschaftlichen Einheitsfaktor, der unbestimmt bleibt.

Ebenso kann man eine beliebige Zahl gegebener Brüche auf gemeinsamen Nenner, den Hauptnenner, bringen, indem man ein gemeinsames Multiplum aller gegebenen Nenner als gemeinschaftlichen Nenner wählt.

Alles das ist in vollkommener Übereinstimmung mit den Regeln der elementaren Arithmetik, und auch die Beweismethoden, die wir hier angewandt haben, sind wesentlich dieselben, die dort gebraucht werden. Der Kernpunkt der Deduktion ist einerseits die weitgehende Verallgemeinerung des Begriffes der Einheit, andererseits die darauf gegründete Definition des größten gemeinschaftlichen Teilers im § 99.

Diese Sätze genügen, um den Gaußschen Satz über ganze Funktionen (§ 20) auf Funktionale auszudehnen.

8. Wenn zwei ganze Funktionen einer Variablen t der Grade h und k:

$$\alpha = \alpha_0 t^h + \alpha_1 t^{h-1} + \cdots + \alpha_h,$$
$$\beta = \beta_0 t^k + \beta_1 t^{k-1} + \cdots + \beta_k,$$

deren Koeffizienten ganze Funktionale sind, die die Variable t nicht enthalten, ein Produkt

$$\gamma = \gamma_0 t^{h+k} + \gamma_1 t^{h+k-1} + \cdots + \gamma_{h+k}$$

haben, in dem die Koeffizienten $\gamma_0, \gamma_1, \ldots \gamma_{h+k}$ einen gemeinschaftlichen Primteiler π haben, so muß π entweder in allen Koeffizienten von α oder in allen Koeffizienten von β aufgehen.

Der Beweis ist genau derselbe wie in § 20. Aus diesem Satze ziehen wir hier eine wichtige Folgerung:

Die Funktionen

$$\varphi = \varphi_0 t^h + \varphi_1 t^{h-1} + \cdots + \varphi_h, \tag{9}$$

in denen die Koeffizienten $\varphi_0, \varphi_1, \ldots \varphi_h$ ganze oder gebrochene Funktionale in Ω sind, aber von den Variablen t frei angenommen werden, gehören selbst zu den Funktionalen im Körper Ω. Von ihnen gilt der Satz:

9. Ein Funktional φ ist nur dann ganz, wenn die Koeffizienten $\varphi_0, \varphi_1, \ldots \varphi_h$ ganze Funktionale sind.

Um ihn zu beweisen, nehmen wir an, es seien $\varphi_0, \varphi_1, \ldots \varphi_h$ nicht alle zugleich ganz. Bestimmen wir ihren Hauptnenner μ und setzen

$$\mu\varphi_0 = \alpha_0, \quad \mu\varphi_1 = \alpha_1, \ldots, \quad \mu\varphi_h = \alpha_h,$$

so sind $\alpha_0, \alpha_1, \ldots \alpha_h$ ganze Funktionale, und μ enthält wenigstens einen Primfaktor π, der nicht zugleich in allen Zählern $\alpha_0, \alpha_1, \ldots \alpha_h$ aufgeht.

Dann ist die Funktion

$$\chi = \mu\varphi = \alpha_0 t^h + \alpha_1 t^{h-1} + \cdots + \alpha_h, \tag{10}$$

deren Koeffizienten ganz sind, gewiß ein ganzes Funktional.

Nehmen wir nun an, es sei φ selbst ganz, so ist $\mu\varphi$ durch π teilbar, während doch nicht sämtliche Koeffizienten $\alpha_0, \alpha_1, \ldots \alpha_h$ durch π teilbar sind. Wenn nun φ ein ganzes Funktional ist, so genügt es einer Gleichung von der Form

$$\varphi^m = C_1\varphi^{m-1} + C_2\varphi^{m-2} + \cdots + C_m, \tag{11}$$

in der die Koeffizienten $C_1, C_2, \ldots C_m$ ganze rationale Funktionale sind.

Diese Funktionale setzen wir nach § 96, (2) in die Form

$$C_1 = \frac{a_1 E_1}{E}, \quad C_2 = \frac{a_2 E_2}{E} \cdots, \quad C_m = \frac{a_m E_m}{E},$$

worin $a_1, a_2, \ldots a_m$ die absoluten Werte von $C_1, C_2, \ldots C_m$, also natürliche ganze Zahlen (oder Null), und $E, E_1, \ldots E_m$ primitive ganze Funktionen, also Einheiten sind.

Dann ergibt die Gleichung (11):

$$E\varphi^m = a_1 E_1 \varphi^{m-1} + a_2 E_2 \varphi^{m-2} + \cdots + a_m E_m,$$

und durch Multiplikation mit μ^m:

$$E\chi^m = \mu(a_1 E_1 \chi^{m-1} + a_2 E_2 \mu\chi^{m-2} + \cdots + a_m E_m \mu^{m-1}). \tag{12}$$

Hier stehen nun rechter und linker Hand ganze Funktionen von t, deren Koeffizienten ganze Funktionale sind. Auf der rechten Seite haben alle diese Koeffizienten den Faktor μ, also auch den Faktor π, während nach Voraussetzung nicht alle Koeffizienten von χ diesen Faktor haben. Da E eine Einheit ist, so enthalten auch die Koeffizienten von E den Faktor π nicht, und folglich können nach dem Satze 8. auch die Koeffizienten von $E\chi^m$ nicht alle durch π teilbar sein, was doch die Gleichung (12) verlangen würde. Daraus ergibt sich, daß unsere Annahme, φ sei ganz, $\varphi_0, \varphi_1 \ldots \varphi_h$ dagegen nicht alle ganz unstatthaft ist, und der Satz 9. ist bewiesen.

Nehmen wir nun an, in φ seien die Koeffizienten $\varphi_0, \varphi_1 \ldots$ selbst wieder ganze Funktionen einer Variablen, und wenden den Satz 9 wiederholt darauf an, so gelangen wir zu dem Schlusse:

10. Eine ganze rationale Funktion beliebig vieler Veränderlicher, deren Koeffizienten Zahlen oder Funktionale mit anderen Variablen sind, ist nur dann ein ganzes Funktional, wenn die Koeffizienten ganz sind.

Wir können jedes Funktional ω als Quotienten zweier ganzer Funktionen in der Weise darstellen, daß der Nenner eine primitive Funktion im Körper R wird.

Denn sind φ, ψ ganze Funktionen in Ω, und ist

$$\omega = \frac{\varphi}{\psi}, \quad N(\psi) = \psi\,\psi',$$

so können wir den Bruch ω durch ψ' erweitern und erhalten im Nenner $N(\psi)$ eine ganze Funktion mit rationalen Koeffizienten, die wir durch Erweiterung mit dem Hauptnenner auch ganz annehmen können. Den Teiler dieser Funktion können wir dann zum Zähler von ω rechnen, und erhalten, wenn E eine primitive Funktion, χ eine ganze Funktion in Ω bedeutet:

$$E\omega = \chi,$$

d. h. man kann jedes Funktional ω in Ω durch Multiplikation mit einer primitiven Funktion E in eine ganze Funktion der Variablen verwandeln, deren Koeffizienten Zahlen in Ω sind.

Ist ω ein ganzes Funktional, so sind die Koeffizienten von χ nach 10. ebenfalls ganz. Also haben wir:

11. Jedes ganze Funktional ω im Körper Ω ist assoziiert mit einer ganzen Funktion χ, deren Koeffizienten ganze Zahlen in Ω sind.

§ 101.

Funktionale und Zahlen in Ω.

1. Ist ω ein beliebiges ganzes Funktional, so kann man eine durch ω teilbare ganze Zahl α so wählen, daß der Quotient $\alpha : \omega$ zu einem beliebig gegebenen Funktional μ teilerfremd ist.

Wir beweisen zunächst, daß es eine ganze Zahl α gibt, die durch ω, aber nicht durch $\omega\pi$ teilbar ist, wenn π ein beliebiges Primfunktional ist.

Bilden wir nämlich nach § 100, 11. eine mit ω assoziierte ganze Funktion φ, so sind die Koeffizienten dieser Funktion zwar

alle durch ω, aber nicht alle durch $\omega\pi$ teilbar, weil sonst auch φ und mithin ω selbst durch $\omega\pi$ teilbar wäre, was nicht möglich ist. Es gibt also unter den Koeffizienten von φ wenigstens einen, der die verlangte Eigenschaft hat.

Es sei jetzt $\pi_1, \pi_2, \pi_3 \ldots$ eine beliebige Anzahl voneinander verschiedener gegebener Primfunktionale. Wir setzen

$$\omega_1 = \omega\pi_2\pi_3 \ldots, \quad \omega_2 = \omega\pi_1\pi_3 \ldots, \quad \omega_3 = \omega\pi_1\pi_2 \ldots, \ldots$$

und bestimmen nach dem, was soeben bewiesen ist, die ganzen Zahlen $\alpha_1, \alpha_2, \alpha_3 \ldots$ in Ω, so daß

$$\begin{aligned}
&\alpha_1 \text{ teilbar wird durch } \omega_1, \text{ aber nicht durch } \omega_1\pi_1,\\
&\alpha_2 \quad \text{„} \quad \text{„} \quad \text{„} \quad \omega_2, \quad \text{„} \quad \text{„} \quad \text{„} \quad \omega_2\pi_2,\\
&\alpha_3 \quad \text{„} \quad \text{„} \quad \text{„} \quad \omega_3, \quad \text{„} \quad \text{„} \quad \text{„} \quad \omega_3\pi_3,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,
\end{aligned}$$

und leiten daraus die ganze Zahl

$$\alpha = \alpha_1 + \alpha_2 + \alpha_3 + \cdots$$

ab. Diese Zahl ist offenbar teilbar durch ω, da alle Summanden $\alpha_1, \alpha_2, \alpha_3 \ldots$ durch ω teilbar sind. Sie ist aber nicht teilbar durch $\omega\pi_1$, weil zwar $\alpha_2, \alpha_3 \ldots$, nicht aber α_1 durch $\omega\pi_1$ teilbar ist; und ebenso ist sie nicht durch $\omega\pi_2$, $\omega\pi_3 \ldots$ teilbar. Wenn wir also

$$\alpha = \omega\eta$$

setzen, so ist η ein ganzes Funktional, das nicht durch $\pi_1, \pi_2, \pi_3 \ldots$ teilbar ist, und das daher, wenn $\pi_1, \pi_2, \pi_3 \ldots$ die voneinander verschiedenen Primfaktoren von μ sind, teilerfremd zu μ ist.

Nehmen wir nun beliebig eine durch ω teilbare ganze Zahl β in Ω an und setzen $\beta = \omega\mu$, dann können wir $\alpha = \omega\eta$ so bestimmen, daß η relativ prim zu μ wird, und dann ist ω der größte gemeinschaftliche Teiler von α und β. Daraus folgt:

2. Jedes ganze Funktional ω des Körpers Ω ist der größte gemeinschaftliche Teiler zweier ganzer Zahlen, und folglich ist ω assoziiert mit einer binären Linearform $\alpha x + \beta y$, in der α und β ganze Zahlen sind.

Die Zahl α kann auf unendlich viele Arten bestimmt und daher auch noch anderen Bedingungen unterworfen werden. So kann z. B. α so gewählt werden, daß die mit α konjugierten Zahlen alle voneinander verschieden sind, d. h. so, daß α eine primitive Zahl des Körpers wird. Dies erreicht man dadurch, daß man α in 2. durch $\alpha + x\xi$ ersetzt, worin ξ eine primitive

ganze Zahl des Körpers ist und x eine durch ω teilbare ganze rationale Zahl ist, die man so bestimmen kann, daß die n Werte von $\alpha + x\xi$ alle voneinander verschieden ausfallen.

Wenn man die Variablen $x, y, z \ldots$ eines Funktionals ω, ohne die Zahlen zu verändern, durch andere, davon unabhängige Variable $x', y', z' \ldots$ ersetzt, so entsteht ein Funktional ω', das ebenfalls in Ω enthalten ist. Jede Gleichung zwischen mehreren der Funktionale ω ist dann auch richtig für die entsprechenden Funktionale ω', da es sich ja um Identitäten handelt. Ist also ω ganz, so ist auch ω' ganz. Sind ω, η ganze Funktionale und ω durch η teilbar, so ist auch ω' durch η' teilbar. Ist π ein Primfunktional, so ist auch π' prim. Sind ω, η teilerfremd, so sind auch ω', η' teilerfremd und ist δ der größte gemeinschaftliche Teiler von ω, η, so ist δ' der größte gemeinschaftliche Teiler von ω' und η'.

Nach 2. können wir zwei ganze Zahlen α und β bestimmen, deren größter gemeinschaftlicher Teiler ein beliebiges Funktional ω ist. Ersetzt man die Variablen $x, y, z \ldots$ durch andere $x' y', z' \ldots$, so ändern sich die Zahlen α, β nicht, die überhaupt keine Variablen enthalten und folglich ist ω' gleichfalls der größte gemeinschaftliche Teiler von α und β, also ω' assoziiert mit ω.

Damit ist bewiesen:

3. Ersetzt man die Variablen $x, y, z \ldots$ durch andere Variable $x', y', z' \ldots$, so geht jedes Funktional in Ω in ein assoziiertes Funktional über.

Ist φ eine ganze Funktion der Variablen $x, y, z \ldots$ mit ganzzahligen Koeffizienten, so ist hiernach $\varphi : \varphi'$ ein ganzes Funktional und die Koeffizienten von φ sind also nach § 100, 10. durch φ' und folglich auch durch φ teilbar. Ebenso ist φ durch jeden Teiler aller Koeffizienten von φ teilbar, und daraus folgt:

4. Eine ganze Funktion φ mit ganzen Zahlen als Koeffizienten ist der größte gemeinschaftliche Teiler aller ihrer Koeffizienten.

Es mag hier noch eine allgemeine, auf ein anderes Gebiet hinübergreifende Bemerkung ihren Platz finden.

Es ist das Hauptergebnis dieses Abschnittes, daß sich die ganzen algebraischen Zahlen in einem bestimmten Körper in eindeutiger Weise in Primfaktoren zerlegen lassen, genau in der-

selben Weise, wie dies bei den ganzen rationalen Zahlen bekannt ist; freilich aber nur dadurch, daß der Inhalt des Körpers (durch Adjunktion von Variablen) vergrößert wird. Es entsteht so ein erweiterter Körper, in dem die Gesetze der Zerlegbarkeit rein gelten.

Den Ausgangspunkt der Definition bildete der Körper R der rationalen Zahlen, und wenn wir uns Rechenschaft darüber geben wollen, auf welchen Eigenschaften des Körpers R die Möglichkeit dieser Erweiterung beruht, so finden wir, daß es einerseits die Existenz der ganzen Zahlen in R, andererseits die eindeutige Zerlegbarkeit dieser ganzen Zahlen in Primfaktoren ist, die allein bei der ganzen Deduktion benutzt wurden. Wenn wir also an Stelle des Körpers R irgend einen anderen Körper treten lassen, dem diese beiden Eigenschaften zukommen, so werden wir dieselben Folgerungen ziehen können. Nehmen wir für R einen anderen algebraischen Zahlkörper, so bekommen wir freilich nichts Neues, wohl aber, wenn wir z. B. an Stelle des Körpers R den Körper der rationalen Funktionen einer Variablen setzen, dem ja die beiden fundamentalen Eigenschaften auch zukommen. So gewinnen wir einen Ausgangspunkt für die Theorie der algebraischen Funktionen einer Variablen[1]).

§ 102.

Algebraische Körper.

Es sei

$$f(\Theta) = \Theta^n + a_1\Theta^{n-1} + \cdots + a_n = 0 \tag{1}$$

die irreduzible Gleichung $n$ten Grades mit rationalen Koeffizienten $a_1, a_2, \ldots a_n$, die uns einen algebraischen Körper $\Omega = R(\Theta)$ definiert. Der Körper Ω ist dann der Inbegriff der Zahlen der Form

$$\omega = h_1 + h_2\Theta + h_3\Theta^2 + \cdots + h_n\Theta^{n-1}, \tag{2}$$

worin $h_1, h_2, \ldots h_n$ rationale Zahlen sind. Setzen wir aber für $h_1, h_2, \ldots h_n$ rationale Funktionale, so erhalten wir aus (2) die Funktionale des Körpers Ω.

Betrachten wir ein beliebiges System von n Zahlen

[1]) Vgl. Dedekind-Weber, Theorie der algebraischen Funktionen einer Veränderlichen. Crelles Journ., Bd. 92. Algebra, große Ausgabe, III. Bd., V. Buch.

(3) $\omega_r = h_{1,r} + h_{2,r}\,\Theta + \cdots + h_{n,r}\,\Theta^{n-1}, \quad r = 1, 2, \ldots n$

unter der Voraussetzung, daß die Determinante

(4) $$H = \Sigma \pm h_{1,1}\, h_{2,2} \ldots h_{n,n}$$

von Null verschieden ist, so kann wegen der Irreduzibilität von f eine Gleichung der Form

(5) $$k_1\,\omega_1 + k_2\,\omega_2 + \cdots + k_n\,\omega_n = 0,$$

in der $k_1, k_2, \ldots k_n$ rationale Zahlen sind, nur dann bestehen, wenn diese Koeffizienten alle Null sind, und dies gilt auch dann noch, wenn in der Gleichung (5) für die Koeffizienten k **rationale Funktionale** zugelassen werden.

Eliminieren wir aber aus den Gleichungen (2) und (3) die Potenzen von Θ, so ergibt sich eine Relation von der Form

(6) $$\omega = k_1\,\omega_1 + k_2\,\omega_2 + \cdots + k_n\,\omega_n.$$

Man kann also jede Zahl und jedes Funktional des Körpers Ω in der Form (6) darstellen, wenn man für $k_1, k_2, \ldots k_n$ rationale Zahlen oder Funktionale setzt. Diese Darstellung ist für ein gegebenes ω [wegen (5)] nur auf eine Art möglich, und ω ist eine Zahl, wenn die $k_1, k_2, \ldots k_n$ Zahlen sind, dagegen ein Funktional, wenn unter den k auch Funktionale vorkommen.

Ein solches System von Zahlen, wie

$$\omega_1, \omega_2, \ldots \omega_n$$

nennen wir eine **Basis des Körpers** Ω. Eine solche Basis bilden auch die Potenzen von Θ:

$$1, \Theta, \Theta^2, \ldots \Theta^{n-1}.$$

> Jedes System von n Zahlen des Körpers ω, zwischen denen keine lineare Relation mit von Null verschiedenen rationalen Koeffizienten von der Form (5) besteht, ist eine Basis von Ω.

Bezeichnen wir mit $\omega_{r,1}, \omega_{r,2}, \ldots \omega_{r,n}$ die mit ω_r konjugierten Zahlen, so ist das Determinantenquadrat

(7) $$\Delta(\omega_1, \omega_2, \ldots \omega_n) = (\Sigma \pm \omega_{1,1}\,\omega_{2,2} \ldots \omega_{n,n})^2$$

eine symmetrische Funktion der konjugierten Werte $\Theta_1, \Theta_2, \ldots \Theta_n$, und folglich eine rationale Zahl.

Diese Zahl heißt die **Diskriminante** des Systems $\omega_1, \omega_2, \ldots \omega_n$. Insbesondere ist (§ 26)

$$(8) \qquad \Delta(1, \Theta, \Theta^2, \dots \Theta^{n-1}) = \begin{vmatrix} 1, & \Theta_1 & \dots & \Theta_1^{n-1} \\ 1, & \Theta_2 & \dots & \Theta_2^{n-1} \\ \dots & \dots & \dots & \dots \\ 1, & \Theta_n & \dots & \Theta_n^{n-1} \end{vmatrix}^2$$

die Diskriminante der Gleichung (1), für die man auch, wenn N das Zeichen für die Norm ist, nach § 26 (5)

$$(9) \qquad (-1)^{\frac{n(n-1)}{2}} N f'(\Theta)$$

setzen kann. Diese Zahl ist also sicher von Null verschieden. Nach dem Multiplikationssatze der Determinanten ergibt sich aus (3) und (7):

$$(10) \qquad \Delta(\omega_1\ \omega_2 \dots \omega_n) = H^2 \Delta(1, \Theta, \Theta^2 \dots \Theta^{n-1}),$$

woraus man schließt, daß die Diskriminante einer Basis von Ω immer von Null verschieden ist. Da H eine rationale Zahl ist, so folgt, daß das Verhältnis der Diskriminanten verschiedener Basen das Quadrat einer rationalen Zahl ist, und daß die Diskriminanten aller Basen von Ω dasselbe Vorzeichen haben.

Die Formel (10) zeigt auch, daß irgend ein System von n Zahlen ω_r des Körpers Ω immer dann eine Basis von Ω ist, wenn das Determinantenquadrat (7) nicht verschwindet.

Ist $\omega_1, \omega_2, \dots \omega_n$ eine Basis von Ω, und bedeuten $c_{r,s}$ rationale Zahlen, so bilden auch die n Zahlen

$$(11) \qquad \omega_r' = c_{r,1}\omega_1 + c_{r,2}\omega_2 + \dots + c_{r,n}\omega_n, \quad r = 1, 2, \dots n$$

eine Basis von Ω, wenn die Determinante

$$(12) \qquad C = \Sigma \pm c_{1,1} c_{2,2} \dots c_{n,n}$$

nicht verschwindet, denn es ist

$$(13) \qquad \Delta(\omega_1', \omega_2', \dots \omega_n') = C^2 \Delta(\omega_1, \omega_2, \dots \omega_n).$$

Wenn wir z. B. die Elemente $\omega_1, \omega_2, \dots \omega_n$ einer Basis mit rationalen Koeffizienten $c_1, c_2 \dots c_n$ multiplizieren, deren keiner verschwindet, so erhalten wir eine neue Basis

$$c_1\omega_1, c_2\omega_2 \dots c_n\omega_n.$$

§ 103.

Die Minimalbasis und die Körperdiskriminante.

Nach der zuletzt gemachten Bemerkung verliert eine Basis von Ω die Eigenschaft, eine Basis zu sein, nicht, wenn man jede ihrer Zahlen mit einer von Null verschiedenen rationalen Zahl

multipliziert. Nun kann man nach § 97, 5. jede Zahl durch Multiplikation mit einer ganzen rationalen Zahl in eine ganze Zahl verwandeln, und daraus folgt, daß es Basen von Ω gibt, deren Elemente lauter ganze Zahlen sind. Die Diskriminante einer solchen Basis ist eine ganze rationale Zahl. Diese ganze rationale Zahl ist von Null verschieden. Sie ändert sich, wenn eine andere ganzzahlige Basis gewählt wird, behält aber für einen bestimmten Körper ein unverändertes Vorzeichen.

Unter all diesen ganzen Zahlen, die als Diskriminanten einer ganzzahligen Basis auftreten können, und die alle in quadratischem Verhältnis zueinander stehen, muß nun eine dem absoluten Werte nach die kleinste sein. Diese kleinste Diskriminante bezeichnen wir mit Δ und nennen sie die Grundzahl oder auch die Diskriminante des Körpers Ω.

Dies Δ ist eine durch Ω völlig bestimmte positive oder negative, aber niemals verschwindende ganze rationale Zahl, und es gibt immer eine aus ganzen Zahlen $\omega_1, \omega_2, \ldots \omega_n$ bestehende Basis von Ω, deren Diskriminante gleich Δ ist.

Eine solche Basis wollen wir eine Minimalbasis von Ω nennen.

Verstehen wir unter $k_1, k_2, \ldots k_n$ irgend welche ganze rationale Zahlen, so ist jede Zahl von der Gestalt

$$\omega = k_1\omega_1 + k_2\omega_2 + \cdots + k_n\omega_n \tag{1}$$

eine ganze algebraische Zahl, und wir beweisen jetzt den fundamentalen Satz:

1. Wenn $\omega_1, \omega_2, \ldots \omega_n$ eine Minimalbasis ist, so sind in der Form (1) alle ganzen Zahlen des Körpers Ω enthalten.

Da $\omega_1, \omega_2, \ldots \omega_n$ eine Basis ist, so kann zunächst jede Zahl in Ω in der Form (1) dargestellt werden, wenn für $k_1, k_2, \ldots k_n$ rationale Brüche zugelassen werden. Nehmen wir also an, es sei eine ganze Zahl ω in der Form

$$\omega = \frac{k_1\omega_1 + k_2\omega_2 + \cdots + k_n\omega_n}{k} \tag{2}$$

darstellbar, worin $k_1, k_2, \ldots k_n, k$ ganze rationale Zahlen sind, so daß nicht alle $k_1, k_2, \ldots k_n$ mit k einen gemeinschaftlichen Teiler haben. Ist p irgend eine in k aufgehende natürliche Primzahl und $k = pk'$, so muß wenigstens einer der Koeffizienten

$k_1, k_2, \ldots k_n$ durch p unteilbar sein. Es sei etwa k_1 durch p nicht teilbar; dann läßt sich die ganze rationale Zahl l so bestimmen, daß $(lk_1 - 1)$ durch p teilbar wird. Es folgt dann aus (2):

$$(3)\quad lk'\omega - \frac{lk_1 - 1}{p}\omega_1 = \frac{\omega_1 + lk_2\omega_2 + \cdots + lk_n\omega_n}{p} = \omega_1',$$

und ω_1' ist gleichfalls eine ganze algebraische Zahl. Setzen wir

$$(4)\qquad \omega_2' = \omega_2, \ldots \omega_n' = \omega_n,$$

so bilden die Zahlen $\omega_1', \omega_2', \ldots \omega_n'$ eine ganzzahlige Basis von Ω, weil sich die Zahlen $\omega_1, \omega_2, \ldots \omega_n$ und folglich alle Zahlen ω linear durch $\omega_1', \omega_2' \ldots \omega_n'$ ausdrücken lassen, und die Formel (13) des vorigen Paragraphen ergibt:

$$(5)\qquad \Delta(\omega_1', \omega_2', \ldots \omega_n') = \frac{1}{p^2}\Delta(\omega_1, \omega_2, \ldots \omega_n).$$

Die Diskriminante $\Delta(\omega_1', \omega_2', \ldots \omega_n')$ ist also kleiner als $\Delta(\omega_1, \omega_2 \ldots \omega_n)$, und dies widerspricht der Annahme, daß $\omega_1, \omega_2, \ldots \omega_n$ eine Minimalbasis sei. Damit ist unser Satz erwiesen.

Bezeichnet man das System der ganzen Zahlen des Körpers Ω mit $\mathfrak{o}$, so erhält man jede Zahl von $\mathfrak{o}$, und jede nur einmal, wenn man in

$$(6)\qquad k_1\omega_1 + k_2\omega_2 + \cdots + k_n\omega_n$$

die Koeffizienten $k_1, k_2, \ldots k_n$ die ganzen rationalen Zahlen durchlaufen läßt.

Aus diesem Grunde wird eine Minimalbasis von Ω auch eine Basis von $\mathfrak{o}$ genannt.

2. Die Diskriminante einer Basis von $\mathfrak{o}$ ist gleich der Grundzahl des Körpers Ω.

Nach der Formel (13) des vorigen Paragraphen können wir aus einer Basis von Ω beliebig viele andere durch lineare Substitution ableiten:

$$(7)\qquad (\omega_1', \omega_2', \ldots \omega_n') = C(\omega_1, \omega_2, \ldots \omega_n),$$

wenn C eine lineare Substitution mit rationalen Koeffizienten bedeutet.

Ist $(\omega_1, \omega_2, \ldots \omega_n)$ hierin eine Minimalbasis, so bilden die ω_r' eine ganzzahlige Basis, wenn die $c_{i,k}$ ganze rationale Zahlen sind, und die ω_r' bilden eine Minimalbasis, wenn die Determinante $C = \pm 1$ ist. Immer aber stehen die beiden Diskriminanten Δ', Δ in dem Verhältnis [§ 102, (13)]:

$$\Delta' = C^2 \Delta$$

und daraus können wir schließen:

3. Ist die Diskriminante einer Basis ω'_r von Ω, die aus ganzen Zahlen besteht, durch kein Quadrat teilbar, so ist ω'_r eine Minimalbasis, und Δ' ist die Grundzahl des Körpers.

Da unter den Zahlen von $\mathfrak{o}$ immer die Zahl 1 enthalten ist, so kann man, wenn $\omega_1, \omega_2, \dots \omega_n$ eine Basis von $\mathfrak{o}$ ist, die ganzen rationalen Zahlen $c_1, c_2, \dots c_n$ so bestimmen, daß die Relation

$$c_1\omega_1 + c_2\omega_2 + \cdots + c_n\omega_n = 1 \tag{8}$$

befriedigt ist.

Es gilt aber auch in bezug auf die Funktionale der Satz:

4. Wenn $\omega_1, \omega_2, \dots \omega_n$ eine Basis von $\mathfrak{o}$ ist, so sind in der Form

$$\omega = u_1\omega_1 + u_2\omega_2 + \cdots + u_n\omega_n, \tag{9}$$

in der $u_1, u_2, \dots u_n$ ganze rationale Funktionale sind, alle ganzen Funktionale in Ω enthalten.

Denn wir haben schon oben (§ 102) gezeigt, daß alle Funktionale überhaupt in der Form (9) enthalten sind, wenn die Koeffizienten $u_1, u_2, \dots u_n$ ganze oder gebrochene rationale Funktionale sind. Wir können aber immer eine ganze primitive rationale Funktion e so bestimmen, daß

$$eu_1 = y_1, \quad eu_2 = y_2 \dots eu_n = y_n$$

ganze Funktionen der Variablen sind, und dann wird

$$e\omega = y_1\omega_1 + y_2\omega_2 + \cdots + y_n\omega_n, \tag{10}$$

und da e eine Einheit ist, so ist $e\omega$ zugleich mit ω ganz. Da nun die Koeffizienten der Potenzen und Produkte der Variablen in der Funktion (10) nach § 100, 10. ganze Zahlen sein müssen, so folgt nach 1., daß die Koeffizienten in den Funktionen $y_1, y_2, \dots y_n$ ganze rationale Zahlen sein müssen, und daß folglich $u_1, u_2, \dots u_n$ ganze rationale Funktionale sind.

Ist η irgend eine Zahl oder ein Funktional des Körpers Ω, so verstehen wir unter der Diskriminante von η die Diskriminante des Systems

$$1, \eta, \eta^2, \dots \eta^{n-1},$$

oder auch, wenn $\eta_1, \eta_2, \ldots \eta_n$ die konjugierten Größen zu η sind, das Differenzenprodukt

$$\Delta(\eta) = \Pi(\eta_i - \eta_k)^2.$$

Stellt man $\Delta(\eta)$ durch eine Determinante dar, so erkennt man aus § 102, (10), angewandt auf die Potenzen von η:

5. Die Diskriminante einer ganzen Zahl oder eines ganzen Funktionals ist immer durch die Grundzahl des Körpers teilbar.

§ 104.

Basen und Normen der Funktionale.

Ist μ ein ganzes Funktional des Körpers Ω, so verstehen wir unter einer Basis von μ ein System von n ganzen Zahlen in Ω:

(1) $$\alpha_1, \alpha_2, \ldots \alpha_n,$$

das eine Basis des Körpers Ω ist, und dem die Eigenschaft zukommt, daß in der Form

(2) $$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$$

alle durch μ teilbaren ganzen Zahlen des Körpers Ω und keine anderen enthalten sind, wenn für $x_1, x_2, \ldots x_n$ ganze rationale Zahlen gesetzt werden.

Es soll jetzt bewiesen werden, daß jedes ganze Funktional eine Basis hat.

Zunächst ist klar, daß eine Basis eines Funktionals μ zugleich Basis aller mit μ assoziierten Funktionale ist, ferner, daß jedes Element $\alpha_1, \alpha_2, \ldots \alpha_n$ einer solchen Basis durch μ teilbar sein muß, endlich daß aus einer Basis von μ alle anderen abgeleitet werden können, wenn man eine lineare Substitution mit ganzen rationalen Koeffizienten und der Determinante ± 1 anwendet.

Um nun eine Basis von μ zu bilden, gehen wir von einer Minimalbasis von Ω aus, die wir, wie oben, mit

$$\omega_1, \omega_2, \ldots \omega_n$$

bezeichnen.

Da es positive ganze rationale Zahlen gibt, die durch μ teilbar sind, so gibt es auch ganze rationale Zahlen a, für die

das Produkt $a\omega_1$ durch μ teilbar ist. Die kleinste positive unter diesen Zahlen wollen wir mit $a_{1,1}$ bezeichnen, und

$$a_{1,1}\,\omega_1 = \alpha_1$$

setzen. Sodann bezeichnen wir mit $a_{2,2}$ die kleinste positive ganze rationale Zahl, für die sich ein ganzes rationales $a_{1,2}$ so bestimmen läßt, daß

$$\alpha_2 = a_{1,2}\,\omega_1 + a_{2,2}\,\omega_2$$

durch μ teilbar wird, und fahren so fort. Wir bilden auf diese Weise das System

$$\begin{aligned}
&\alpha_1 = a_{1,1}\,\omega_1,\\
&\alpha_2 = a_{1,2}\,\omega_1 + a_{2,2}\,\omega_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\alpha_n = a_{1,n}\,\omega_1 + a_{2,n}\,\omega_2 + \cdots + a_{n,n}\,\omega_n,
\end{aligned} \tag{3}$$

worin $a_{1,1}, a_{2,2}, \ldots a_{n,n}$ die kleinsten positiven ganzen rationalen Zahlen sind, die eine Bestimmung der ganzen rationalen Zahlen $a_{1,2}, \ldots a_{n-1,n}$ so ermöglichen, daß die ganzen Zahlen $\alpha_1, \alpha_2, \ldots \alpha_n$ durch μ teilbar werden.

Jede Zahl $a_{r,r}$ ist hiernach unabhängig von den übrigen dadurch definiert, daß sie die kleinste natürliche Zahl ist, für die sich die zugehörigen ganzen rationalen Zahlen

$$a_{1,r},\ a_{2,r}, \ldots a_{r-1,r}$$

so bestimmen lassen, daß

$$\alpha_r = a_{1,r}\,\omega_1 + a_{2,r}\,\omega_2 + \cdots + a_{r-1,r}\,\omega_{r-1} + a_{r,r}\,\omega_r \tag{4}$$

durch μ teilbar wird. In dem besonderen Falle, wo ω_r selbst durch μ teilbar ist, hat man demnach $a_{r,r} = 1$ zu setzen, und die $a_{1,r}, a_{2,r}, \ldots a_{r-1,r}$ können alle gleich Null angenommen werden.

Bezeichnen wir mit Δ die Grundzahl des Körpers Ω, so ergibt sich die Diskriminante des durch (3) bestimmten Systems $\alpha_1, \alpha_2, \ldots \alpha_n$ nach der Formel § 102, (13):

$$\Delta(\alpha_1, \alpha_2, \ldots \alpha_n) = a_{1,1}^2\, a_{2,2}^2 \ldots a_{n,n}^2\, \Delta. \tag{5}$$

Dies ist eine von Null verschiedene ganze rationale Zahl, und folglich sind die Größen α eine Basis von Ω.

Um also zu zeigen, daß das so bestimmte System $\alpha_1, \alpha_2, \ldots \alpha_n$ eine Basis von μ ist, bleibt noch nachzuweisen, daß jede durch μ teilbare ganze Zahl α in der Form (2) dargestellt werden kann. Nehmen wir, um diesen Beweis zu führen, irgend einen Index $r \leqq n$ an, und suchen die Bedingung dafür, daß eine ganze Zahl von der Form

(6) $$\gamma_r = h_1\omega_1 + h_2\omega_2 + \cdots + h_r\omega_r$$

durch μ teilbar ist, wenn $h_1, h_2, \ldots h_r$ ganze rationale Zahlen sind.

Zunächst folgt, daß h_r durch $a_{r,r}$ teilbar sein muß. Denn bezeichnen wir mit q_r den Rest der Division von h_r durch $a_{r,r}$ und setzen

$$h_r = l_r a_{r,r} + q_r, \quad 0 \leqq q_r < a_{r,r},$$

so ergibt sich nach (6):

$$\gamma_r - l_r\alpha_r = (h_1 - l_r a_{1,r})\omega_1 + (h_2 - l_r a_{2,r})\omega_2 + \cdots + q_r\omega_r,$$

und diese Zahl müßte auch durch μ teilbar sein. Dies ist aber nach der Definition von $a_{r,r}$ nur möglich, wenn $q_r = 0$ ist. Dann aber erhält $\gamma_r - l_r\alpha_r$ den Ausdruck:

(7) $$\gamma_r - l_r\alpha_r = h_1'\omega_1 + h_2'\omega_2 + \cdots + h_{r-1}'\omega_{r-1},$$

wird also von derselben Form wie (6), nur daß $r - 1$ an Stelle von r tritt, und in (7) ist dieselbe Schlußweise zu wiederholen. Demnach ergibt sich durch vollständige Induktion der Satz:

> Eine in der Form $h_1\omega_1 + h_2\omega_2 + \cdots + h_r\omega_r$ darstellbare ganze Zahl des Körpers Ω ist immer dann und nur dann durch μ teilbar, wenn sie in der Form
>
> $$l_1\alpha_1 + l_2\alpha_2 + \cdots + l_r\alpha_r$$
>
> darstellbar ist, in der die Koeffizienten $l_1, l_2, \ldots l_r$ ganze rationale Zahlen sind.

Setzt man in diesem Satze $r = n$, so hat man den Beweis dafür, daß das Zahlensystem (3) eine Basis von μ ist.

Wie man aus einer Basis von μ alle anderen ableiten kann, haben wir schon oben gesehen.

Wir verstehen jetzt unter $\alpha_1, \alpha_2, \ldots \alpha_n$ eine beliebige Basis von μ und stellen das Funktional μ nach § 100 als Quotienten zweier ganzer Funktionen dar, deren Nenner eine rationale Einheit ist. Die Koeffizienten des Zählers sind dann ganze durch μ teilbare Zahlen (§ 101, 4.) und können daher in der Form (2) dargestellt werden.

Fassen wir diese Darstellung gehörig zusammen, so ergibt sich also für μ ein Ausdruck

(8) $$\mu = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n,$$

worin die $u_1, u_2, \ldots u_n$ ganze rationale Funktionale sind.

Hiermit wollen wir die Linearform

(9) $$\lambda = \alpha_1 t_1 + \alpha_2 t_2 + \cdots + \alpha_n t_n$$

vergleichen, in der $t_1, t_2, \ldots t_n$ Variable sind. Die Linearform ist (nach § 99) der größte gemeinschaftliche Teiler der Zahlen $\alpha_1, \alpha_2, \ldots \alpha_n$ und ist durch μ teilbar, weil die Zahlen $\alpha_1, \alpha_2, \ldots \alpha_n$ durch μ teilbar sind. Andererseits ist aber auch, wie die Darstellung (8) zeigt, μ durch λ teilbar, und folglich sind die beiden Funktionale μ und λ miteinander assoziiert.

Das Funktional λ wollen wir eine Basisform des Funktionals μ nennen: es ist dann λ zugleich Basisform von allen mit μ assoziierten Funktionalen.

Eine Basisform λ von μ hat die Eigenschaft, daß man aus ihr alle durch μ (oder durch λ) teilbaren ganzen Zahlen in Ω erhält, wenn man für die Variablen ganze rationale Zahlen setzt.

Die Basis $\alpha_1, \alpha_2, \ldots \alpha_n$ steht zu dem Funktional μ in einer ähnlichen Beziehung, wie die Basis $\omega_1, \omega_2, \ldots \omega_n$ des Systems $\mathfrak{o}$ aller ganzen Zahlen in Ω zu den Einheiten. In der Tat ist die Linearform

$$\tau = \omega_1 t_1 + \omega_2 t_2 + \cdots + \omega_n t_n \tag{10}$$

eine Einheit; denn sie ist der größte gemeinschaftliche Teiler von $\omega_1, \omega_2, \ldots \omega_n$ und muß also, wie die Formel § 103, (8) zeigt, ein Teiler von 1, also eine Einheit sein.

Demnach wollen wir das Funktional τ eine Basisform von $\mathfrak{o}$ nennen.

Aus einer solchen Basisform erhält man alle ganzen Zahlen des Körpers Ω, wenn man für die Variablen ganze rationale Zahlen setzt.

Die Linearform τ ist die Wurzel einer irreduziblen Funktion $n$ten Grades der Variablen t:

$$F(t) = N(t - \tau) = 0,$$

in der die Koeffizienten der Potenzen von t ganze rationale (und homogene) Funktionen der Variablen $t_1, t_2, \ldots t_n$ sind.

Die Elemente $\alpha_1, \alpha_2, \ldots \alpha_n$ einer Basis des Funktionals μ können als ganze Zahlen in Ω linear und ganzzahlig ausgedrückt werden durch eine Basis $\omega_1, \omega_2, \ldots \omega_n$ von $\mathfrak{o}$ in der Form

$$(\alpha_1, \alpha_2, \ldots \alpha_n) = A\ (\omega_1, \omega_2, \ldots \omega_n), \tag{11}$$

worin A eine lineare Substitution mit ganzen rationalen Koeffizienten bedeutet. Einen Spezialfall hiervon bieten die Formeln (3).

Eine Basisform λ von μ wollen wir so darstellen:

$$\lambda = \sum_{\nu}^{\nu} \alpha_\nu t_\nu, \tag{12}$$

und die Substitution (11) schreiben wir ausführlicher:

$$\alpha_s = \sum_{\nu}^{\nu} a_{s,\nu}\, \omega_\nu, \tag{13}$$

worin t_ν Variable, $a_{s,\nu}$ die Substitutionskoeffizienten sind, und der Summationsbuchstabe ν von 1 bis n läuft.

Da die Produkte $\alpha_s \omega_r$ alle durch μ teilbar sind, so können sie nach der Bedeutung der Basis in der Form dargestellt werden:

$$\alpha_\nu \omega_r = \sum_{s}^{s} g_{\nu,r}^{(s)} \alpha_s, \tag{14}$$

worin die $g_{\nu,r}^{(s)}$ ganze rationale Zahlen sind. Hieraus erhält man dann nach (12):

$$\lambda \omega_r = \sum_{s}^{s} \alpha_s t_{s,r}, \tag{15}$$

wenn

$$t_{s,r} = \sum_{\nu}^{\nu} g_{\nu,r}^{(s)} t_\nu \tag{16}$$

ganze rationale Linearformen sind.

Substituiert man in (15) wieder die Ausdrücke (13), so folgt:

$$\lambda \omega_r = \sum_{\nu}^{\nu} \omega_\nu \sum_{s}^{s} a_{s,\nu}\, t_{s,r}. \tag{17}$$

Eliminieren wir aus diesen linearen Gleichungen die ω_ν, so können wir die Gleichung $n$ten Grades für λ in Determinantenform darstellen, und das Produkt der Wurzeln dieser Gleichung, also die Norm von λ, erhalten wir als die Determinante aus den n^2 Größen

$$\sum_{s}^{s} a_{s,\nu}\, t_{s,r}$$

[§ 97, (8)]. Diese Determinante läßt sich aber nach dem Multiplikationssatze der Determinanten zerlegen, und gibt, wenn

$$A = \Sigma \pm a_{1,1}\, a_{2,2} \ldots a_{n,n}, \quad T = \Sigma \pm t_{1,1}\, t_{2,2} \ldots t_{n,n}$$

gesetzt wird,

$$N(\lambda) = A\, T^{\cdot}. \tag{18}$$

Hierin ist T eine ganze Funktion der Variablen t_ν mit ganzen rationalen Zahlenkoeffizienten, von der wir nun noch nachweisen werden, daß sie primitiv ist.

Nehmen wir also im Gegenteil an, im Teiler von T gehe irgend eine natürliche Primzahl p auf. Dann können wir n ganze rationale Formen $y_1, y_2, \ldots y_n$ der Variablen t so bestimmen, daß die n Summen

(19) $$u_s = t_{s,1} y_1 + t_{s,2} y_2 + \cdots + t_{s,n} y_n$$

alle durch p teilbar sind, ohne daß alle $y_1, y_2, \ldots y_n$ durch p teilbar sind.

Die Richtigkeit dieser Behauptung folgt aus elementaren Determinantensätzen.

Wenn nämlich die $t_{s,r}$ alle durch p teilbar sind, so können wir die y_r ganz beliebig, z. B. gleich 1 annehmen.

Anderenfalls nehmen wir an, daß außer der Determinante T auch alle m-reihigen Unterdeterminanten durch p teilbar seien ($m \gtreqless n$), und daß unter den $(m-1)$-reihigen Unterdeterminanten wenigstens eine nicht durch p teilbar sei. Ist dann etwa unter den $(m-1)$-reihigen Determinanten der Matrix

$$\begin{matrix} t_{1,1}, & t_{1,2}, & \ldots\ t_{1,m} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ t_{m-1,1}, & t_{m-1,2}, & \ldots\ t_{m-1,m} \end{matrix}$$

eine durch p nicht teilbar, so setzen wir für $y_1, y_2, \ldots y_m$ eben diese $(m-1)$-reihigen Determinanten und nehmen $y_{m+1}, \ldots y_n = 0$ an. Diese y genügen dann der gestellten Forderung.

Sind die y so bestimmt, so setzen wir

(20) $$\omega = \omega_1 y_1 + \omega_2 y_2 + \cdots + \omega_n y_n,$$

und leiten aus (15) die Gleichung ab:

(21) $$\lambda \omega = \sum^{s} \alpha_s u_s.$$

Da nun die α_s durch λ und die u durch p teilbar sind, so folgt, daß ω durch p teilbar ist, und daß mithin nach § 103, 1., $y_1, y_2, \ldots y_n$ durch p teilbar sein müssen, was unserer Voraussetzung entgegen ist.

Damit ist bewiesen, daß T eine primitive Funktion ist, und daß also der absolute Wert der Determinante A gleich der absoluten Norm von λ und folglich auch von μ ist.

Wenden wir das Ergebnis auf die spezielle Basis (3) an, so ergibt sich die Formel:

(22) $$N_a(\mu) = a_{1,1}\, a_{2,2} \ldots a_{n,n}.$$

Hieran knüpfen wir noch folgende Bemerkungen: Wenn man in einer Basisform eines Funktionals μ

$$\lambda = \alpha_1 t_1 + \alpha_2 t_2 + \cdots + \alpha_n t_n$$

auf die Variablen $t_1, t_2 \ldots t_n$ eine ganzzahlige lineare Substitution mit der Determinante ± 1 anwendet, so entsteht eine neue Basisform von μ. Denn allen ganzzahligen rationalen Werten

der Variablen $t_1, t_2, \ldots t_n$ entsprechen ganzzahlige rationale Werte der neuen Variablen und umgekehrt.

Die Anwendung einer linearen Substitution auf die Variablen t ist aber gleichbedeutend mit der Anwendung der transponierten Substitution auf die Koeffizienten $\alpha_1, \alpha_2, \ldots \alpha_n$, d. h. mit dem Übergange zu einer neuen Basis von μ:

$$(\beta_1, \beta_2 \ldots \beta_n) = B(\alpha_1, \alpha_2 \ldots \alpha_n), \tag{23}$$

die dann durch Zusammensetzung mit (11) ergibt:

$$(\beta_1, \beta_2 \ldots \beta_n) = BA(\omega_1, \omega_2 \ldots \omega_n). \tag{24}$$

Die Diskriminante der Basis $\alpha_1, \alpha_2, \ldots \alpha_n$ von μ ist nach (13) gleich $A^2\Delta$, wenn A die absolute Norm von μ und Δ die Körperdiskriminante ist.

Wenn die Diskriminante der durch eine Substitution (23) bestimmten Zahlen β mit der Diskriminante der α übereinstimmt, so muß die Substitutionsdeterminante $B = \pm 1$ sein; und wir kommen also zu den Sätzen:

1. Die Diskriminante einer Basis von μ ist gleich dem Quadrat der absoluten Norm von μ, multipliziert mit der Grundzahl des Körpers;

und umgekehrt:

2. Ist $\beta_1, \beta_2, \ldots \beta_n$ ein System ganzer durch μ teilbarer Zahlen, dessen Diskriminante gleich ist dem Quadrat der absoluten Norm von μ, multipliziert mit der Grundzahl des Körpers, so ist $\beta_1, \beta_2, \ldots \beta_n$ eine Basis von μ.

§ 105.

Kongruenzen.

1. **Definition:** Zwei ganze algebraische Zahlen ξ, η, deren Differenz $\xi - \eta$ durch ein ganzes Funktional μ teilbar ist, heißen miteinander kongruent nach dem Modul μ.

Der Begriff der Kongruenz läßt sich auch auf Funktionale ausdehnen, was wir aber fürs erste noch nicht tun. Dagegen ist es wesentlich, als Moduln der Kongruenzen nicht bloß Zahlen, sondern auch Funktionale zu berücksichtigen. Jede Kongruenz bleibt bestehen, wenn der Modul durch ein assoziiertes Funktional

ersetzt wird. Bei dem Modul können beliebige Einheitsfaktoren hinzugefügt werden.

Wir gebrauchen für die Kongruenz das Gaußsche Zeichen

$$\xi \equiv \eta \pmod{\mu}. \tag{1}$$

Eine solche Kongruenz ist gleichbedeutend mit der Gleichung

$$\xi = \eta + \omega\mu, \tag{2}$$

worin ω irgend ein ganzes Funktional sein kann.

Aus dieser Darstellung erkennt man dann sofort, daß, ebenso wie in Kongruenzen zwischen rationalen Zahlen, wenn man in einem durch Addition, Subtraktion und Multiplikation von ganzen Zahlen zusammengesetzten Ausdruck jede Zahl durch eine nach dem Modul μ kongruente Zahl ersetzt, eine nach demselben Modul kongruente Zahl das Resultat ist; in Zeichen:

Sind $\xi_1, \eta_1, \xi_2, \eta_2 \ldots$ ganze Zahlen, die den Kongruenzen

$$\xi_1 \equiv \eta_1, \ \xi_2 \equiv \eta_2, \ \ldots \pmod{\mu}$$

genügen, und ist $\psi(x_1, x_2, \ldots)$ eine ganze Funktion mit ganzzahligen rationalen Koeffizienten, so ist auch

$$\psi(\xi_1, \xi_2, \ldots) \equiv \psi(\eta_1, \eta_2, \ldots) \pmod{\mu}. \tag{3}$$

2. Ist m die kleinste positive ganze rationale Zahl, die durch μ teilbar ist, so sind zwei nach dem Modul μ kongruente ganze rationale Zahlen auch nach dem Modul m kongruent (§ 98, 13.) und umgekehrt.

Ein Hauptsatz über die Kongruenzen, zu dessen Beweis wir jetzt schreiten, ist der:

3. Die Anzahl der nach einem Modul μ inkongruenten ganzen Zahlen eines Körpers Ω ist endlich, und zwar gleich der absoluten Norm von μ.

Um ihn zu beweisen, nehmen wir eine Basis des Funktionals μ an, und zwar wählen wir gerade die im § 104, (3) bestimmte spezielle Basis:

$$\begin{aligned} \alpha_1 &= a_{1,1}\omega_1, \\ \alpha_2 &= a_{1,2}\omega_1 + a_{2,2}\omega_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_n &= a_{1,n}\omega_1 + a_{2,n}\omega_2 + \cdots + a_{n,n}\omega_n, \end{aligned} \tag{4}$$

worin $\omega_1, \omega_2, \ldots \omega_n$ eine Basis von $\mathfrak{o}$ ist.

Jede ganze Zahl η in Ω läßt sich dann in der Form darstellen:

(5) $$\eta = y_1\omega_1 + y_2\omega_2 + \cdots + y_n\omega_n,$$

wenn die $y_1, y_2, \ldots y_n$ ganze rationale Zahlen sind. Damit verbinden wir das Zahlensystem

(6) $$\xi = x_1\omega_1 + x_2\omega_2 + \cdots + x_n\omega_n,$$

in dem wir x_1 als ganze rationale Zahl ein volles Restsystem nach dem Modul $a_{1,1}$, x_2 nach dem Modul $a_{2,2}\ldots$, x_n nach dem Modul $a_{n,n}$ durchlaufen lassen. Die Anzahl der so gebildeten Zahlen ξ beträgt $a_{1,1}\, a_{2,2} \ldots a_{n,n}$, eine Zahl, die nach § 104 gleich der absoluten Norm von μ ist.

Wir wollen nachweisen, daß jede Zahl η einer und nur einer dieser Zahlen ξ nach dem Modul μ kongruent ist.

Wenn $\eta - \xi$ durch μ teilbar sein soll, so muß nach der Bedeutung der Basis diese Differenz in der Form

(7) $$\eta - \xi = h_1\alpha_1 + h_2\alpha_2 + \cdots + h_n\alpha_n$$

darstellbar sein, worin die $h_1, h_2, \ldots h_n$ ganze rationale Zahlen sind, und wenn diese Bedingung erfüllt ist, so ist auch umgekehrt η mit ξ nach dem Modul μ kongruent. Ordnen wir die rechte Seite mittels (4) nach $\omega_1, \omega_2, \ldots \omega_n$, so ergibt sich aus (5) und (6):

(8)
$$\begin{aligned}
y_1 &= x_1 + h_1 a_{1,1} + h_2 a_{1,2} + \cdots + h_n a_{1,n},\\
y_2 &= x_2 + \qquad\quad h_2 a_{2,2} + \cdots + h_n a_{2,n},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
y_n &= x_n \qquad\qquad\qquad\qquad + h_n a_{n,n}.
\end{aligned}$$

Darin aber liegt der Beweis unserer Behauptung. Denn sind $y_1, y_2, \ldots y_n$ irgend welche gegebene ganze rationale Zahlen, so ist zunächst aus

$$y_n = x_n + h_n a_{n,n}$$

x_n und h_n vollständig bestimmt. Denn es ist h_n der Quotient und x_n der Rest der Division von y_n durch $a_{n,n}$. Darauf erhalten wir aus der vorletzten Gleichung

$$y_{n-1} - h_n a_{n-1,n} = x_{n-1} + h_{n-1} a_{n-1,n-1},$$

wodurch x_{n-1} und h_{n-1} völlig bestimmt sind. Die nächst vorhergehende Gleichung ergibt:

$$y_{n-2} - h_n a_{n-2,n} - h_{n-1} a_{n-2,n-1} = x_{n-2} + h_{n-2} a_{n-2,n-2},$$

wodurch x_{n-2} und h_{n-2} völlig bestimmt sind, und so fahren wir fort, bis alle x_i und h_i bestimmt sind.

Hiernach können wir zweckmäßig die Gesamtheit der Zahlen ξ oder irgend ein System mit diesen kongruenter Zahlen ein volles Restsystem nach dem Modul μ nennen.

Die Anzahl der Individuen eines vollen Restsystems für den Modul μ ist gleich der absoluten Norm von μ.

Aus der Existenz eines vollen Restsystems nach einem Modul μ, die hierdurch nachgewiesen ist, ergibt sich eine Reihe von Sätzen, die mit bekannten elementaren Sätzen der rationalen Zahlentheorie übereinstimmen.

Bedeutet α irgend eine ganze Zahl in Ω und μ ein als Modul dienendes Funktional, und ist α relativ prim zu μ, so sind zwei Zahlen $\alpha\xi$ und $\alpha\xi'$ nur dann kongruent nach dem Modul μ, wenn die ganzen Zahlen ξ, ξ' kongruent sind. Hieraus ergibt sich, daß, wenn ξ ein volles Restsystem nach dem Modul μ durchläuft, dasselbe auch von dem Produkte $\alpha\xi$ gilt, wodurch der Satz bewiesen ist:

4. Ist α eine ganze Zahl in Ω, relativ prim zu dem Modul μ, ferner γ eine beliebige ganze Zahl in Ω, so ist die Kongruenz:

(9) $$\alpha\xi \equiv \gamma \pmod{\mu}$$

immer durch eine ganze Zahl ξ lösbar, und auch nur durch eine, wenn für ξ ein volles Restsystem nach dem Modul μ vorgeschrieben ist.

Wenn hierin μ selbst eine ganze Zahl ist, die wir mit β bezeichnen, so ist auch der Quotient

$$\frac{\alpha\xi - \gamma}{\beta} = -\eta$$

eine ganze Zahl, und dann nimmt der vorstehende Satz die Form an:

5. Sind α, β, γ drei ganze Zahlen in Ω und α, β relativ prim, so kann man zwei andere ganze Zahlen ξ, η in Ω so bestimmen, daß

(10) $$\alpha\xi + \beta\eta = \gamma$$

wird. Insbesondere kann man also auch für zwei beliebige relative Primzahlen α, β die Gleichung

(11) $$\alpha\xi + \beta\eta = 1$$

durch ganze Zahlen ξ, η befriedigen.

Dieser Satz läßt sich auf ein System von mehreren Zahlen übertragen.

Wenn die Zahlen α, β, γ, ... in $\mathfrak{o}$ keinen gemeinsamen Teiler haben, so können wir zunächst Zahlen η_1, ζ_1, ... in $\mathfrak{o}$ so be-

stimmen, daß

$$\beta_1 = \beta\eta_1 + \gamma\zeta_1 + \cdots$$

relativ prim zu α wird. Wir haben nämlich, wenn $\pi_1, \pi_2 \ldots$ die verschiedenen Primfaktoren von α sind, deren keiner in allen $\beta, \gamma \ldots$ aufgehen kann, und wenn etwa β durch π_1 nicht teilbar ist, η_1 durch π_1 unteilbar, die übrigen Zahlen $\zeta_1 \ldots$ durch π_1 teilbar usf. anzunehmen, und erhalten für jede der Zahlen $\eta_1, \zeta_1 \ldots$ die Bedingung, daß sie durch einige der Primfaktoren π_i teilbar, durch andere nicht teilbar sein soll, und dieser Forderung kann nach § 101, 1. immer genügt werden. Dann können wir nach 5. die ganze Zahl τ so bestimmen, daß

$$\alpha\xi + \beta_1\tau = 1$$

wird, und wenn wir $\tau\eta_1 = \eta, \tau\zeta_1 = \zeta, \ldots$ setzen, so erhalten wir den Satz:

6. Sind $\alpha, \beta, \gamma \ldots$ Zahlen in $\mathfrak{o}$ ohne gemeinschaftlichen Teiler, so lassen sich andere Zahlen $\xi, \eta, \zeta \ldots$ in $\mathfrak{o}$ so bestimmen, daß

$$\alpha\xi + \beta\eta + \gamma\zeta + \cdots = 1 \tag{12}$$

wird.

Daraus läßt sich weiter auf folgenden Satz schließen:

7. Sind $\alpha, \beta, \gamma \ldots$ beliebige Zahlen in $\mathfrak{o}$ und μ eine durch den größten gemeinschaftlichen Teiler aller dieser Zahlen teilbare Zahl in $\mathfrak{o}$, so kann man die Zahlen $\xi, \eta, \zeta \ldots$ in $\mathfrak{o}$ so bestimmen, daß

$$\mu = \alpha\xi + \beta\eta + \gamma\zeta + \cdots$$

wird.

Wenn wir nämlich den größten gemeinschaftlichen Teiler der Zahlen $\alpha, \beta, \gamma \ldots$ nach § 99 in der Form

$$\delta = \alpha u + \beta v + \gamma w + \cdots$$

annehmen, worin $u, v, w \ldots$ Variable sind, so gibt es nach Voraussetzung ein ganzes Funktional ω, so daß $\mu = \delta\omega$ ist. Das Funktional ω stellen wir als Quotienten zweier ganzer Funktionen $\varphi : \varepsilon$ dar, von denen ε eine Einheit, und folglich φ eine Funktion mit ganzzahligen Koeffizienten ist, und erhalten so

$$\varepsilon\mu = (\alpha u + \beta v + \gamma w + \cdots)\varphi. \tag{13}$$

Ordnet man beide Seiten dieser Gleichung nach den darin vorkommenden Variablen, so erhält man, wenn $\varepsilon_1, \varepsilon_2 \ldots$ die

Koeffizienten in ε sind, ein System von Gleichungen von folgender Form:

$$\begin{aligned} \varepsilon_1 \mu &= \alpha\xi_1 + \beta\eta_1 + \gamma\zeta_1 + \cdots \\ \varepsilon_2 \mu &= \alpha\xi_2 + \beta\eta_2 + \gamma\zeta_2 + \cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \end{aligned} \tag{14}$$

worin die ξ_i, η_i, ζ_i ... Zahlen in $\mathfrak{o}$ sind. Nun haben aber die Zahlen ε_1, ε_2 ... als Koeffizienten einer Einheit keinen gemeinsamen Teiler, und folglich kann man nach 6. die Zahlen τ_1, τ_2 ... in $\mathfrak{o}$ so bestimmen, daß

$$\varepsilon_1\tau_1 + \varepsilon_2\tau_2 + \cdots = 1 \tag{15}$$

wird. Aus (14) folgt aber, wenn man mit τ_1, τ_2 ... multipliziert und addiert, und dann

$$\xi = \tau_1\xi_1 + \tau_2\xi_2 + \cdots, \quad \eta = \tau_1\eta_1 + \tau_2\eta_2 + \cdots$$

setzt, mittels (15) der zu beweisende Satz 7.

Wir beweisen noch den folgenden Satz:

8. Ist μ, ν, ϱ ... ein System von Funktionalen, deren je zwei relativ prim sind, und α, β, γ ... beliebige ganze Zahlen in Ω, so kann man eine Zahl Θ (nach dem Modul $\mu\nu\varrho$ ...) bestimmen, die den Kongruenzen

$$\Theta \equiv \alpha \pmod{\mu}, \quad \Theta \equiv \beta \pmod{\nu}, \quad \Theta \equiv \gamma \pmod{\varrho} \ldots \tag{16}$$

genügt.

Um ihn zu beweisen, wähle man eine ganze Zahl a in Ω, die relativ prim zu μ, aber durch $\nu\varrho$ ... teilbar ist. Ebenso sei b relativ prim zu ν, aber durch $\mu\varrho$ ... teilbar, und entsprechendes gelte für c ... Dann löse man nach 4. die Kongruenzen

$$a\xi \equiv 1 \pmod{\mu}, \quad b\eta \equiv 1 \pmod{\nu}, \quad c\zeta \equiv 1 \pmod{\varrho} \ldots \tag{17}$$

und setze

$$\Theta = a\alpha\xi + b\beta\eta + c\gamma\zeta \cdots, \tag{18}$$

und diese Zahl genügt offenbar den Kongruenzen (16).

Wenn das Funktional δ ein Teiler des Funktionals μ ist und

$$\mu = \nu\delta,$$

so werden in einem vollen Restsystem ξ nach dem Modul μ alle Reste nach dem Modul ν vorkommen. Jeder dieser Reste wird aber gleich oft unter den Zahlen ξ auftreten. Denn wenn ξ_0 die durch ν teilbaren unter den Resten ξ sind, so sind zwei Zahlen ξ und ξ_1 nur dann nach dem Modul ν kongruent, wenn

$$\xi \equiv \xi_1 + \xi_0 \pmod{\mu}$$

ist. Da nun die Anzahl der nach dem Modul ν verschiedenen Reste $N_a(\nu)$ beträgt, und $N_a(\mu) = N_a(\nu)\, N_a(\delta)$ ist, so folgt

9. In einem vollen Restsystem nach dem Modul μ kommt jeder Rest nach dem Modul ν gleich oft, nämlich $N_a(\delta)$mal vor.

Ist δ der größte gemeinschaftliche Teiler der Zahl α und des Funktionals μ, so wird $\alpha\xi$ mit $\alpha\xi'$ dann und nur dann nach dem Modul μ kongruent sein, wenn

$$\xi \equiv \xi' \pmod{\nu}.$$

Wenn also ξ ein Restsystem (mod μ) durchläuft, so wird $\alpha\xi$ eine durch δ teilbare Zahl γ genau $N_a(\delta)$mal darstellen; und es muß also dabei auch jede der $N_a(\nu)$ durch δ teilbaren Zahlen als Rest erscheinen. Hieraus folgt:

10. Haben α und μ den größten gemeinschaftlichen Teiler δ, so hat die Kongruenz

(19) $$\alpha\xi \equiv \gamma \pmod{\mu}$$

nur dann Lösungen, wenn auch γ durch δ teilbar ist, und in diesem Falle ist die Anzahl der mod μ verschiedenen Lösungen gleich $N_a(\delta)$.

Wenn α und α' zwei nach dem Modul μ kongruente Zahlen sind, so gibt es ein ganzes Funktional η, so daß

(20) $$\alpha' = \alpha + \mu\eta$$

ist, und diese Gleichung ist mit der Kongruenz $\alpha' \equiv \alpha \pmod{\mu}$ gleichbedeutend.

Um den Begriff der Kongruenz auf ganze Funktionale auszudehnen, muß man ein Funktional ω als Quotient zweier ganzer Funktionen $\varphi : e$ darstellen, so daß e eine funktionale Einheit ist. Zwei Funktionale

$$\omega = \frac{\varphi}{e}, \qquad \omega' = \frac{\varphi'}{e'}$$

heißen dann nach dem Modul μ kongruent:

(21) $$\omega \equiv \omega' \pmod{\mu},$$

wenn in der nach den Variablen geordneten Funktion $\varphi e' - e\varphi'$ alle Koeffizienten durch μ teilbar sind. Dies findet z. B. dann statt, wenn φ, φ' und e, e' dieselben Variablen haben, und wenn entsprechende Koeffizienten nach dem Modul μ kongruent sind. Ist dann η ein ganzes Funktional, so gilt auch hier die Gleichung

(22) $$\omega' = \omega + \mu\eta,$$

die wieder mit (21) gleichwertig ist. Man kann daher auch solche Funktionalkongruenzen addieren, subtrahieren und multiplizieren, wie Gleichungen.

§ 106.

Anzahl der zu einem Modul teilerfremden Zahlklassen.

Wenn eine Zahl ω in $\mathfrak{o}$ relativ prim zu einem Funktional μ ist, so gilt das gleiche von allen mit ω nach dem Modul μ kongruenten Zahlen. Wenn wir daher die Zahlen in $\mathfrak{o}$ nach dem Modul μ in Klassen kongruenter Zahlen einteilen, so ist die Zahl dieser Klassen nach § 105, 3. gleich $N_a(\mu)$, und unter diesen Klassen wird sich eine gewisse Anzahl befinden, die nur teilerfremde Zahlen zu μ enthalten. Die Anzahl dieser Zahlklassen, die wir jetzt näher bestimmen wollen, und die ein Analogon zu der Zahl $\varphi(n)$ (§ 67) ist, soll jetzt mit $\psi(\mu)$ bezeichnet sein.

Ist μ in zwei Faktoren ϱ, σ zerlegt, die zueinander teilerfremd sind, so wählen wir zwei Zahlen in $\mathfrak{o}$, nämlich:

α relativ prim zu ϱ, teilbar durch σ

β „ „ „ σ, „ „ ϱ.

Nach § 101, 1. gibt es immer solche Zahlen. Setzen wir dann

(1) $$\zeta = \alpha\xi + \beta\eta$$

und lassen ξ und η volle Restsysteme nach den Moduln ϱ und σ durchlaufen, so durchläuft ζ zugleich ein volles Restsystem nach dem Modul μ. Dies erkennt man leicht daraus, daß ζ nur dann durch μ teilbar ist, wenn zugleich ξ durch ϱ und η durch σ teilbar ist. Es ist aber ζ dann und nur dann relativ prim zu μ, wenn ξ relativ prim zu ϱ und η relativ prim zu σ ist; daraus erhält man

(2) $$\psi(\mu) = \psi(\varrho)\,\psi(\sigma).$$

Hiernach genügt es, wenn wir $\psi(\mu)$ unter der Voraussetzung bestimmen, daß

$$\mu = \pi^r$$

eine Potenz eines Primfunktionals π ist.

Unter dieser Voraussetzung ist $N_a(\pi)^r$ die Anzahl aller Zahlklassen nach dem Modul μ. Um festzustellen, wie viele darunter durch π teilbare Zahlen enthalten, nehmen wir eine durch π, aber nicht durch π^2 teilbare Zahl ω an, und lassen ξ in $\omega\xi$ ein volles Restsystem nach dem Modul π^{r-1} durchlaufen. Dann

erhalten wir alle Zahlklassen nach dem Modul π^r, deren Zahlen durch π teilbar sind. Hiernach ist

$$\psi(\mu) = N_a(\pi)^r - N_a(\pi)^{r-1} = N_a(\mu)\left(1 - \frac{1}{N_a(\pi)}\right)$$

die Anzahl der Zahlklassen nach dem Modul μ, deren Zahlen durch π nicht teilbar sind, und es ergibt sich nach (2) allgemein:

$$\psi(\mu) = N_a(\mu)\,\Pi\left(1 - \frac{1}{N_a(\pi)}\right), \tag{3}$$

worin sich das Produktzeichen Π auf alle voneinander verschiedenen Primfaktoren π von μ erstreckt.

Es sei nun das Funktional μ in zwei funktionale Faktoren δ, δ' zerlegt und eine Zahl ν eines vollen Restsystems (mod μ), sei durch δ teilbar. δ wird der größte gemeinschaftliche Teiler von ν und μ sein, wenn ν/δ relativ prim zu δ' ist. Da jede Zahl eines solchen Restsystems, das $N(\mu)$ Zahlen ν enthält, irgend einen größten gemeinschaftlichen Teiler mit μ gemein haben muß, so folgt, wenn man δ, und folglich auch δ', alle Teiler von μ durchlaufen läßt:

$$\Sigma\psi(\delta) = N_a(\mu), \tag{4}$$

worin sich die Summe Σ auf alle Divisoren δ des Funktionals μ bezieht.

§ 107.

Der Fermatsche Satz.

Aus der Theorie der Kongruenzen lassen sich Folgerungen ziehen, die den aus dem Fermatschen Lehrsatze abgeleiteten Sätzen der rationalen Zahlentheorie genau entsprechen.

Wir wollen unter π ein Primfunktional $f$ten Grades verstehen, also, wenn p die durch π teilbare natürliche Primzahl ist,

$$N_a(\pi) = p^f \tag{1}$$

setzen (§ 100). Ist dann α irgend eine durch π nicht teilbare Zahl in $\mathfrak{o}$, so wird, wie wir schon im vorigen Paragraphen gesehen haben, das Produkt $\alpha\xi$ zugleich mit ξ ein volles Restsystem nach dem Modul π durchlaufen. Lassen wir die durch π teilbare Zahl weg, so bleiben $N_a(\pi) - 1$ Zahlen übrig, und wenn wir das Produkt bilden, so folgt:

$$\alpha^{p^f-1}\,\Pi(\xi) \equiv \Pi(\xi) \pmod{\pi}, \tag{2}$$

wenn $\Pi(\xi)$ das Produkt aller Zahlen eines vollen Restsystems

(mit Ausschluß der Null) bedeutet, und daher durch π nicht teilbar ist. Demnach folgt aus (2):

$$\alpha^{p^f-1} \equiv 1 \pmod{\pi}, \tag{3}$$

oder, wenn man mit α multipliziert,

$$\alpha^{p^f} \equiv \alpha \pmod{\pi}, \tag{4}$$

und in der letzten Form gilt der Satz auch noch, wenn α durch π teilbar ist.

Wir haben also den Fermatschen Lehrsatz:

1. Ist π ein Primteiler der natürlichen Primzahl p, so ist für jede ganze Zahl ω im Körper Ω

$$\omega^{N_a(\pi)} \equiv \omega \pmod{\pi}.$$

Hieran knüpfen sich wichtige Folgerungen:

2. Bezeichnet $f(t)$ eine ganze Funktion $m$ten Grades, deren Koeffizienten ganze Zahlen in Ω sind, und π ein Primfunktional, so hat die Kongruenz

$$f(t) \equiv 0 \pmod{\pi} \tag{5}$$

höchstens m Wurzeln, d. h. es gibt höchstens m inkongruente ganze Zahlen in Ω, die, für t gesetzt, die Kongruenz befriedigen.

Bedeutet nämlich α irgend eine Zahl in $\mathfrak{o}$, so können wir

$$f(t) = (t - \alpha) f_1(t) + f(\alpha) \tag{6}$$

setzen, worin $f_1(t)$ eine ebensolche Funktion wie $f(t)$ ist, aber nur vom $(m-1)$ten Grade. Ist aber $f(\alpha) \equiv 0 \pmod{\pi}$, so muß jede Wurzel von (5) der Kongruenz

$$(t - \alpha) f_1(t) \equiv 0 \pmod{\pi}$$

genügen. Sie muß also entweder mit α kongruent oder eine Wurzel der Kongruenz $(m-1)$ten Grades

$$f_1(t) \equiv 0 \pmod{\pi}$$

sein. Setzen wir unseren Satz als bewiesen voraus für Kongruenzen $(m-1)$ten Grades, so gilt er demnach auch für Kongruenzen $m$ten Grades; und da er für Kongruenzen ersten Grades gilt, so ist er allgemein richtig.

Jede durch π nicht teilbare Zahl in $\mathfrak{o}$ genügt, wie wir gesehen haben, der Kongruenz

$$\omega^{p^f-1} \equiv 1 \pmod{\pi}$$

und die Kongruenz

(7) $$t^{p^f-1} - 1 \equiv 0$$

hat folglich so viele Wurzeln, als ihr Grad angibt, nämlich $p^f - 1$.

Ist nun a die kleinste natürliche Zahl, für die die Kongruenz

(8) $$\omega^a \equiv 1 \pmod{\pi}$$

befriedigt ist, so läßt sich durch das schon oft angewandte Schlußverfahren zeigen, daß jeder andere Exponent l, für den $\omega^l \equiv 1$ ist, ein Vielfaches von a sein muß. Denn wäre l nicht durch a teilbar, so wäre auch, wenn a' der Rest der Division von a durch l ist, $\omega^{a'} \equiv 1$, was nach der Voraussetzung über a nur für $a' = 0$ möglich ist. Also ist a ein Teiler von $p^f - 1$, und wir nennen ω eine zum Exponenten a gehörige Zahl.

Gehört ω zum Exponenten a, so sind die Potenzen

(9) $$1, \omega, \omega^2 \ldots \omega^{a-1}$$

alle inkongruent und bilden also, da sie alle der Kongruenz (8) genügen, nach 2. die Gesamtheit der Wurzeln dieser Kongruenz. Unter den Zahlen (9) müssen daher alle anderen zum Exponenten a gehörigen Zahlen ω gesucht werden. Es wird aber ω^l nur dann zum Exponenten a gehören, wenn l relativ prim zu a ist, und es folgt:

Wenn es überhaupt Zahlen ω gibt, die zum Exponenten a gehören, so ist ihre Anzahl so groß, wie die Anzahl der relativen Primzahlen zu a in der Reihe der Zahlen $0, 1, 2, \ldots a - 1$. Diese Zahl bezeichnen wir, wie schon früher § 67, mit $\varphi(a)$. Daß aber zu jedem Teiler a von $p^f - 1$ immer wenigstens eine Zahl ω und folglich $\varphi(a)$ Zahlen gehören, kann so bewiesen werden:

Bezeichnen wir die Anzahl der Zahlen, die zu einem Teiler a von $p^f - 1$ gehören, mit $\varphi_0(a)$, so ist $\varphi_0(a)$ entweder Null oder $\varphi(a)$; und da jede der $p^f - 1$ Zahlen eines vollen Restsystemmoduls $p^f - 1$ zu einem Teiler a des Moduls gehören muß, so ist:

(10) $$\Sigma \varphi_0(a) = p^f - 1.$$

Ebenso ergibt sich, wenn wir die Formel (4) des vorigen Paragraphen auf die Körper der rationalen Zahlen anwenden, § 68, (4):

(11) $$\Sigma \varphi(a) = p^f - 1,$$

und dies ist nur mit (10) verträglich, wenn

$$\varphi_0(a) = \varphi(a)$$

ist.

Die Zahlen ω, die zu dem Exponenten $p^f - 1$ gehören, deren es hiernach immer $\varphi(p^f - 1)$ nach dem Modul π inkongruente gibt, heißen primitive Wurzeln von π. Ist γ eine solche primitive Wurzel, so bilden die Potenzen

$$1, \gamma, \gamma^2 \dots \gamma^{p^f-2}$$

ein volles Restsystem nach dem Modul π, mit Ausschluß der durch π teilbaren Zahl.

Nach dem Fermatschen Lehrsatze für rationale Zahlen ist $t^{p-1} - 1$ für $t = 1, 2, \dots p - 1$ durch p, und folglich auch durch π teilbar. Die Kongruenz

$$t^{p-1} \equiv 1 \pmod{\pi} \tag{12}$$

hat daher die Wurzeln

$$1, 2, \dots p - 1,$$

und diese sind, da nach § 100, 3. eine rationale Zahl nur dann durch π teilbar ist, wenn sie durch p teilbar ist, untereinander inkongruent. Nach dem Satze 2. hat also die Kongruenz (12) keine anderen Wurzeln als diese.

Multiplizieren wir die Kongruenz (12) noch mit t, so folgt, daß die Kongruenz $p$ten Grades

$$t^p - t \equiv 0 \pmod{\pi}$$

die p Wurzeln

$$0, 1, 2, \dots p - 1,$$

und keine anderen hat. Darin liegt der Beweis des folgenden Satzes:

3. Eine Zahl ω in $\mathfrak{o}$ ist dann und nur dann nach dem Modul π mit einer rationalen Zahl kongruent, wenn sie der Bedingung

$$\omega^p \equiv \omega \pmod{\pi}$$

genügt.

Beachtet man noch, daß die Polynomialkoeffizienten in der $p$ten Potenz eines Polynoms alle durch p teilbar sind, mit Ausnahme derer, die zu den $p$ten Potenzen der einzelnen Glieder des Polynoms gehören (S. 330), so ergibt sich noch folgender Satz, der sich auf ganze Funktionen in Ω von beliebigen Veränderlichen $x, y \dots$ bezieht:

4. Ist $\psi(x, y \dots)$ eine ganze Funktion der Variablen $x, y \dots$ mit ganzzahligen Koeffizienten aus Ω, so ist das Bestehen der Kongruenz

$$[\psi(x, y \ldots)]^p \equiv \psi(x^p, y^p \ldots) \pmod{\pi}$$

die notwendige und hinreichende Bedingung dafür, daß alle Koeffizienten von ψ mit ganzen rationalen Zahlen nach dem Modul π kongruent sind.

§ 108.

Die Dedekindschen Ideale.

Dedekind gründet die Theorie der algebraischen Zahlen auf den Begriff des Ideals.

Wir wollen jetzt nachweisen, daß die Theorie der Ideale im Wesen übereinstimmt mit der Theorie der Funktionale, indem wir zeigen, wie der Übergang von der einen zur anderen bewirkt werden kann.

Das System aller ganzen Zahlen eines algebraischen Zahlkörpers Ω soll, wie oben, mit $\mathfrak{o}$ bezeichnet werden. Ein in $\mathfrak{o}$ enthaltenes Zahlensystem $\mathfrak{a}$ wird ein Ideal genannt, wenn es den beiden Forderungen genügt:

I. Summe und Differenz irgend zweier Zahlen in $\mathfrak{a}$ geben immer wieder Zahlen in $\mathfrak{a}$.

II. Das Produkt irgend einer Zahl in $\mathfrak{a}$ und einer Zahl in $\mathfrak{o}$ gehört dem System $\mathfrak{a}$ an[1]).

Dieser Forderung würde das aus der einzigen Zahl Null bestehende System genügen, was aber der Einfachheit halber nicht als ein Ideal bezeichnet wird.

Das System $\mathfrak{o}$ dagegen ist ein eigentliches Ideal. Ebenso ist das System aller durch eine bestimmte Zahl μ in $\mathfrak{o}$ teilbarer Zahlen $\mathfrak{o}\mu$ ein Ideal, und ein solches wird ein Hauptideal genannt.

Unter dem Produkte $\mathfrak{a}\mathfrak{b}$ zweier Ideale $\mathfrak{a}$ und $\mathfrak{b}$ versteht man den Inbegriff aller Zahlen, die man erhält, wenn man irgend eine Zahl α aus $\mathfrak{a}$ mit einer Zahl β aus $\mathfrak{b}$ multipliziert und eine beliebige Anzahl solcher Zahlenprodukte addiert, also den Inbegriff aller Zahlen von der Form $\Sigma\alpha\beta$. Daß dieses Produkt $\mathfrak{a}\mathfrak{b}$ wieder ein Ideal ist, leuchtet unmittelbar ein. Nach dieser Definition ist z. B. $\mathfrak{o}\mathfrak{a} = \mathfrak{a}$, und das Ideal $\mathfrak{o}$ spielt bei dieser Multiplikation die Rolle der Einheit.

[1]) Vgl. Dirichlet-Dedekind, Vorlesungen über Zahlentheorie im § 167 der dritten, § 177 der vierten Auflage.

Man kann nun die Ideale und Funktionale in der Weise aufeinander beziehen, daß dabei folgende Gesetze obwalten:

1. Jedem ganzen Funktional entspricht ein bestimmtes Ideal, und assoziierten Funktionalen entspricht dasselbe Ideal.
2. Jedem Ideal entsprechen unendlich viele, aber nur assoziierte ganze Funktionale.
3. Dem Produkt zweier oder mehrerer ganzer Funktionale entsprechen die Produkte der den Faktoren entsprechenden Ideale.
4. Einer ganzen Zahl entspricht ein Hauptideal.
5. Den funktionalen Einheiten entspricht das Ideal $\mathfrak{o}$.

Um dieses Entsprechen zu definieren, ordnen wir zunächst dem System aller Einheiten das Ideal $\mathfrak{o}$ zu. Ist dann ferner φ irgend ein ganzes Funktional, das keine Einheit ist, so genügt der Inbegriff der durch φ teilbaren ganzen Zahlen α des Körpers Ω nach § 96 den Forderungen I., II., und ist also ein Ideal, das wir mit $\mathfrak{a}$ bezeichnen[1]) und dem Funktional φ zuordnen. Dasselbe Ideal $\mathfrak{a}$ ist dann auch sämtlichen mit φ assoziierten Funktionalen zugeordnet. Diese Zuordnung hat die Eigenschaften 1., 4., 5.

Sind φ und φ_1 zwei nicht assoziierte Funktionale, so ist gewiß eines von ihnen, etwa φ, nicht durch das andere φ_1 teilbar, und folglich gibt es (nach § 101, 1.) ganze Zahlen, die durch φ, aber nicht durch φ_1 teilbar sind. Folglich sind nicht assoziierten Funktionalen immer verschiedene Ideale $\mathfrak{a}$, $\mathfrak{a}_1$ zugeordnet.

Es ist aber nun auch zu zeigen, daß auf diese Weise alle Ideale des Körpers Ω erhalten werden können, mit anderen Worten, daß jedes von $\mathfrak{o}$ verschiedene Ideal $\mathfrak{a}$ aus der Gesamtheit der durch einen gewissen Funktionalfaktor teilbaren Zahlen besteht.

Wir gehen also jetzt von irgend einem Ideal $\mathfrak{a}$ aus und wählen eine beliebige endliche Menge von Zahlen daraus, $\alpha_1, \alpha_2, \ldots \alpha_r$, deren größter gemeinschaftlicher Teiler δ_r sein mag. Dieses Funktional δ_r hat eine endliche Anzahl von Primfaktoren.

---

[1]) Zur Bezeichnung der Ideale gebrauchen wir mit Dedekind die kleinen deutschen Buchstaben.

Gibt es nun eine Zahl α_{r+1} in $\mathfrak{a}$, die nicht durch δ_r teilbar ist, so hat der größte gemeinschaftliche Teiler δ_{r+1} von δ_r und α_{r+1} weniger Primfaktoren als δ_r. Wenn wir mit dieser Schlußweise fortfahren, so kommen wir zu dem Ergebnis, daß sich aus $\mathfrak{a}$ eine endliche Zahl von Zahlen $\alpha_1, \alpha_2, \ldots \alpha_m$ so auswählen läßt, daß der größte gemeinschaftliche Teiler δ dieser Zahlen in allen Zahlen von $\mathfrak{a}$ aufgeht.

Andererseits gehört jede durch δ teilbare Zahl in $\mathfrak{o}$ zu $\mathfrak{a}$. Denn nach § 105, 7. kann jede durch δ teilbare Zahl α in die Form gesetzt werden:

$$\alpha = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \cdots + \alpha_m \xi_m,$$

worin $\xi_1, \xi_2, \ldots \xi_m$ Zahlen in $\mathfrak{o}$ sind, und folglich gehört nach I. und II. α zum Ideal $\mathfrak{a}$.

Es ergibt sich hieraus, daß, wenn δ eine Einheit ist, das Ideal $\mathfrak{a}$ mit $\mathfrak{o}$ identisch ist. Jedes Ideal $\mathfrak{a}$ ist also dadurch charakterisiert, daß alle seine Zahlen einen gewissen größten gemeinschaftlichen Teiler haben.

Es bleibt noch zu zeigen, daß, wenn die Funktionale φ, ψ den beiden Idealen $\mathfrak{a}$, $\mathfrak{b}$ entsprechen, das Produkt $\varphi\psi$ dem Ideal $\mathfrak{a}\mathfrak{b}$ entspricht.

Da alle Zahlen aus $\mathfrak{a}\mathfrak{b}$ von der Form $\Sigma\alpha\beta$ sind, so ist zunächst klar, daß alle diese Zahlen durch $\varphi\psi$ teilbar sind.

Wenn wir aber nach § 101, 2. das Funktional φ als größten gemeinschaftlichen Teiler zweier Zahlen α_1, α_2 darstellen, so gehören diese Zahlen, als durch φ teilbar, dem Ideal $\mathfrak{a}$ an, und wir können, da es auf einen Einheitsfaktor bei φ nicht ankommt,

$$\varphi = \alpha_1 x_1 + \alpha_2 x_2$$

setzen, wenn x_1, x_2 Variable sind. Ebenso können wir, wenn β_1, β_2 zwei Zahlen aus $\mathfrak{b}$ und y_1, y_2 Variable bedeuten,

$$\psi = \beta_1 y_1 + \beta_2 y_2$$

setzen, und daraus ergibt sich:

$$\varphi\psi = \alpha_1\beta_1 x_1 y_1 + \alpha_1\beta_2 x_1 y_2 + \alpha_2\beta_1 x_2 y_1 + \alpha_2\beta_2 x_2 y_2.$$

Es ist also $\varphi\psi$ nach § 101, 4. der größte gemeinschaftliche Teiler der vier Zahlen $\alpha_1\beta_1$, $\alpha_1\beta_2$, $\alpha_2\beta_1$, $\alpha_2\beta_2$, die dem Ideal $\mathfrak{a}\mathfrak{b}$ angehören, und folglich ist $\varphi\psi$ der größte gemeinschaftliche Teiler aller Zahlen des Ideals $\mathfrak{a}\mathfrak{b}$.

Damit ist die gegenseitige Zuordnung der ganzen Funktionale und Ideale den Forderungen 1. bis 5. gemäß bewerkstelligt.

Den Primfunktionalen entsprechen bei dieser Zuordnung Primideale, und die Zerlegung der Ideale in Primfaktoren und überhaupt die Gesetze der Teilbarkeit der Ideale ergeben sich in völliger Übereinstimmung mit den entsprechenden Sätzen aus der Theorie der Funktionale.

Bei dieser vollständigen Übereinstimmung kann es zu keiner Unzuträglichkeit führen, wenn wir das System der untereinander assoziierten ganzen Funktionale zu einem Gemeinbegriffe zusammenfassen und dafür den Namen Ideal brauchen.

Wir sagen dann auch, daß ein Funktional ein bestimmtes Ideal erzeugt, und alle untereinander assoziierten Funktionale, und nur diese, erzeugen dasselbe Ideal. Eine ganze Zahl erzeugt ein Hauptideal.

Wenn irgend eine Zahl oder ein Funktional durch die Funktionale eines Ideals teilbar ist, so nennen wir es durch das Ideal teilbar, und wenn ein Funktional φ in Faktoren zerlegt ist, denen die Ideale $\mathfrak{a}$, $\mathfrak{b}$ ... entsprechen, so setzen wir auch, indem die Einheitsfaktoren in der Bezeichnung weggelassen werden:

$$\varphi = \mathfrak{a}\,\mathfrak{b} \ldots \tag{1}$$

Eine Basis des Funktionals ist zugleich eine Basis des Ideals, und die absolute Norm des repräsentierenden Funktionals stimmt mit der Zahl überein, die bei Dedekind die Norm des Ideals heißt. Sie soll also auch hier so genannt werden, und wir setzen demnach, wenn das Funktional φ zu dem Ideal $\mathfrak{a}$ gehört,

$$N_a(\varphi) = N(\mathfrak{a}). \tag{2}$$

Die Norm eines Ideals ist also immer eine natürliche Zahl.

Ist $\mathfrak{p}$ ein Primideal und p die durch $\mathfrak{p}$ teilbare natürliche Primzahl, so ist

$$N(\mathfrak{p}) = p^f, \tag{3}$$

und f heißt der Grad des Primideals $\mathfrak{p}$.

Von den Normen der Ideale gilt wie von den Normen überhaupt, der Satz

$$N(\mathfrak{a}\,\mathfrak{b} \ldots) = N(\mathfrak{a})\,N(\mathfrak{b}) \ldots \tag{4}$$

In allen Fragen, die sich auf Teilbarkeit beziehen, können die Ideale an Stelle der Funktionale treten. So werden in den die Kongruenzen betreffenden Betrachtungen der §§ 105 bis 107

durchweg die Moduln durch Ideale ersetzt werden können. An die Stelle der absoluten Normen der Funktionale treten die Normen der Ideale[1]).

§ 109.
Äquivalenz.

Wir nennen jetzt zwei Funktionale, die sich nur durch einen Einheitsfaktor unterscheiden, auch wenn sie gebrochen sind, assoziiert, und bezeichnen die Gesamtheit aller mit einem gebrochenen Funktional assoziierten Funktionale als gebrochenes Ideal. Nun stellen wir folgende Definition auf:

1. Zwei ganze oder gebrochene Funktionale φ, ψ im Körper Ω heißen äquivalent, wenn ihr Quotient $\varphi:\psi$ mit einer Zahl assoziiert ist.

Es heißen also die beiden Funktionale φ und ψ äquivalent, wenn eine Einheit ε und eine Zahl α in Ω existieren, so daß

$$\frac{\varphi}{\psi} = \alpha\varepsilon$$

ist. Als Zeichen für die Äquivalenz gebrauchen wir das folgende:

(1) $$\varphi \sim \psi$$ [2]).

Ist φ äquivalent mit ψ, so ist auch ψ äquivalent mit φ, und ist φ äquivalent mit ψ und mit ψ_1, so folgt aus (1):

$$\frac{\varphi}{\psi} = \alpha\varepsilon, \qquad \frac{\varphi}{\psi_1} = \alpha_1\varepsilon_1,$$

folglich:

$$\frac{\psi}{\psi_1} = \frac{\alpha_1}{\alpha}\,\frac{\varepsilon_1}{\varepsilon},$$

und da $\alpha_1:\alpha$ eine Zahl, $\varepsilon_1:\varepsilon$ eine Einheit ist, so folgt der erste Satz:

---

[1]) Der Begriff der Funktionale, d. h. der Gebrauch unbekannter Größen (Variablen) als Rechnungssymbole in der Algebra geht auf Kronecker zurück (Festschrift). Ich glaube, ihn wesentlich vereinfacht zu haben durch ausgiebige Benutzung der funktionalen Einheiten und die dadurch bedingte Möglichkeit, die Variablen auch im Nenner auftreten zu lassen. Den Namen „Funktional" verdanke ich Dedekind, der übrigens dieser Darstellungsweise der Theorie weit weniger zugeneigt ist als der arithmetischen Reinheit seiner Idealtheorie. Mir scheinen die Funktionale zur Einführung aus dem Grunde nicht ungeeignet, weil sie direkter an die elementaren und vertrauten Rechenoperationen anknüpfen und nicht von vornherein ganz neue Begriffsbildungen nötig machen. Erst neuerdings hat Ernst Jacobsthal die Theorie wieder auf die Funktionale begründet. Crelles Journ., Bd. 140.

[2]) Sprich φ (ist) äquivalent (mit) ψ.

2. Zwei Funktionale, die mit einem dritten äquivalent sind, sind auch untereinander äquivalent.

Teilt man hiernach alle Funktionale des Körpers Ω in Klassen ein, indem man zwei Funktionale in dieselbe oder in verschiedene Klassen wirft, je nachdem sie äquivalent sind oder nicht, so ergibt sich, daß zwei dieser Klassen, die ein einziges gemeinsames Element enthalten, vollständig identisch sein müssen, und die Klasseneinteilung ist also durchaus eindeutig. Jede Klasse ist durch ein beliebiges in ihr enthaltendes Funktional, einen Repräsentanten, völlig bestimmt.

3. Zwei miteinander assoziierte Funktionale sind auch äquivalent und kommen daher in derselben Klasse vor.

Denn wenn φ und ψ assoziiert sind, so ist ihr Quotient $\varphi : \psi$ eine Einheit, und φ und ψ sind also auch äquivalent.

Eine Klasse enthält also nicht bloß die einzelnen Funktionale φ, sondern alle durch diese Funktionale bestimmten Ideale, und wir nennen diese Klassen daher, wenn eine genauere Bezeichnung nötig ist, Funktionalklassen oder, häufiger noch, dem üblichen Sprachgebrauche gemäß, Idealklassen.

4. Die Gesamtheit der ganzen und gebrochenen Zahlen des Körpers Ω, verbunden mit den Einheiten und den Produkten von Zahlen mit Einheiten, bilden unter sich eine Klasse, die die Hauptklasse genannt wird.

Als Repräsentanten der Hauptklasse kann man z. B. die Zahl 1 betrachten. Die Hauptklasse, als Idealklasse aufgefaßt, enthält das Ideal $\mathfrak{o}$ und wird daher in der Folge durch den Buchstaben O bezeichnet.

5. In jeder Idealklasse gibt es ganze Funktionale.

Denn nach der Definition ist, wenn φ irgend ein Funktional und α eine Zahl ist, φ mit $\alpha\varphi$ äquivalent. Wir können aber nach § 97, 5. die Zahl α, sogar rational, so bestimmen, daß $\alpha\varphi$ ein ganzes Funktional wird.

6. Aus jeder Idealklasse C können wir einen Repräsentanten φ auswählen, der nicht nur selbst ein ganzes Funktional ist, sondern auch zu einem beliebig gegebenen ganzen Funktional ω relativ prim ist.

Nehmen wir, um diesen Satz zu beweisen, zunächst nach 5. einen beliebigen ganzen Repräsentanten φ der Klasse C und eine durch φ teilbare ganze Zahl α, so ist

$$\varphi\chi = \alpha, \tag{2}$$

und χ ein ganzes Funktional. Nun wählen wir (nach § 101, 1.) eine durch χ teilbare Zahl β so, daß $\beta:\chi$ relativ prim zu ω wird, und setzen

$$\psi\chi = \beta. \tag{3}$$

Da jetzt $\varphi:\psi$ eine Zahl ist, so ist ψ mit φ äquivalent, und ψ ist ein zu ω teilerfremder Repräsentant der Klasse C[1]).

Wir kommen nun zum Beweise des wichtigen Satzes:

7. Die Anzahl der Idealklassen eines Körpers Ω ist endlich.

Dieser Satz ist gleichbedeutend mit dem folgenden:

8. In jeder Klasse gibt es ganze Funktionale, deren absolute Norm eine bestimmte endliche, nur von der Natur des Körpers Ω abhängige Zahl nicht übersteigt.

Denn weil jedes ganze Funktional ein Faktor seiner absoluten Norm ist, und jede ganze Zahl nur eine endliche Anzahl von Idealfaktoren hat, so gibt es nur eine endliche Anzahl von ganzen Idealen, deren Norm unter einer gegebenen Zahl liegt. Wenn nun bewiesen werden kann, daß in jeder Klasse ein ganzes Ideal vorkommt, dessen Norm unter einer durch den Körper bestimmten endlichen Zahl liegt, so ist die Endlichkeit der Anzahl der Klassen nachgewiesen.

---

[1]) Wir wollen hier im Vorübergehen auf eine Analogie der Äquivalenz der Ideale hinweisen. Die ganze Theorie der algebraischen Zahlen läßt sich, mit den notwendigen Modifikationen, übertragen auf die Theorie der algebraischen Funktionen, die wieder ihren geometrischen Ausdruck in der Theorie der algebraischen Kurven findet. Den Idealen entsprechen dann Punktsysteme auf einer festen Grundkurve und den Hauptidealen volle Schnittpunktsysteme der Grundkurve mit einer anderen algebraischen Kurve. Wenn sich zwei Punktsysteme zu einem vollen Schnittpunktsystem ergänzen, was in der obigen Formel $\varphi\chi = \alpha$ seinen Ausdruck finden würde, so wird von den beiden Punktsystemen φ, χ jedes der Rest des anderen genannt. Zwei Punktsysteme, die, wie φ, ψ, denselben Rest haben, werden in der Geometrie korresidual genannt. Dieser Begriff entspricht also der Äquivalenz. (Vgl. Brill u. Nöther, Mathem. Annalen, Bd. 7. Salmon, Higher plane Curves, deutsch von Fiedler.)

Lassen wir $\omega_1, \omega_2, \ldots \omega_n$ eine Minimalbasis von Ω bedeuten, so werden alle ganzen Zahlen des Körpers aus

$$\omega = \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n \tag{4}$$

erhalten, wenn wir den $x_1, x_2, \ldots x_n$ ganze rationale Zahlwerte erteilen. Wenn wir eine positive ganze Zahl k annehmen und festsetzen, daß keine der Zahlen x_i aus dem Intervall $\pm k$ heraustreten soll, so wird der absolute Wert von ω nicht über der Grenze

$$(|\omega_1| + |\omega_2| + \cdots + |\omega_n|)k = rk$$

liegen, wenn unter $|\omega_1|, |\omega_2|, \ldots$ die absoluten Werte der (reellen oder komplexen) Größen ω_i verstanden sind, und r die Summe $|\omega_1| + |\omega_2| + \cdots$ bedeutet. Bilden wir den Ausdruck (4) für die n konjugierten Körper, und nehmen wir das Produkt, so erhalten wir, wenn wir mit R eine positive reelle Zahl bezeichnen, die über dem Produkt der n Werte r liegt,

$$N_a(\omega) < Rk^n. \tag{5}$$

Die Zahl R ist nur von der Natur des Körpers Ω, nicht aber von k abhängig.

Jetzt sei μ irgend ein ganzes Funktional in Ω, und $N_a(\mu)$ seine absolute Norm. Wenn wir die ganze Zahl k so bestimmen, daß

$$k^n \leqq N_a(\mu) < (k+1)^n, \tag{6}$$

und wenn wir ferner in (4) den Zahlen x_i die Werte 0, 1, 2 ... k erteilen, so ist die Anzahl der verschiedenen Werte, die aus (4) hervorgehen, $(k+1)^n$, also größer als $N_a(\mu)$. Nach § 105, 3 ist aber die Zahl der nach dem Modul μ inkongruenten Zahlen gleich $N_a(\mu)$, und folglich müssen unter den so bestimmten Zahlen ω mindestens zwei verschiedene nach dem Modul μ kongruente Zahlen vorkommen. Ist also $\omega' \equiv \omega''$ (mod μ), so wird die Differenz

$$\alpha = \omega' - \omega'' = (x_1' - x_1'')\omega_1 + \cdots + (x_n' - x_n'')\omega_n \tag{7}$$

durch μ teilbar sein, und zugleich sind die ganzen Zahlen

$$x_1' - x_1'' = a_1, \ldots, \quad x_n' - x_n'' = a_n$$

absolut genommen nicht größer als k. Es gibt eine durch μ teilbare von Null verschiedene Zahl:

$$\alpha = a_1\omega_1 + a_2\omega_2 + \cdots + a_n\omega_n, \tag{8}$$

in der die ganzzahligen Koeffizienten $a_1, a_2 \ldots a_n$ die Grenzen $\pm k$ nicht überschreiten, und folglich ist nach (5) und (6)

$$N_a(\alpha) < Rk^n \leqq RN_a(\mu). \tag{9}$$

Da nun α durch μ teilbar ist, so setzen wir

$$\alpha = \mu\varphi, \quad N_a(\alpha) = N_a(\mu)\, N_a(\varphi), \tag{10}$$

und erhalten aus (9):

$$N_a(\varphi) < R. \tag{11}$$

Wenn nun ψ ein Repräsentant einer beliebig gegebenen Klasse C ist, so wählen wir μ so, daß

$$\beta = \mu\psi$$

eine Zahl ist, und wenn dann nach (10)

$$\alpha = \mu\varphi$$

ist, so ist

$$\frac{\varphi}{\psi} = \frac{\alpha}{\beta},$$

also φ und ψ äquivalent. φ ist also gleichfalls ein Repräsentant der Klasse C, und dieser genügt der Bedingung (11). Es kommt also, wie bewiesen werden sollte, in jeder Klasse ein Funktional vor, dessen absolute Norm unter R liegt.

Die Anzahl der Idealklassen, die wir mit h bezeichnen wollen, ist hiernach eine dem Körper Ω eigentümliche natürliche Zahl und wird die Klassenzahl des Körpers genannt.

In dem einfachsten Falle, wo die Klassenzahl gleich 1 ist, ist jedes Funktional mit einer Zahl assoziiert, d. h. es kann jedes (ganze oder gebrochene) Funktional durch Absonderung eines Zahlenfaktors in eine Einheit verwandelt werden. In diesem Falle läßt sich jede ganze Zahl des Körpers Ω in Primzahlfaktoren, zerlegen, und diese Körper haben eine Theorie, die im wesentlichen mit der rationalen Zahlentheorie übereinstimmt. Für solche Körper ist die Einführung der Funktionale und Ideale nicht notwendig.

Hierher gehören neben dem Körper der rationalen Zahlen unter anderen der Körper der Gaußschen imaginären Zahlen (§ 78).

Die Idealklassen können auf Grund des folgenden Satzes komponiert werden:

9. Sind φ, ψ, φ_1, ψ_1 Funktionale, und φ äquivalent mit φ_1, ψ äquivalent mit ψ_1, so ist auch $\varphi\psi$ äquivalent mit $\varphi_1\psi_1$.

Die Richtigkeit hiervon ergibt sich unmittelbar aus der Definition. Denn wenn $\varphi : \varphi_1$ und $\psi : \psi_1$ mit Zahlen assoziiert sind, so ist auch $\varphi\psi : \varphi_1\psi_1$ mit einer Zahl assoziiert.

Ebenso ergibt sich auch der umgekehrte Satz:

10. Ist φ äquivalent mit φ_1 und $\varphi\psi$ äquivalent mit $\varphi_1\psi_1$, so ist auch ψ äquivalent mit ψ_1.

Betrachten wir also zwei Idealklassen A, B, die auch identisch sein können, und bilden das Produkt $\varphi\psi$ irgend eines Funktionals φ aus A und eines Funktionals ψ aus B, so ist die Klasse C, in der das Produkt $\varphi\psi$ vorkommt, unabhängig von der Wahl von φ und ψ, und die Klasse C ist durch die beiden Klassen A, B völlig bestimmt. Wir nennen C aus A und B komponiert und schreiben symbolisch:

$$C = AB = BA. \tag{12}$$

Die Klasse C enthält alle Produkte eines Elementes von A mit einem Elemente von B, kann aber auch noch andere Funktionale enthalten.

Da diese Komposition aus der wahren Multiplikation abgeleitet ist, so gelten auch die Gesetze der Multiplikation für diese Komposition, nämlich das kommutative und assoziative Gesetz.

Es folgt ferner aus dem Satze 10, daß, wenn $AB = AB_1$ ist, auch $B = B_1$ sein muß, und folglich erzeugen die Idealklassen bei dieser Komposition eine endliche Abelsche Gruppe vom Grade h, auf die wir alle Sätze anwenden können, die wir von solchen Gruppen kennen (achter Abschnitt).

Die Einheit dieser Gruppe ist die schon im § 109 definierte Hauptklasse O, die ja, wie wir gesehen haben, den Repräsentanten 1 hat. Um entgegengesetzte Klassen zu definieren, nehmen wir einen Repräsentanten φ einer Klasse A und eine durch φ teilbare Zahl α. Ist dann $\alpha = \varphi\chi$, so ist χ ein Repräsentant der Klasse A^{-1}, und es ist $AA^{-1} = O$.

Ist A eine beliebige Klasse, und h die Klassenzahl, so ist immer (§ 44, S. 188)

$$A^h = O,$$

und wenn k die kleinste positive Zahl ist, die der Bedingung

$$A^k = O$$

genügt, so ist k ein Teiler von h. Daraus ergibt sich der Satz:

11. Jedes Funktional φ in Ω gehört zu einem bestimmten Exponenten k, der ein Teiler der Klassenzahl h ist, so daß φ^k assoziiert ist mit einer Zahl in Ω.

§ 110.
Primfaktoren der natürlichen Primzahlen.

Eine genauere Untersuchung der in § 104 erklärten Basisform τ von $\mathfrak{o}$ des Körpers Ω soll uns das Mittel geben, jede beliebig gegebene natürliche Primzahl p in ihre Primfaktoren zu zerlegen, und damit alle Primfunktionale des Körpers Ω, also Repräsentanten aller Primideale darzustellen.

Es sei $\mathfrak{p}$ ein beliebiges Primideal vom Grade f, so daß

$$N(\mathfrak{p}) = p^f \tag{1}$$

ist, worin p die natürliche Primzahl bedeutet, die durch den Primfaktor $\mathfrak{p}$ teilbar ist. f ist ein positiver Exponent $\leqq n$. Die kleinste ganze rationale Zahl, die durch $\mathfrak{p}$ teilbar ist, ist p, und jede andere ganze rationale Zahl, die durch $\mathfrak{p}$ teilbar ist, ist daher auch durch p teilbar.

Wir müssen nun auch Kongruenzen nach dem Modul $\mathfrak{p}$ zwischen ganzen Funktionen beliebiger Variablen mit Koeffizienten in $\mathfrak{o}$ betrachten (§ 107, 2.).

Den Körper der rationalen Zahlen bezeichnen wir wie immer mit R und nennen demnach eine Funktion mit rationalen Koeffizienten auch eine Funktion in R.

Die Basisform von $\mathfrak{o}$

$$\tau = \omega_1 t_1 + \omega_2 t_2 + \cdots + \omega_n t_n \tag{2}$$

genügt, wie wir im § 104 gesehen haben, einer Gleichung $n$ten Grades $F(\tau) = 0$, deren Koeffizienten ganze rationale Funktionen von $t_1, t_2, \ldots t_n$ sind. Es ist also $F(\tau)$ jedenfalls durch $\mathfrak{p}$ teilbar, und daraus folgt, daß es ganze Funktionen $\Phi(t)$ in R gibt, die außer t irgend welche Variable enthalten können, die durch die Substitution $t = \tau$ in durch $\mathfrak{p}$ teilbare Funktionale in $\mathfrak{o}$ übergehen, die also der Kongruenz

$$\Phi(\tau) \equiv 0 \pmod{\mathfrak{p}} \tag{3}$$

genügen, und wir werden also sagen können, τ ist eine Wurzel der Kongruenz

$$\Phi(t) \equiv 0 \pmod{\mathfrak{p}}. \tag{4}$$

Die Funktion Φ wird gewiß die Variablen $t_1, t_2, \ldots t_n$ enthalten müssen; sie kann aber auch noch andere Variable enthalten, und wenn wir also die Variablen von Φ mit $t, u_1, u_2 \ldots$ bezeichnen, werden wir auch setzen:

$$\Phi(t) = \Phi(t, u_1, u_2 \ldots), \tag{5}$$

oder kürzer $\Phi(t, u)$, und diese Funktion hat ganze rationale Zahlenkoeffizienten.

Wir wollen jetzt das Übereinkommen treffen, daß für jede Variable x, wie sie in den Funktionalen auftreten,

$$x^p \equiv x \pmod{p} \tag{6}$$

sei. Nach dem Fermatschen Satze bleibt diese Kongruenz richtig, wenn für die Variablen ganze rationale Zahlen gesetzt werden, und daher kann aus dieser Bezeichnung, die das Folgende wesentlich vereinfacht, nicht leicht ein Mißverständnis entstehen.

Wenn wir die Funktion $\Phi(t)$ in die $p$te Potenz erheben, und die Formel § 73, 2. anwenden, so folgt aus der Kongruenz (3) mittels des Fermatschen Satzes und der Bezeichnung (6) eine neue Kongruenz

$$\Phi(\tau^p, u_1, u_2 \ldots) \equiv 0 \pmod{\mathfrak{p}}. \tag{7}$$

Ebenso ist aber auch

$$\tau^p \equiv \omega_1^p t_1 + \omega_2^p t_2 + \cdots + \omega_n^p t_n \pmod{\mathfrak{p}}. \tag{8}$$

Setzen wir demnach

$$\tau_1 = \omega_1^p t_1 + \omega_2^p t_2 + \cdots + \omega_n^p t_n,$$

so ergibt sich aus (7):

$$\Phi(\tau_1, u_1, u_2 \ldots) \equiv 0 \pmod{\mathfrak{p}},$$

und folglich ist τ_1 auch eine Wurzel der Kongruenz (4).

Dieses nämliche Verfahren läßt sich wiederholt anwenden, und wir finden, daß auch

$$\tau_2 = \omega_1^{p^2} t_1 + \omega_2^{p^2} t_2 + \cdots + \omega_n^{p^2} t_n$$

Wurzel der Kongruenz (4) ist, usf.

Wenn wir also ein System von Formen τ_r definieren durch

$$\tau_r = \omega_1^{p^r} t_1 + \omega_2^{p^r} t_2 + \cdots + \omega_n^{p^r} t_n \tag{9}$$

für beliebige positive Exponenten r, so sind alle diese Größen τ_r zugleich Wurzeln der Kongruenz (4). Es ist noch die Frage zu beantworten, wieviele von diesen Formen τ_r voneinander verschieden sind.

Nach § 107, (4) ist sicher

$$\tau_r \equiv \tau_{r'} \pmod{\mathfrak{p}},$$

wenn

$$r \equiv r' \pmod{f}$$

ist, und folglich gibt es unter den τ_r gewiß nicht mehr als f nach dem Modul $\mathfrak{p}$ verschiedene

$$\tau, \tau_1, \tau_2, \ldots \tau_{f-1}. \tag{10}$$

Daß diese Formen aber wirklich voneinander verschieden sind, ergibt sich daraus, daß wir nach §§ 103, 107 für die Variablen $t_1, t_2, \ldots t_n$ solche ganze rationale Zahlen setzen können, daß τ in eine primitive Wurzel γ des Primideals $\mathfrak{p}$ übergeht. Durch dieselbe Substitution werden die Größen (10)

$$\equiv \gamma, \gamma^p, \gamma^{p^2}, \ldots \gamma^{p^{f-1}} \pmod{\mathfrak{p}}, \tag{11}$$

die nach dem Modul $\mathfrak{p}$ alle voneinander verschieden sind. Es können also auch nicht zwei der Formen (10) nach dem Modul $\mathfrak{p}$ kongruent sein, weil sonst auch die beiden entsprechenden Zahlen (11) kongruent ausfallen würden.

Dies fassen wir als Satz so zusammen:

1. Jede Kongruenz (4), deren eine Wurzel $t = \tau$ ist, hat die f verschiedenen Wurzeln

$$\tau, \tau_1, \tau_2, \ldots \tau_{f-1}.$$

Hiernach können wir, indem wir $\Phi(t)$ durch $t - \tau$ algebraisch dividieren,

$$\Phi(t) \equiv (t - \tau)\, \Phi_1(t) \pmod{\mathfrak{p}}$$

setzen, und $\tau_1, \tau_2 \ldots \tau_{f-1}$ sind Wurzeln von $\Phi_1(t) \equiv 0$, worin aber $\Phi_1(t)$ nicht rationale Koeffizienten, sondern Koeffizienten in $\mathfrak{o}$ hat. Dividieren wir $\Phi_1(t)$ wieder durch $t - \tau_1$, und fahren so fort, so folgt endlich, wenn wir die Funktion $f$ten Grades

$$\Pi(t) = (t - \tau)\,(t - \tau_1) \ldots (t - \tau_{f-1}) \tag{12}$$

einführen,

$$\Phi(t) \equiv \Pi(t)\, \Phi_0(t) \pmod{\mathfrak{p}}, \tag{13}$$

worin $\Phi_0(t)$ eine ganze Funktion mit Koeffizienten in $\mathfrak{o}$ ist.

Die Funktion $\Pi(t)$ hängt von den Variablen $t, t_1, \ldots t_n$ ab, und um dies auszudrücken, setzen wir

$$\Pi(t) = \Pi(t, t_1, \ldots t_n).$$

Die Koeffizienten dieser Form sind Zahlen in $\mathfrak{o}$, und es läßt sich noch nachweisen, daß sie mit ganzen rationalen Zahlen nach dem Modul $\mathfrak{p}$ kongruent sind. Dieser Beweis ergibt sich durch Erheben in die $p$te Potenz:

$$[\Pi(t)]^p \equiv (t - \tau^p)\,(t - \tau_1^p) \ldots (t - \tau_{f-1}^p) \pmod{\mathfrak{p}}.$$

Nun ist aber nach (9)

$$\tau_r^p \equiv \omega_1^{p^{r+1}} t_1 + \omega_2^{p^{r+1}} t_2 + \cdots + \omega_n^{p^{r+1}} t_n \text{ (riod } \mathfrak{p}),$$

und wir erhalten also $\tau_r^p \equiv \tau_{r+1}$, und τ_f ist kongruent mit τ. Demnach erhalten wir die Kongruenz:

$$[\Pi(t, t_1, t_2 \dots t_n)]^p \equiv \Pi(t, t_1, t_2, \dots t_n) \pmod{\mathfrak{p}}.$$

Damit ist nach § 107, 4. der Satz bewiesen:

2. Die Funktion

$$\Pi(t) = (t - \tau)(t - \tau_1) \dots (t - \tau_{f-1})$$

ist nach dem Modul $\mathfrak{p}$ mit einer ganzen und homogenen Form $f$ten Grades in R der Variablen $t, t_1, t_2, \dots t_n$ kongruent.

Diese Funktion bezeichnen wir mit $P(t)$ und da auch $\Phi(t)$ in R enthalten, also

$$\Phi(t)^p \equiv \Phi(t), \qquad \Pi(t)^p \equiv \Pi(t) \equiv P \pmod{\mathfrak{p}} \tag{14}$$

ist, und $\Pi(t)$ nicht durch $\mathfrak{p}$ teilbar ist (weil die höchste Potenz t^f den Koeffizienten 1 hat), so folgt aus (13) durch Erhebung zur $p$ten Potenz:

$$\Phi_0(t)^p \equiv \Phi_0(t), \pmod{\mathfrak{p}}. \tag{15}$$

Demnach können wir in (13) auch $\Phi_0(t)$ als ganze Form in R annehmen, und dann muß nach § 100, 3. die Kongruenz (13) nicht nur für den Modul $\mathfrak{p}$, sondern für den Modul p bestehen. Daraus erhalten wir den Satz:

3. Unter den Funktionen $\Phi(t)$, die für $t = \tau$ in ein durch $\mathfrak{p}$ teilbares Funktional übergehen, ist $P(t)$ vom Grade f in bezug auf t vom niedrigsten Grade, und wenn $\Phi(t)$ eine beliebige unter ihnen ist, so läßt sich eine ganze Funktion $\Phi_0(t)$ in R so bestimmen, daß

$$\Phi(t) \equiv P(t)\, \Phi_0(t) \pmod{p}$$

wird.

Lassen wir in $\Phi(t)$ und $\Phi_0(t)$ Glieder weg, deren Koeffizienten durch p teilbar sind, so ist der Grad von $\Phi(t)$ in bezug auf t um f größer als der Grad von $\Phi_0(t)$.

$P(t)$ ist eine ganze Funktion in R der Variablen $t, t_1, t_2 \dots t_n$, die durch die Substitution $t = \tau$ in ein durch $\mathfrak{p}$ teilbares Funktional übergeht, das natürlich nicht mehr in R enthalten ist.

Die Bedeutung dieser Form $P(t)$ tritt nun noch deutlicher hervor, wenn wir den Satz beweisen:

4. Der Primfaktor $\mathfrak{p}$ ist der größte gemeinschaftliche Teiler von p und $P(\tau)$.

Dazu haben wir nachzuweisen, daß, wenn p durch $\mathfrak{p}\mathfrak{p}_1$ teilbar ist, wo $\mathfrak{p}$, $\mathfrak{p}_1$ zwei gleiche oder verschiedene Primfaktoren sind, $P(\tau)$ zwar durch $\mathfrak{p}$, nicht aber durch $\mathfrak{p}\mathfrak{p}_1$ teilbar ist.

Nach § 101, 1. existiert immer eine Zahl ξ in $\mathfrak{o}$, die zwar durch $\mathfrak{p}$, aber nicht durch $\mathfrak{p}\mathfrak{p}_1$ teilbar ist. Diese Zahl wird die Form haben:

$$\xi = a_1\omega_1 + a_2\omega_2 + \cdots + a_n\omega_n, \tag{16}$$

worin die $a_1, a_2, \ldots a_n$ ganze rationale Zahlen sind, und geht also aus der Form τ hervor durch die Substitution

$$(t_1, t_2 \ldots t_n) = (a_1, a_2, \ldots a_n). \tag{17}$$

Wenn wir dieselbe Substitution in den in (9) definierten Formen τ_r machen, so geht τ_r in eine Zahl ξ_r über, die nach dem Fermatschen Satze der Kongruenz

$$\xi^{p^r} \equiv \xi_r \pmod{\mathfrak{p}}$$

genügt und also sicher auch durch $\mathfrak{p}$ teilbar ist.

Wenn wir daher in der Form

$$\Pi(t) = (t - \tau)\,(t - \tau_1) \ldots (t - \tau_{f-1})$$

$\tau_r + \xi_r$ an Stelle von τ_r setzen, so bleibt diese Form mit sich selbst nach dem Modul $\mathfrak{p}$ kongruent. Diese Substitution kommt aber darauf hinaus, daß wir $t_1 + a_1$, $t_2 + a_2 \ldots t_n + a_n$ an Stelle von $t_1, t_2 \ldots t_n$ substituieren und wir erhalten demnach

$$\Pi(t, t_1 + a_1, \ldots t_n + a_n) \equiv \Pi(t, t_1, \ldots t_n) \pmod{\mathfrak{p}},$$

und wenn wir Π durch die kongruente Form P ersetzen:

$$P(t, t_1 + a_1, \ldots t_n + a_n) \equiv P(t, t_1, \ldots t_n) \pmod{\mathfrak{p}}.$$

Da aber die Form P lauter rationale Koeffizienten hat, so muß die letztere Kongruenz auch nach dem Modul p stattfinden, also

$$P(t, t_1 + a_1, \ldots t_n + a_n) \equiv P(t, t_1, \ldots t_n) \pmod{p}. \tag{18}$$

Diese Kongruenz besteht für variable $t, t_1, \ldots t_n$.

Nehmen wir nun an, daß, entgegen dem zu beweisenden Satze, $P(\tau)$ durch $\mathfrak{p}\mathfrak{p}_1$ teilbar sei, so besteht die Kongruenz

$$P(\tau, t_1, \ldots t_n) \equiv 0 \pmod{\mathfrak{p}\mathfrak{p}_1}, \tag{19}$$

und diese bleibt richtig, wenn für die unabhängigen Variablen $t_1, t_2, \ldots t_n$ die Substitution $t_1 + a_1, t_2 + a_2, \ldots t_n + a_n$ gemacht wird, und da hierdurch τ in $\tau + \xi$ übergeht, so folgt:

$$P(\tau + \xi, t_1 + a_1, \ldots t_n + a_n) \equiv 0 \pmod{\mathfrak{p}\mathfrak{p}_1}. \tag{20}$$

Machen wir andererseits die Substitution $t = \tau + \xi$, so folgt aus (18) und (20):

(21) $$P(\tau + \xi, t_1, \ldots t_n) = P(\tau + \xi) \equiv 0 \pmod{\mathfrak{p}\mathfrak{p}_1}.$$

Nehmen wir nun zunächst an, $\mathfrak{p}_1$ sei von $\mathfrak{p}$ verschieden, dann können wir so schließen:

Wir setzen in (21) $t_1 = t_2 = \cdots = t_n = 0$, also auch $\tau = 0$. Dadurch aber geht $P(t)$ nach (14) in t^f über (abgesehen von Vielfachen von p), und es ergibt sich aus (21):

$$\xi^f \equiv 0 \pmod{\mathfrak{p}\mathfrak{p}_1},$$

was aber der Annahme widerspricht, daß ξ nicht durch $\mathfrak{p}_1$ teilbar sein soll.

Ist aber $\mathfrak{p}_1 = \mathfrak{p}$, so ordnen wir (21) nach Potenzen von ξ und erhalten

(22) $$P(\tau + \xi) = P(\tau) + \xi P'(\tau) + \frac{\xi^2}{2} P''(\tau) \ldots,$$

wenn $P'(t)$, $P''(t)$, ... die Derivierten von $P(t)$ sind, wobei zu beachten ist, daß die Formen

$$\frac{1}{2} P''(t), \qquad \frac{1}{2.3} P'''(t) \ldots$$

trotz der scheinbaren Nenner ganze Formen in R sind. Da nun nach (19) und (21) $P(\tau + \xi)$ und $P(\tau)$ durch $\mathfrak{p}^2$ teilbar sind, und ebenso nach Voraussetzung $\xi^2, \xi^3 \ldots$, während ξ nicht durch $\mathfrak{p}^2$ teilbar ist, so folgt aus (22):

$$P'(\tau) \equiv 0 \pmod{\mathfrak{p}},$$

woraus wegen

$$P(t) \equiv \Pi(t) \pmod{\mathfrak{p}}$$

nach der Bedeutung (12) von $\Pi(t)$ folgt:

$$\Pi'(\tau) \equiv (\tau - \tau_1)(\tau - \tau_2) \ldots (\tau - \tau_{f-1}) \equiv 0 \pmod{\mathfrak{p}}.$$

Dies ist aber unmöglich, weil die τ, $\tau_1 \ldots \tau_{f-1}$, wie wir gesehen haben, nach dem Modul $\mathfrak{p}$ inkongruent sind. Damit ist also unser Satz 4 vollständig bewiesen. Wir können hiernach, wenn x, y zwei neue Variable bedeuten, ein Funktional π des Ideals $\mathfrak{p}$ bilden:

(23) $$\pi = xp + yP(\tau).$$

Wir betrachten nun ganze Formen $\Phi(t)$ in R, die durch die Substitution $t = \tau$ nicht nur durch $\mathfrak{p}$, sondern durch die natürliche Primzahl p teilbar werden, also der Kongruenz

(24) $$\Phi(\tau) \equiv 0 \pmod{p}$$

genügen. Die Primzahl p möge folgendermaßen in ihre Primfaktoren in Ω zerlegt sein:

$$p = \mathfrak{p}\mathfrak{p}_1\mathfrak{p}_2 \ldots, \tag{25}$$

worin

$$\mathfrak{p}, \mathfrak{p}_1, \mathfrak{p}_2 \ldots$$

gleiche oder verschiedene Primideale der Grade

$$f, f_1, f_2 \ldots$$

sind. Diesen Primfaktoren entspricht (nach dem Satze 3.) eine Reihe ganzer rationaler Funktionen

$$P(t),\ P_1(t),\ P_2(t) \ldots$$

der Grade $f, f_1, f_2 \ldots$, und wenn etwa $\mathfrak{p}$ mit $\mathfrak{p}_1$ identisch ist, so ist auch $P(t)$ mit $P_1(t)$ identisch. Wenn man in (25) rechts und links die Norm nimmt, und die Formeln $N(\mathfrak{p}) = p^f$, $N(p) = p^n$ berücksichtigt, so folgt:

$$n = f + f_1 + f_2 + \cdots \tag{26}$$

Wenn nun $\Phi(t)$ eine der Bedingung (24) genügende ganze rationale Form ist, so folgt aus dem Satze 3.:

$$\Phi(t) \equiv P(t)\ \Phi_1(t) \pmod{p}, \tag{27}$$

worin $\Phi_1(t)$ eine ganze Funktion in R ist, die der Bedingung

$$P(\tau)\ \Phi_1(\tau) \equiv 0 \pmod{\mathfrak{p}\mathfrak{p}_1\mathfrak{p}_2 \ldots}$$

genügt. Nach dem Satze 4. folgt hieraus:

$$\Phi_1(\tau) \equiv 0 \pmod{\mathfrak{p}_1\mathfrak{p}_2 \ldots}, \tag{28}$$

und daraus schließt man wieder nach Satz 3. (auf $\mathfrak{p}_1$ angewandt)

$$\Phi_1(t) \equiv P_1(t)\ \Phi_2(t) \pmod{p}.$$

Hierin läßt sich dieselbe Betrachtung wiederholen, die zu der Kongruenz

$$\Phi_2(t) \equiv 0 \pmod{\mathfrak{p}_2 \ldots}$$

führt, woraus wieder nach 3.

$$\Phi_2(t) \equiv P_2(t)\ \Phi_3(t) \pmod{p}$$

zu schließen ist. Fährt man damit fort, bis alle Primfaktoren von p berücksichtigt sind, so ergibt sich der folgende Satz:

5. Ist $\Phi(t)$ eine ganze Funktion in R, die durch die Substitution $t = \tau$ durch p teilbar wird, so läßt sich eine andere ganze Funktion $\Phi_0(t)$ in R so bestimmen, daß

$$\Phi(t) \equiv \Phi_0(t)\ P(t)\ P_1(t)\ P_2(t) \ldots \pmod{p} \tag{29}$$

wird.

Das Produkt $P(t)\,P_1(t)\,P_2(t)\ldots$ ist vom Grade

$$n = f + f_1 + f_2 + \cdots$$

und ist die ganze rationale Form niedrigsten Grades, die durch die Substitution $t = \tau$ durch p teilbar wird.

Zu den im Satze 5. vorkommenden Funktionen $\Phi(t)$ gehört auch die ganze Funktion $n$ten Grades

$$F(t) = N(t - \tau),$$

die für $t = \tau$ verschwindet. Für diese Funktion wird $\Phi_0(t)$ von t unabhängig, und da sowohl in $F(t)$ als in $P(t)$, $P_1(t)$, $P_2(t)\ldots$ die höchste Potenz von t den Koeffizienten 1 hat, so wird $\Phi_0(t) = 1$. Es gilt also die Kongruenz

$$N(t - \tau) \equiv P(t)\,P_1(t)\,P_2(t)\ldots \pmod{p}. \tag{30}$$

Fassen wir in der Zerlegung (25) der Primzahl p die gleichen Primfaktoren zu Potenzen zusammen, so können wir, wenn $\mathfrak{p}_1, \mathfrak{p}_2\ldots$ verschiedene Primideale der Grade $f_1, f_2\ldots$ sind, die positiven Exponenten $e_1, e_2, \ldots$ so annehmen, daß

$$p = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\ldots \tag{31}$$

und

$$n = e_1 f_1 + e_2 f_2 + \cdots \tag{32}$$

wird. Die Formel (30) nimmt dann die Gestalt an:

$$N(t - \tau) \equiv [P_1(t)]^{e_1}\,[P_2(t)]^{e_2}\ldots \pmod{p}. \tag{33}$$

§ 111.

Dedekinds Satz über die Körperdiskriminante.

Die in der Diskriminante Δ des Körpers Ω aufgehenden natürlichen Primzahlen haben in bezug auf ihre Zerlegung in Primfaktoren einen besonderen Charakter, über den ein Satz von Dedekind die bündigste Auskunft gibt, zu dessen Ableitung wir jetzt schreiten[1]).

Wir bilden nach § 103, 4. die Potenzen der Basisform

$$\tau = t_1\omega_1 + t_2\omega_2 + \cdots + t_n\omega_n,$$

und erhalten die Ausdrücke:

$$\tau^k = u_{1,k}\,\omega_1 + u_{2,k}\,\omega_2 + \cdots + u_{n,k}\,\omega_n, \tag{1}$$

[1]) Dedekind, „Über die Diskriminanten endlicher Körper“ im 29. Bande der Abhandlungen der Gesellschaft der Wissenschaften in Göttingen (1882).

worin die $u_{i,k}$ ganze rationale Funktionen der Variablen $t_1, t_2, \ldots t_n$ sind. Wir setzen wie oben

$$F(t) = N(t - \tau), \tag{2}$$

so daß $F(\tau) = 0$ ist. Nun bilden wir nach § 102, (9) die Diskriminante von τ:

$$\Delta(\tau) = (-1)^{\frac{n(n-1)}{2}} N(F')(\tau), \tag{3}$$

wofür sich nach (1) der Wert $U^2\Delta$ ergibt, wenn Δ die Körperdiskriminante und

$$U = \begin{vmatrix} u_{1,0}, & u_{2,0} & \ldots & u_{n,0} \\ u_{1,1}, & u_{2,1} & \ldots & u_{n,1} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ u_{1,n-1} & u_{2,n-1} & \ldots & u_{n,n-1} \end{vmatrix}, \tag{4}$$

also eine ganze Funktion in R ist.

Wir müssen nachweisen, daß die Form U eine Einheit ist. Angenommen, es gehe in U irgend eine Primzahl p auf, so können wir, wie schon im § 104 bewiesen ist, ein System ganzer Funktionen $y_0, y_1 \ldots y_{n-1}$ in R, die nicht alle durch p teilbar sind, so bestimmen, daß für $i = 1, 2 \ldots n$

$$y_0 u_{i,0} + y_1 u_{i,1} + \cdots + y_{n-1} u_{i,n-1} \equiv 0 \pmod{p}$$

wird. Dann aber ergibt sich aus (1):

$$y_0 + y_1 \tau + \cdots + y_{n-1} \tau^{n-1} \equiv 0 \pmod{p}.$$

Dies widerspricht aber dem Satze 5 des vorigen Paragraphen, nach dem τ nach einem Primzahlmodul p keiner Kongruenz von niedrigerem als $n$ten Grade genügen kann.

Demnach ergibt sich aus (3), daß die Grundzahl Δ des Körpers, vom Vorzeichen abgesehen, die absolute Norm der Funktion $F'(\tau)$ ist, daß also

$$\pm\Delta = N_a[F'(\tau)] \tag{5}$$

zu setzen ist.

Wenn nun statt der ω_i eine andere Basis ω_i' von $\mathfrak{o}$ zugrunde gelegt wird, so tritt an Stelle von τ eine andere Form:

$$\tau' = \omega_1' t_1 + \omega_2' t_2 + \cdots + \omega_n' t_n, \tag{6}$$

wofür wir auch setzen können:

$$\tau' = \omega_1 t_1' + \omega_2 t_2' + \cdots + \omega_n t_n', \tag{7}$$

und darin sind die ω_i' mit den ω_i durch eine lineare Substitution mit der Determinante ± 1 verbunden (§ 103):

$$(8) \qquad (\omega_1', \omega_2', \dots \omega_n') = C(\omega_1, \omega_2, \dots \omega_n).$$

Führt man diese Substitution in (6) aus, und ordnet nach $\omega_1, \omega_2, \dots \omega_n$, so ergibt sich:

$$(9) \qquad (t_1', t_2', \dots t_n') = C_1(t_1, t_2, \dots t_n),$$

wenn C_1 die transponierte Substitution von C ist. Bildet man die Funktion $F(t)$ für die Funktion τ', also $N(t - \tau')$ so mag sich $F_1(t)$ ergeben; die Ableitung für $t = \tau'$ sei $F_1'(\tau')$. Setzen wir

$$F'(\tau) = \Psi(t_1, t_2, \dots t_n),$$

so ist Ψ eine ganze Funktion in R, und es ergibt sich:

$$F_1'(\tau') = \Psi(t_1', t_2', \dots t_n').$$

Da nach (9) ebensowohl die τ_i also die τ_i' als unabhängige Variable betrachtet werden können, so sind die Funktionale als $F'(\tau)$ und $F_1'(\tau)$ assoziiert (§ 101, 3.).

Die Funktion $F'(\tau)$, deren absolute Norm gleich dem absoluten Werte der Grundzahl ist, nennen wir das Grundfunktional, und das durch $F'(\tau)$ repräsentierte Ideal das Grundideal des Körpers Ω. Nun sei $\mathfrak{p}^e$ die höchste Potenz des Primideals $\mathfrak{p}$, die in p aufgeht, und $e \lesseqgtr 1$, dann können wir nach (33) des vorigen Paragraphen

$$(10) \qquad F(t) \equiv P(t)^e \Phi(t) \pmod{p}$$

setzen, worin $\Phi(t)$ eine ganze Funktion in R von der Beschaffenheit ist, daß $\Phi(\tau)$ nicht durch $\mathfrak{p}$ teilbar ist. Aus (10) aber folgt, indem wir die Ableitung nach t bilden, was offenbar gestattet ist,

$$F'(t) \equiv eP(t)^{e-1} P'(t) \Phi(t) + P(t)^e \Phi'(t) \pmod{p}.$$

Setzen wir hierin $t = \tau$, so geht $P'(t)$ in eine durch $\mathfrak{p}$ unteilbare Form $P'(\tau)$ über (was wir schon im Beweis von § 110, 4. gezeigt und benutzt haben) und wir erhalten:

$$(11) \qquad F'(\tau) \equiv eP(\tau)^{e-1} P'(\tau) \Phi(\tau) \pmod{\mathfrak{p}^e}.$$

Nun ist (nach § 110, 4.) $P(\tau)$ durch $\mathfrak{p}$, aber nicht durch $\mathfrak{p}^2$ teilbar, $P'(\tau)$ und $\Phi(\tau)$ sind durch $\mathfrak{p}$ nicht teilbar, und so gibt uns also die Formel (11) den Beweis des folgenden Satzes:

1. Ist $\mathfrak{p}$ ein beliebiges Primideal, p die durch $\mathfrak{p}$ teilbare natürliche Primzahl, und $\mathfrak{p}^e$ die höchste in p aufgehende Potenz von $\mathfrak{p}$, so ist die Grundform

$F'(\tau)$ allemal teilbar durch $\mathfrak{p}^{e-1}$; ist ferner der Exponent e nicht teilbar durch p, so ist $F'(\tau)$ nicht teilbar durch $\mathfrak{p}^e$; ist aber e teilbar durch p, so ist $F'(\tau)$ teilbar durch $\mathfrak{p}^e$ und vielleicht durch noch höhere Potenzen von $\mathfrak{p}$.

Wenn e größer als 1 ist, so ist hiernach $F'(\tau)$ durch $\mathfrak{p}$ teilbar, und folglich ist $N_a[F'(\tau)]$ und also auch die Grundzahl Δ durch p teilbar.

Wenn aber alle Primideale $\mathfrak{p}$ nur in erster Potenz in p aufgehen, so ist $F'(\tau)$ relativ prim zu p. Zerlegt man also $F'(\tau)$ in seine Primfaktoren, so kommt darunter keiner vor, dessen Norm eine Potenz von p ist, folglich ist auch $N_a[F'(\tau)]$ und Δ durch p nicht teilbar. Daraus folgt dann der Satz:

2. Eine natürliche Primzahl p ist dann und nur dann im Körper Ω durch das Quadrat eines Primfaktors teilbar, wenn p in der Grundzahl von Ω aufgeht.

Aus diesen Betrachtungen können wir noch andere wichtige Schlüsse ziehen.

Wenn wir das System der linearen Gleichungen (1) in bezug auf $\omega_1, \omega_2 \dots \omega_n$ auflösen, so erhalten wir Ausdrücke mit dem Nenner U, der, wie wir gesehen haben, eine Einheit ist. Substituiert man diese Ausdrücke in

$$\omega = x_1\omega_1 + x_2\omega_2 + \cdots + x_n\omega_n, \tag{12}$$

worin die $x_1, x_2, \dots x_n$ ganze rationale Zahlen oder ganze rationale Funktionale sind, so erhält man einen Ausdruck von der Form

$$\omega = A_0 + A_1\tau + A_2\tau^2 + \cdots + A_{n-1}\tau^{n-1}, \tag{13}$$

worin die $A_0, A_1, \dots A_{n-1}$ gleichfalls ganze rationale Funktionale bedeuten. Nach § 103, 4. wird aber in dieser Form jede ganze Zahl und jedes ganze Funktional des Körpers Ω dargestellt, und wir können daher den Satz aussprechen:

3. Die Potenzen

$$1, \tau, \tau^2, \dots \tau^{n-1}$$

bilden eine Basis der ganzen Funktionale des Körpers Ω.

Wendet man auf die Darstellung (13) die Sätze § 17, (19) an, so ergibt sich, wenn S das Zeichen für die im § 55 erklärte Spur ist,

$$S\frac{\omega}{F'(\tau)} = A_{n-1}, \tag{14}$$

wodurch der Satz bewiesen ist:

4. Ist ω eine ganze Zahl oder ein ganzes Funktional des Körpers Ω, so ist die Spur von $\frac{\omega}{F'(\tau)}$ ein ganzes rationales Funktional[1]).

Ersetzt man die Variablen $t_1, t_2, \ldots t_n$ durch ganze rationale Zahlen, so entsteht aus der Form τ eine ganze Zahl Θ des Körpers Ω, also eine Zahl in $\mathfrak{o}$, und wegen der Natur der Minimalbasis kann auf diese Weise jede Zahl in $\mathfrak{o}$ erzeugt werden. Durch dieselbe Substitution geht U in eine ganze rationale Zahl über, die von Dedekind der Index der Zahl Θ genannt wird.

Eine natürliche Primzahl, die in dem Index nicht aufgeht, gestattet nach Dedekind eine einfachere Behandlung als die anderen. Er versucht also zunächst für eine gegebene Primzahl die für die t_i zu setzenden Zahlen so zu bestimmen, daß der Index durch diese Primzahl nicht teilbar wird, beweist aber dann durch ein Beispiel, daß dies nicht immer möglich ist, d. h. daß es unter Umständen Primzahlen gibt, die im Index einer jeden Zahl in $\mathfrak{o}$ aufgehen. Es würde dies auf solche Formen U führen, die, obwohl sie primitiv sind, doch für rationale ganze Werte der Variablen stets durch eine bestimmte Primzahl teilbar sind. Das Beispiel, das Dedekind wählt, ist der Körper Ω, der zu der kubischen Gleichung

$$\alpha^3 - \alpha^2 - 2\alpha - 8 = 0 \tag{15}$$

gehört.

Setzt man

$$\beta = \tfrac{1}{2}\alpha(\alpha - 1) - 1, \tag{16}$$

so erhält man nach (15):

---

[1]) Zu diesem Paragraphen ist zu vergleichen: Dedekind, Über den Zusammenhang der Theorie der Ideale und der Theorie der höheren Kongruenzen. Abhandlungen der Gesellschaft der Wissenschaften zu Göttingen, Bd. 23, 1878.

$$
\begin{aligned}
\alpha^2 &= 2 + \alpha + 2\beta \\
\alpha\beta &= 4 \\
\beta^2 &= \tfrac{1}{2}\alpha^2\beta - \tfrac{1}{2}\alpha\beta - \beta \\
&= -2 + 2\alpha - \beta \\
\beta^3 &= -\beta^2 - 2\beta + 8
\end{aligned} \tag{17}
$$

und β ist daher eine ganze Zahl.

Die Diskriminante $\Delta(1, \alpha, \alpha^2)$ ist die Diskriminante der kubischen Gleichung (15) und hat nach § 15 den Wert -4.503, und nach (17) ergibt sich die lineare Substitution:

$$
(1, \alpha, \alpha^2) = \begin{pmatrix} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 2, & 1, & 2 \end{pmatrix} (1, \alpha, \beta), \tag{18}
$$

deren Determinante $= 2$ ist. Hieraus folgt:

$$
\begin{aligned}
\Delta(1, \alpha, \alpha^2) &= 4\Delta(1, \alpha, \beta), \\
\Delta(1, \alpha, \beta) &= -503.
\end{aligned}
$$

Da aber 503 eine Primzahl ist und folglich keinen quadratischen Teiler hat, so ist nach § 103, 3. die Körperdiskriminante $= -503$, und $(1, \alpha, \beta)$ ist eine Minimalbasis.

Wir bilden jetzt die Basisform:

$$
\begin{aligned}
\tau &= t_1 + \alpha t_2 + \beta t_3 \\
\tau^2 &= t_1^2 + \alpha^2 t_2^2 + \beta^2 t_3^2 \\
&\quad + 2\alpha t_1 t_2 + 2\beta t_1 t_3 + 2\alpha\beta t_2 t_3
\end{aligned}
$$

und nach (17):

$$
\begin{aligned}
u_{10} &= 1, & u_{20} &= 0, \\
u_{11} &= t_1, & u_{21} &= t_2, \\
u_{12} &= t_1^2 + 2t_2^2 - 2t_3^2 + 8t_1 t_2, & u_{22} &= t_2^2 + 2t_3^2 + 2t_1 t_2, \\
u_{30} &= 0 \\
u_{31} &= t_3 \\
u_{32} &= 2t_2^2 - t_3^3 + 2t_1 t_3,
\end{aligned}
$$

folglich:

$$
U = u_{21} u_{32} - u_{31} u_{22} = 2t_2^3 - t_2 t_3^2 - t_3 t_2^2 - 2t_3^3. \tag{19}
$$

Die Funktion U ist zwar primitiv, erhält aber doch für jede Bestimmung von t_2, t_3 als ganze Zahlen einen durch 2 teilbaren Wert. Es stellt sich hieraus die noch wenig untersuchte Aufgabe, solche primitive Funktionen zu ermitteln, die einen festen Teiler haben, der sich bei allen ganzzahligen Werten der Variablen einstellt. Ein anderes Beispiel einer solchen Funktion ist $t^p - t$, die, wenn p eine Primzahl ist, für jedes ganzzahlige t nach dem Fermatschen Satze durch p teilbar ist.

§ 112.

Primideale in Normalkörpern.

Es sei jetzt Ω ein Normalkörper über R vom Grade n, es seien also die konjugierten Körper alle mit Ω identisch. Die Substitutionen (Θ, Θ_1), (Θ, Θ_2), ... (Θ, Θ_n), durch die irgend eine Zahl aus Ω in eine konjugierte Zahl übergeht, bilden eine Gruppe, die Gruppe des Körpers Ω. Diese Gruppe bezeichnen wir mit Φ, und ihre Substitutionen mit φ, φ_1, φ_2, ... Ist diese Gruppe kommutativ, so heißt Ω ein Abelscher, und ist die Gruppe zyklisch, ein zyklischer Körper.

Wir bezeichnen mit

$$\omega \mid \varphi \tag{1}$$

die Zahl, die durch die Substitution φ aus ω hervorgeht, so daß ω und $\omega \mid \varphi$ in Ω enthalten sind.

Ebenso wie die Zahlen ω gehen auch alle Funktionale und damit zugleich alle Ideale $\mathfrak{a}$ des Körpers Ω durch eine Substitution φ in bestimmte Funktionale und Ideale $\mathfrak{a} \mid \varphi$ über, und wenn ω eine durch $\mathfrak{a}$ teilbare ganze Zahl ist, so ist $\omega \mid \varphi$ durch $\mathfrak{a} \mid \varphi$ teilbar.

Wenn $\mathfrak{a}$ irgend ein Ideal in Ω ist, so gibt es gewisse Substitutionen ψ in Φ, die der Bedingung

$$\mathfrak{a} \mid \psi = \mathfrak{a} \tag{2}$$

genügen, darunter immer die identische Substitution, und diese Substitutionen ψ bilden eine Gruppe Ψ, die ein Teiler von Φ ist. Wir nennen Ψ die zum Ideal $\mathfrak{a}$ gehörige Gruppe, und wir sagen auch, $\mathfrak{a}$ gehört zu der Gruppe Ψ.

Da man in jeder Gleichung zwischen Zahlen und Funktionalen in Ω alle Substitutionen der Gruppe Φ ausführen darf, ohne daß die Gleichung zu bestehen aufhört, so folgt, daß Einheiten, ganze Funktionale, assoziierte Funktionale, durcheinander teilbare Funktionale diese Eigenschaften nicht verlieren, wenn irgend eine der Substitutionen von Φ ausgeführt wird. Wir heben den Satz hervor:

> Ist $\mathfrak{p}$ ein Primideal, so sind auch alle mit $\mathfrak{p}$ konjugierten Ideale $\mathfrak{p} \mid \varphi$ Primideale. Ist $\mathfrak{a}$ durch irgend eine Potenz von $\mathfrak{p}$ teilbar, so ist $\mathfrak{a} \mid \varphi$ durch die gleiche Potenz von $\mathfrak{p} \mid \varphi$ teilbar.

Ist $\mathfrak{p}$ ein Primideal vom Grade f, das in der natürlichen Primzahl p aufgeht, so ist

$$N\mathfrak{p} = p^f, \tag{3}$$

und da die Norm eines Ideals gleich dem Produkte aller konjugierten Ideale ist, so ist p durch kein anderes als die mit $\mathfrak{p}$ konjugierten Ideale $\mathfrak{p}\,|\,\varphi$ teilbar. Da die Ideale $\mathfrak{p}\,|\,\varphi$ alle dieselbe Norm haben, so haben sie auch denselben Grad f.

Ist $\mathfrak{p}^g$ die höchste in p aufgehende Potenz von $\mathfrak{p}$, so ist nach 1. p auch durch die $g$te Potenz aller mit $\mathfrak{p}$ konjugierten Primfaktoren und durch keine höhere Potenz teilbar. Wenn nun $\mathfrak{p}_1, \mathfrak{p}_2, \ldots \mathfrak{p}_e$ die voneinander verschiedenen unter den mit $\mathfrak{p}$ konjugierten Primidealen sind, so ist

$$p = (\mathfrak{p}_1 \mathfrak{p}_2 \ldots \mathfrak{p}_e)^g, \tag{4}$$

und wenn man die Norm auf beiden Seiten nimmt, und mit n den Grad des Körpers Ω bezeichnet, so daß die Norm von p gleich p^n wird, so folgt aus (3):

$$n = efg. \tag{5}$$

Es sind also die natürlichen Zahlen e, f, g Teiler von n.

Wenn $\mathfrak{p}$ durch φ_1 in $\mathfrak{p}_1$ übergeht, wenn also

$$\mathfrak{p}_1 = \mathfrak{p}\,|\,\varphi_1,$$

so ist $\mathfrak{p} = \mathfrak{p}_1\,|\,\varphi_1^{-1}$, und wenn Ψ die zu $\mathfrak{p}$ gehörige Gruppe ist, so ist auch für jede in Ψ enthaltene Substitution ψ

$$\mathfrak{p}_1 = \mathfrak{p}\,|\,\psi\varphi_1, \quad \mathfrak{p}_1 = \mathfrak{p}_1\,|\,\varphi_1^{-1}\psi\varphi_1.$$

Ist umgekehrt $\mathfrak{p}_1\,|\,\varphi = \mathfrak{p}_1$, so folgt $\mathfrak{p}\,|\,\varphi_1\,\varphi\,\varphi_1^{-1} = \mathfrak{p}$, d. h. $\varphi_1\,\varphi\,\varphi_1^{-1} = \psi$ oder $\varphi = \varphi_1^{-1}\psi\,\varphi_1$.

Es gehört daher $\mathfrak{p}_1$ zur Gruppe $\varphi_1^{-1}\Psi\varphi_1$, und wenn

$$\mathfrak{p}_1 = \mathfrak{p}\,|\,\varphi_1, \quad \mathfrak{p}_2 = \mathfrak{p}\,|\,\varphi_2 \ldots \quad \mathfrak{p}_e = \mathfrak{p}\,|\,\varphi_e$$

ist, so ist

$$\Phi = \Psi\varphi_1 + \Psi\varphi_2 + \cdots + \Psi\varphi_e. \tag{6}$$

Ψ ist also ein Teiler von Φ vom Index e und vom Grade gf.

Ist

$$\tau = t_1\omega_1 + t_2\omega_2 + \cdots + t_{mn}\omega_{mn} \tag{7}$$

eine Basisform von $\mathfrak{o}$, so wird es gewisse Substitutionen χ in Φ geben, und darunter immer die identische, die der Kongruenz

$$\tau\,|\,\chi \equiv \tau \pmod{\mathfrak{p}} \tag{8}$$

genügen. Alle diese Substitutionen bilden eine in Φ enthaltene Gruppe X, die auch ein Teiler der Gruppe Ψ ist, zu der $\mathfrak{p}$ gehört. Denn aus τ erhält man (nach § 112) eine Basisform von $\mathfrak{p}$, wenn man für die Variablen t gewisse lineare Funktionen neuer Variablen mit ganzen rationalen Koeffizienten setzt; wenn also π eine solche Basisform von $\mathfrak{p}$ ist, so folgt aus (8), daß $\pi \mid \chi$ durch π teilbar und daher auch $\mathfrak{p} \mid \chi = \mathfrak{p}$ ist.

Nun genügt τ einer Gleichung $n$ten Grades $f(t) = 0$, deren Koeffizienten ganze Funktionen in R mit den Variablen $t_1, t_2, \ldots$ sind, und es ist, wenn t eine neue Variable bedeutet, und $\tau_1, \tau_2, \ldots \tau_n$ die zu τ konjugierten Formen sind,

$$(9) \qquad f(t, t_1, t_2 \ldots) = f(t) = (t - \tau_1)(t - \tau_2) \ldots (t - \tau_n).$$

Hieraus ergibt sich, wenn wir in die Potenz p erheben und die Bezeichnung von § 110 ($t^p \equiv t \bmod p$) anwenden:

$$(10) \qquad \begin{aligned} f(t, t_1, t_2 \ldots) &\equiv [f(t)]^p \\ &\equiv (t^p - \tau_1)(t^p - \tau_2) \ldots (t^p - \tau_n) \pmod{\mathfrak{p}}. \end{aligned}$$

Setzen wir nun $t = \tau$ in (10), so verschwindet $f(t)$ und es folgt, daß einer der Faktoren des letzten Produktes durch $\mathfrak{p}$ teilbar sein muß. Dies aber kann so ausgedrückt werden, daß es in Φ eine Substitution ψ_0 gibt, die der Bedingung

$$(11) \qquad \tau^p \equiv \tau \mid \psi_0 \pmod{\mathfrak{p}}$$

genügt.

Aus τ entstehen alle ganzen Zahlen in Ω, wenn man für die Variablen ganze rationale Zahlen setzt. Wendet man dann noch den Fermatschen Lehrsatz für rationale Zahlen an, so erkennt man, daß die Kongruenzen (8) und (11) gleichbedeutend sind mit den für jede Zahl ω in $\mathfrak{o}$ gültigen Formeln

$$(12) \qquad \omega \mid \chi \equiv \omega, \qquad \omega^p \equiv \omega \mid \psi_0 \pmod{\mathfrak{p}}.$$

Aus der zweiten Kongruenz (12) folgt, daß, wenn ω durch $\mathfrak{p}$ teilbar ist, immer auch $\omega \mid \psi_0$ durch $\mathfrak{p}$ teilbar sein muß. Folglich ist $\mathfrak{p} \mid \psi_0$ durch $\mathfrak{p}$ teilbar, und als Primideal $= \mathfrak{p}$. Daraus folgt, daß ψ_0 in der Gruppe Ψ enthalten ist.

Wir verstehen jetzt unter γ eine Primitivwurzel von $\mathfrak{p}$ und wenden den Satz § 110, 2. an, nach dem

$$(13) \qquad (t - \gamma)(t - \gamma^p) \ldots (t - \gamma^{p^{f-1}}) \equiv F(t) \pmod{\mathfrak{p}},$$

wo $F(t)$ eine ganze Funktion von t in R ist.

Daraus folgt:

$$F(\gamma) \equiv 0 \pmod{\mathfrak{p}},$$

und wenn nun ψ irgend eine Substitution aus $\mathfrak{P}$ ist:

$$F(\gamma \mid \psi) \equiv 0 \pmod{\mathfrak{p}}.$$

Daraus ergibt sich aber, daß $\gamma \mid \psi$ mit einer der in (13) vorkommenden Potenzen von γ nach $\mathfrak{p}$ kongruent sein muß, also etwa:

$$\gamma^{p^\nu} \equiv \gamma \mid \psi \pmod{\mathfrak{p}}. \tag{14}$$

Nun ist jede durch $\mathfrak{p}$ nicht teilbare Zahl ω in $\mathfrak{o}$ mit einer Potenz von γ kongruent, und wenn man also (14) zu dieser Potenz erhebt, so folgt:

$$\omega^{p^\nu} \equiv \omega \mid \psi \pmod{\mathfrak{p}}, \tag{15}$$

und diese Kongruenz gilt offenbar auch noch, wenn ω durch $\mathfrak{p}$ teilbar ist, also für alle Zahlen in $\mathfrak{o}$.

Wenn wir nun die zweite der Kongruenzen (12) νmal nacheinander anwenden, so folgt:

$$\omega^{p^\nu} \equiv \omega \mid \psi_0^\nu \pmod{\mathfrak{p}}, \tag{16}$$

also

$$\omega \mid \psi \equiv \omega \mid \psi_0^\nu \pmod{\mathfrak{p}}, \tag{17}$$

und wenn man hierin ω durch $\omega \mid \psi^{-1}$ ersetzt und auf (17) die Substitution ψ^{-1} anwendet:

$$\omega \equiv \omega \mid \psi^{-1}\psi_0^\nu \equiv \omega \mid \psi_0^\nu \psi^{-1} \pmod{\mathfrak{p}}. \tag{18}$$

Es sind also sowohl $\psi^{-1}\psi_0^\nu$ als auch $\psi_0^\nu \psi^{-1}$ in X enthalten woraus sich ergibt, daß jede Substitution ψ aus $\mathfrak{P}$ in einem der Systeme (Nebengruppen)

$$X,\ X\psi_0,\ X\psi_0^2, \ldots$$

enthalten ist.

Wenn $\mathfrak{p}$ vom Grade f ist, so ist nach dem Fermatschen Satze (§ 107):

$$\omega^{p^f} \equiv \omega \pmod{\mathfrak{p}},$$

und p^f ist die niedrigste Potenz mit positivem Exponenten, die dieser Kongruenz für alle ω genügt. Setzt man also in (16) $\nu = f$, so folgt, daß ψ_0^f in X enthalten ist, daß aber keine Potenz von ψ_0 mit niedrigerem positiven Exponenten diese Eigenschaft hat. Die Nebengruppen

$$X,\ X\psi_0,\ X\psi_0^2, \ldots X\psi_0^{f-1}$$

sind also alle voneinander verschieden und es ergibt sich:

$$\mathfrak{P} = X + X\psi_0 + X\psi_0^2 + \cdots + X\psi_0^{f-1}.$$

Nun ist aber die Gesamtheit der Substitutionen $\psi_0^\nu X$ für $\nu = 0, 1, 2 \ldots f - 1$ mit $X\psi_0^\nu$ identisch; denn aus (12) ergibt sich, wenn χ ein beliebiges Element aus X ist:

$$\omega \mid \psi_0^{-\nu} \chi \psi_0^\nu \equiv (\mathrm{mod}\ \mathfrak{p})$$

und folglich ist $\psi_0^{-\nu} \chi \psi_0^\nu$ in X enthalten. Es ist also

$$\psi_0^\nu X = X \psi_0^\nu, \tag{19}$$

woraus folgt, daß X ein Normalteiler von Ψ ist.

Da, wie wir oben gesehen haben, Ψ vom Grade gf ist, so ist X vom Grade g. Die Gruppe $\Psi \mid X$ ist zyklisch und vom Grade f.

Die Zahl g ist nach § 111 nur für solche Primzahlen, die in der Grundzahl aufgehen, also für eine endliche Anzahl, größer als 1. Diese Primzahlen heißen kritische Primzahlen und g ihr Gewicht. Die Gruppe X heißt nach Hilbert die Trägheitsgruppe, Ψ die Zerlegungsgruppe von $\mathfrak{p}$.

Fünfzehnter Abschnitt.

Anwendung auf Kreisteilungskörper.

§ 113.

Zerlegung der Primzahl q in Primfaktoren im Kreisteilungskörper $\Omega_{q^\varkappa}$.

Wir machen eine Anwendung der allgemeinen Theorie der algebraischen Zahlen auf die Kreisteilungskörper.

Die Betrachtungen, die wir zunächst anzustellen haben, lassen sich mit kleinen Modifikationen auf jeden vollen Kreisteilungskörper anwenden. Der Einfachheit halber beschränken wir uns hier aber auf den einfacheren Fall, daß der Grad der Einheitswurzel eine Primzahlpotenz ist.

Es sei q eine natürliche Primzahl (mit Einschluß von 2) und

$$m = q^\varkappa \tag{1}$$

eine Potenz von q, deren positiver Exponent $\varkappa$ für $q = 2$ größer als 1 vorausgesetzt wird.

Es sei ferner r eine primitive $m$te Einheitswurzel und Ω_m der volle Kreisteilungskörper für den Exponenten m, dessen Grad

$$\mu = \varphi(m) = q^{\varkappa-1}(q-1) \tag{2}$$

ist (§ 73).

Wir bezeichnen durchweg mit n die μ Zahlen eines vollen Restsystems für den Modul m mit Ausschluß der durch q teilbaren Zahlen, und es sind dann die μ Zahlen r^n die Wurzeln der irreduziblen Gleichung μten Grades $f(x) = 0$, worin

$$f(x) = \frac{x^m - 1}{x^{\frac{m}{q}} - 1} = x^{q^{\varkappa-1}(q-1)} + x^{q^{\varkappa-1}(q-2)} + \cdots + 1 \tag{3}$$

zu setzen ist; r ist daher eine ganze Zahl in Ω_m, und ihre Norm ist $= 1$; folglich ist r eine Einheit. Ω_m ist ein Normalkörper, der durch die μ Substitutionen

$$s_n = (r, r^n)$$

in sich selber übergeht. Diese μ Substitutionen s_n bilden eine Abelsche Gruppe $\mathfrak{N}$, welche die Galoissche Gruppe des Körpers Ω_m ist. Die Zahlen n oder deren Reste nach dem Modul m bilden bei der Zusammensetzung durch Multiplikation gleichfalls eine Abelsche Gruppe, die mit der Gruppe $\mathfrak{N}$ isomorph ist und die gleichfalls mit $\mathfrak{N}$ bezeichnet werden soll.

Die Funktion $f(x)$ läßt sich in die μ Faktoren $x - r^n$ zerlegen, und wir setzen daher

$$f(x) = \overset{n}{\Pi}(x - r^n). \tag{4}$$

Setzen wir darin $x = 1$, und beachten, daß [nach (3)] $f(1) = q$ ist, so folgt:

$$q = \overset{n}{\Pi}(1 - r^n). \tag{5}$$

Die Zahlen

$$\sigma_n = 1 - r^n, \quad \sigma = 1 - r \tag{6}$$

sind ganze Zahlen des Körpers Ω_m, und die Formel (5) zeigt zunächst, daß q im Körper Ω_m in μ Faktoren σ_n zerlegbar ist.

Wir beweisen zunächst, daß die μ Zahlen σ_n miteinander assoziiert sind. Bedeuten n, n' zwei durch q nicht teilbare Zahlen, so können wir eine natürliche Zahl a so bestimmen, daß

$$n' \equiv an \pmod{m}$$

wird. Dann ist aber

$$\frac{\sigma_{n'}}{\sigma_n} = \frac{1 - r^{an}}{1 - r^n} = 1 + r^n + r^{2n} + \cdots + r^{(a-1)n}$$

eine ganze Zahl, also $\sigma_{n'}$ durch σ_n teilbar. Da hierin n mit n' vertauscht werden kann, so ist auch σ_n durch $\sigma_{n'}$ teilbar, also sind beide Zahlen assoziiert (§ 98). Wenn daher ε eine Einheit bedeutet, so ist nach (5):

$$q = \varepsilon\sigma^\mu, \quad N(\sigma) = q. \tag{7}$$

Es ist also die natürliche Primzahl q assoziiert mit der μten Potenz einer ganzen Zahl σ im Körper Ω_m. Diese Zahl σ ist aber auch im Körper Ω_m noch Primzahl, und zwar vom ersten Grade. Denn hat σ irgend einen Teiler σ', so muß die Norm von σ' ein Teiler der Norm von σ sein, also, wenn σ' keine Einheit ist, so ist $N(\sigma') = q$. Folglich ist die Norm von $\sigma : \sigma'$ gleich 1, d. h. $\sigma : \sigma'$ ist eine Einheit und σ mit σ' assoziiert. Wir haben also den ersten Satz:

I. Die natürliche Primzahl q ist in dem vollen Kreisteilungskörper Ω_m mit der μten Potenz einer Primzahl ersten Grades assoziiert.

Nach § 69 ist die Diskriminante D der Kreisteilungsgleichung, also das Produkt aller Differenzen $(r^n - r^{n'})$, eine Potenz von q, nämlich:

$$D = (-1)^{\frac{\mu}{2}} q^{q^{\varkappa-1}[\varkappa(q-1)-1]}. \tag{8}$$

Hieraus ergibt sich der Satz:

II. Die Zahlen

$$1,\ r,\ r^2 \ldots r^{\mu-1} \tag{9}$$

bilden eine Minimalbasis von Ω_m.

Um dies zu zeigen, genügt nach § 103 der Nachweis, daß die ganze Zahl in Ω_m

$$\omega = x_0 + x_1 r + x_2 r^2 + \cdots + x_{\mu-1} r^{\mu-1}, \tag{10}$$

worin $x_0, x_1 \ldots x_{\mu-1}$ ganze rationale Zahlen sind, nur dann durch eine rationale Primzahl p teilbar sein kann, wenn die Zahlen $x_0, x_1, \ldots x_{\mu-1}$ alle durch p teilbar sind.

Nehmen wir also an, es sei ω durch p teilbar; dann sind alle Zahlen ω_n, die aus ω durch eine der μ Substitutionen (r, r^n) entstehen, durch p teilbar, und wir erhalten also aus (10) ein System von μ Gleichungen, die in bezug auf die x_i linear sind:

$$\omega_n = x_0 + x_1 r^n + x_2 r^{2n} + \cdots + x_{\mu-1} r^{(\mu-1)n}. \tag{11}$$

Die Determinante dieses Systems, d. h. das Differenzenprodukt der r^n, ist aber nach § 102 gleich der Quadratwurzel $\sqrt{D}$, und wenn wir also das System der Gleichungen (11) auflösen, so folgt, daß die sämtlichen rationalen Zahlen Dx_i durch p teilbar sein müssen.

Ist nun p nicht $= q$, so ist D nicht durch p teilbar, und sämtliche x_i müssen durch p teilbar sein.

Ist aber $p = q$, so setzen wir

$$\varphi(t) = x_0 + x_1 t + x_2 t^2 + \cdots + x_{\mu-1} t^{\mu-1},$$

so daß $\omega_n = \varphi(r^n)$ wird, und setzen darin nach (6) $r = 1 - \sigma$. Dann ist nach der Taylorschen Entwickelung:

$$(12)\qquad \omega = \varphi(1-\sigma) = \varphi(1) - \sigma\varphi'(1) + \frac{\sigma^2}{1.2}\varphi''(1) - \cdots - \frac{\sigma^{\mu-1}}{\Pi(\mu-1)}\varphi^{(\mu-1)}(1),$$

und hierin sind die Koeffizienten

$$(13)\qquad \varphi(1),\ \varphi'(1),\ \tfrac{1}{2}\varphi''(1) \cdots \frac{1}{\Pi(\mu-1)}\varphi^{(\mu-1)}(1)$$

ganze rationale Zahlen, nämlich

$$\frac{1}{\Pi(\mu-1)}\varphi^{(\mu-1)}(1) = x_{\mu-1},$$

$$\frac{1}{\Pi(\mu-2)}\varphi^{(\mu-2)}(1) = x_{\mu-2} + (\mu-1)\,x_{\mu-1},$$

$$\frac{1}{\Pi(\mu-3)}\varphi^{(\mu-3)}(1) = x_{\mu-3} + (\mu-2)\,x_{\mu-2} + \frac{(\mu-1)(\mu-2)}{1.2}\,x_{\mu-1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\varphi(1) = x_0 + x_1 + \cdots\cdots\cdots\cdots + x^{\mu-1}.$$

Da nun ω durch q und folglich durch σ^μ teilbar ist, so muß $\varphi(1)$ nach (12) durch q teilbar sein. Da aber $\mu > 1$ und folglich der Quotient $q : \sigma$ noch durch σ teilbar ist, so folgt, daß auch

$$\varphi'(1) - \frac{\sigma}{1.2}\varphi''(1) + \cdots$$

durch σ, folglich $\varphi'(1)$ durch q teilbar ist, und wenn man so weiter schließt, ergibt sich, daß die sämtlichen rationalen Zahlen (13) durch q teilbar sind. Folglich sind auch die $x_{\mu-1}, x_{\mu-2}, \ldots x_0$ durch q teilbar. Damit ist der Satz II. vollständig bewiesen.

Anstatt der Reihe (9) kann man auch eine beliebige Reihe von μ aufeinander folgenden Potenzen von r als Basis von $\mathfrak{o}$ nehmen:

$$(14)\qquad r^a,\ r^{a+1},\ r^{a+2} \ldots r^{a+\mu-1},$$

die man erhält, wenn man alle Glieder der Reihe (9) mit r^a multipliziert.

Die Richtigkeit dieser Bemerkung ergibt sich daraus, daß die $r^{\pm a}$ Einheiten sind, und daß also ω und $r^a\omega$ stets gleichzeitig ganze oder gebrochene Zahlen sind.

Beispielsweise bilden die Zahlen

$$(15)\qquad r^{-\frac{1}{2}\mu+1},\ r^{-\frac{1}{2}\mu+2} \ldots r^{-1},\ 1,\ r,\ r^2 \ldots r^{\frac{1}{2}\mu}$$

eine Basis von $\mathfrak{o}$.

Damit ist auch die Grundzahl Δ des Körpers Ω_m bestimmt. Sie ist nach der allgemeinen Definition (§ 103) gleich der Diskriminante

$$\Delta(1, r, r^2 \dots r^{\mu-1}),$$

d. h. gleich der Diskriminante der Kreisteilungsgleichung.

§ 114.

Die Primideale im Körper Ω_m.

Wir haben im § 113 gesehen, daß die natürliche Primzahl q die μte Potenz einer Primzahl σ im Körper Ω_m ist.

Um alle Primideale, die im Körper Ω_m existieren, zu ermitteln, sind also noch sämtliche von q verschiedene natürliche Primzahlen p in Primfaktoren zu zerlegen, und es sind darunter die nicht assoziierten auszusuchen.

Die Grundlage für diese Untersuchung bilden die Sätze des § 112.

Jede Zahl ω des Körpers Ω_m geht durch eine der μ Substitutionen

$$s_n = (r, r^n)$$

in eine bestimmte andere Zahl ω_n über, die gleichfalls in Ω_m enthalten ist. Ist

$$(1) \qquad \omega = \omega_1 = a_0 + a_1 r + a_2 r^2 + \cdots + a_{\mu-1} r^{\mu-1},$$

so ist

$$(2) \qquad \omega_n = a_0 + a_1 r^n + a_2 r^{2n} + \cdots + a_{\mu-1} r^{(\mu-1)n}.$$

Wenn eine dieser Zahlen ω_n eine ganze Zahl ist, wie wir jetzt annehmen wollen, so sind es auch alle anderen, und dies tritt immer dann und nur dann ein, wenn die rationalen Zahlen $a_0, a_1, \dots a_{\mu-1}$ ganz sind.

Sind h, k irgend zwei Exponenten, so ist

$$r^h - r^k = r^k(r^{h-k} - 1),$$

also mit σ assoziiert, wenn $\mathfrak{p}$ ein in p aufgehendes Primideal bedeutet, durch $\mathfrak{p}$ nicht teilbar, außer wenn $h \equiv k \pmod m$ ist. Daraus ergibt sich, daß zwei Zahlen ω_h, ω_k nur dann nach dem Modul $\mathfrak{p}$ kongruent sein können, wenn $h \equiv k \pmod m$, also $\omega_h = \omega_k$ ist, und daß also die Gruppe X des § 112, deren Substitutionen χ der Bedingung

$$\omega \mid \chi \equiv \omega \pmod{\mathfrak{p}}$$

genügen, nur die identische Substitution enthält. Daraus folgt aber, daß $g = 1$, d. h. daß p nicht durch das Quadrat eines Primideals teilbar ist, in Übereinstimmung mit dem Satz von Dedekind, nach dem nur die Primzahl q in Ω_m kritisch ist.

Bilden wir aus (1) die $p$te Potenz von ω, und beachten den Fermatschen Lehrsatz für rationale Zahlen $[a^p \equiv a \pmod p]$, so ergibt sich:

$$\omega^p \equiv a_0 + a_1 r^p + a_2 r^{2p} + \cdots + a_{\mu-1} r^{(\mu-1)p} \pmod p,$$

oder nach (2):

$$\omega^p \equiv \omega_p \pmod p. \tag{3}$$

Dadurch ist die Gruppe Ψ bestimmt, zu der das Ideal $\mathfrak{p}$ gehört. Es ist nämlich nach (3) auch

$$\omega^p \equiv \omega_p \pmod{\mathfrak{p}},$$

daher ist $\psi_0 = (r, r^p)$, und die Gruppe Ψ besteht nach § 112 aus den Potenzen dieser Substitution, soweit sie voneinander verschieden sind.

Ist daher f der kleinste positive Exponent, für den

$$p^f \equiv 1 \pmod m \tag{4}$$

ist, d. h. gehört p zu dem Exponenten f für den Modul m, so ist

$$\Psi = (r, r), (r, r^p), (r, r^{p^2}) \ldots (r, r^{p^{f-1}}), \tag{5}$$

und ist vom $f$ten Grade. Demnach ist auch $\mathfrak{p}$ ein Primideal $f$ten Grades und

$$N(\mathfrak{p}) = p^f. \tag{6}$$

Die Anzahl e der voneinander verschiedenen konjugierten Primfaktoren von p ist $\mu : f$, und wir haben also den Satz:

III. Ist p eine von q verschiedene Primzahl, die für den Modul m zum Exponenten f gehört, ist ferner $\mu = \varphi(m) = ef$, so zerfällt p im Körper Ω_m in e voneinander verschiedene konjugierte Primfaktoren $f$ten Grades.

Die Gruppe Ψ ist isomorph mit einer in $\mathfrak{N}$ enthaltenen Gruppe nach dem Modul m genommener ganzer rationaler Zahlen, die wir mit $\mathfrak{A}$ bezeichnen, die aus allen, einer Kongruenz

$$a \equiv p^h \pmod m \tag{7}$$

genügenden Zahlen a besteht, und die wir daher symbolisch durch

$$\mathfrak{A} \equiv p^h \pmod m \tag{8}$$

darstellen können, wenn der Exponent h alle ganzzahligen Werte durchläuft.

Wir können also, wenn wir die Zahlen $\xi_1, \xi_2, \ldots \xi_e$ aus $\mathfrak{R}$ passend auswählen,

$$\mathfrak{R} = \mathfrak{A}\xi_1 + \mathfrak{A}\xi_2 + \mathfrak{A}\xi_3 + \cdots + \mathfrak{A}\xi_e \tag{9}$$

setzen, und jeder dieser Nebengruppen entspricht eines der konjugierten Ideale $\mathfrak{p}_{\xi_1}, \mathfrak{p}_{\xi_2} \ldots \mathfrak{p}_{\xi_e}$, die alle voneinander verschieden sind, und deren Produkt p ist.

Wir setzen also

$$p = \mathfrak{p}_{\xi_1} \cdot \mathfrak{p}_{\xi_2} \cdots \mathfrak{p}_{\xi_e}, \tag{10}$$

und bemerken noch, daß $\mathfrak{p}_{a\xi} = \mathfrak{p}_\xi$ ist, wenn a in $\mathfrak{A}$ enthalten ist.

Ist m ungerade und c eine primitive Wurzel von m, so ist $p \equiv c^\gamma$, (mod m) für einen gewissen Exponenten γ, und e ist der größte gemeinschaftliche Teiler von μ und γ. Die Gruppe $\mathfrak{A}$ besteht dann aus allen Zahlen von der Form c^{eh}, und für $\xi_1, \xi_2 \ldots \xi_e$ kann man die Zahlen

$$1, c, c^2 \ldots c^{e-1} \tag{11}$$

wählen. Die e Primideale (10) entsprechen den Kreisteilungsperioden.

Durch eine Substitution (r, r^n) vertauschen sich die Primfaktoren $\mathfrak{p}$ nach dem Gesetz

$$(r, r^n) = (\mathfrak{p}_{\xi_i}, \mathfrak{p}_{n\xi_i}). \tag{12}$$

§ 115.

Darstellung der Primfaktoren von p.

Zur Darstellung der Primfunktionale des Körpers Ω_m können wir ein Verfahren anwenden, das wir im § 110 kennen gelernt haben, welches sich hier infolge des Umstandes, daß

$$1, r, r^2 \ldots r^{\mu-1}$$

eine Basis von $\mathfrak{o}$ ist, wesentlich vereinfacht.

Die natürliche Primzahl q ist, wie wir schon gesehen haben, die μte Potenz einer im Körper Ω_m existierenden Primzahl σ; mit dieser brauchen wir uns nicht weiter zu beschäftigen, und betrachten daher hier nur die von q verschiedenen Primzahlen p. Es sei $\mathfrak{p}$ ein Primfaktor von p, und $\mathfrak{A}$ habe dieselbe Bedeutung wie in § 114, (8). Wenn p zum Exponenten f gehört und

$$ef = \varphi(m) = \mu$$

ist, so besteht die Gruppe $\mathfrak{A}$ aus den Zahlen

$$1,\ p,\ p^2 \dots p^{f-1}.$$

Wenn nun

$$f(t) = N(t - r) = \prod_{n}^{n} (t - r^n)$$

das Polynom von t vom Grade μ bedeutet, dessen Wurzeln die μ Größen r^n sind, so können wir dies so in Faktoren zerlegen, daß jeder Faktor einer der Nebengruppen von $\mathfrak{A}$ entspricht. Setzen wir nämlich

$$F_1(t) = \prod_{0,f-1}^{h} (t - r^{p^h}), \tag{1}$$

und allgemein für jedes durch q nicht teilbare n:

$$F_n(t) = \prod_{0,f-1}^{h} (t - r^{n p^h}), \tag{2}$$

so ist $F_n(t)$ mit $F_{n'}(t)$ identisch, wenn n und n' in dieselbe Nebengruppe $\mathfrak{A}\xi_i$ gehören. Wenn wir also mit $\xi_1, \xi_2, \dots \xi_e$ das im § 114, (9) angewandte Zahlensystem verstehen, so ist

$$f(t) = F_{\xi_1}(t)\, F_{\xi_2}(t) \dots F_{\xi_e}(t). \tag{3}$$

Nun ist aber nach dem binomischen Lehrsatze

$$(t - r^{p^h})^p \equiv (t^p - r^{p^{h+1}}) \pmod{p},$$

und wenn wir dies auf jeden Faktor von $F_n(t)$ anwenden und beachten, daß p^{h+1} nach dem Modul m dieselbe Zahlenreihe durchläuft, wie p^h, so folgt:

$$[F_n(t)]^p \equiv F_n(t) \pmod{p},$$

und diese Kongruenz gilt dann natürlich auch für den Modul $\mathfrak{p}$.

Dies ist aber das Kennzeichen dafür, daß $F_n(t)$ mit einem Polynom $P_n(t)$ mit ganzen rationalen Zahlenkoeffizienten nach dem Modul $\mathfrak{p}$ kongruent ist (§ 107, 4.), also:

$$F_n(t) \equiv P_n(t) \pmod{\mathfrak{p}}. \tag{4}$$

Setzen wir dies in (3) ein, so ergibt sich zunächst eine Kongruenz nach dem Modul $\mathfrak{p}$, die aber, da es eine Kongruenz zwischen rationalen Funktionen ist, auch für den Modul p bestehen muß, also

$$f(t) \equiv P_{\xi_1}(t)\, P_{\xi_2}(t) \dots P_{\xi_e}(t) \pmod{p}. \tag{5}$$

Nach der Definition (1), (4) ist

$$P_1(r) \equiv 0 \pmod{\mathfrak{p}}, \tag{6}$$

und die sämtlichen Wurzeln dieser Kongruenz sind

$$r,\ r^p \dots r^{p^{f-1}}. \tag{7}$$

Ebenso hat die Kongruenz

$$P_n(t) \equiv 0 \pmod{\mathfrak{p}}$$

die Wurzeln

$$r^n,\ r^{np} \dots r^{np^{f-1}},$$

und diese Wurzeln sind, wenn man n das System der Zahlen $\xi_1, \xi_2 \dots \xi_e$ durchlaufen läßt, alle inkongruent nach dem Modul $\mathfrak{p}$.

Macht man in der Kongruenz

$$F_1(t) \equiv P_1(t) \pmod{\mathfrak{p}}$$

die Substitution (r, r^n), so geht $F_1(t)$ in $F_n(t)$, $\mathfrak{p}$ in das konjugierte Ideal $\mathfrak{p}_n$ über, und $P_1(t)$ bleibt als rationale Funktion ungeändert. Wir haben also

$$F_n(t) \equiv P_1(t) \pmod{\mathfrak{p}_n}, \tag{8}$$

und wenn wir hierin $t = r$ setzen, so folgt:

$$F_n(r) \equiv P_1(r) \pmod{\mathfrak{p}_n}; \tag{9}$$

$F_n(r)$ ist aber, wenn n nicht in $\mathfrak{A}$ enthalten ist, relativ prim zu p, also nicht durch $\mathfrak{p}_n$ teilbar.

Es ist also $P_1(r)$ nach (9) durch $\mathfrak{p}_1$, aber durch keinen der mit $\mathfrak{p}_1$ konjugierten Primfaktoren teilbar, und es folgt, da p nicht durch $\mathfrak{p}^2$ teilbar ist:

1. Der Primfaktor $\mathfrak{p}_1$ ist der größte gemeinschaftliche Teiler von p und $P_1(r)$; in gleicher Weise ergibt sich, daß $\mathfrak{p}_n$ der größte gemeinschaftliche Teiler von p und $P_1(r^n)$ ist.

Wenn man $t = r^n$ in (8) einsetzt, so erhält man:

$$P_1(r^n) \equiv 0 \pmod{\mathfrak{p}_n} \tag{10}$$

und aus (4) erhält man:

$$P_n(r^n) \equiv 0 \pmod{\mathfrak{p}};$$

macht man aber darin die Substitution $(r, r^{n'})$, so folgt:

$$P_n(r^{nn'}) \equiv 0 \pmod{\mathfrak{p}_{n'}},$$

also wenn $nn' \equiv 1 \bmod m$ genommen wird:

$$P_n(r) \equiv 0 \pmod{\mathfrak{p}_{n'}}. \tag{11}$$

Man kann also die verschiedenen konjugierten Primideale entweder durch die Funktion P_1 in R für verschiedene Argument-

werte oder durch verschiedene Funktionen P_n in R für ein und dasselbe Argument r definieren.

Wir wollen den Fall noch etwas näher betrachten, wo $f = 1$ ist, also

$$p \equiv 1 \pmod{m}.$$

In diesem Falle ist $e = \varphi(m)$; die Gruppe $\mathfrak{A}$ reduziert sich auf die Einheit, und die sämtlichen μ konjugierten Faktoren $\mathfrak{p}$ von p sind voneinander verschieden. Die Funktion $P_1(t)$ ist linear, d. h. r ist einer rationalen Zahl nach jedem der Moduln $\mathfrak{p}_n$ kongruent.

Dies ergibt sich auch daraus, daß hier die Anzahl der nach dem Modul $\mathfrak{p}$ inkongruenten Zahlen gleich $N(\mathfrak{p}) = p$ ist, und daß also $0, 1, 2 \ldots p - 1$ ein volles Restsystem ist.

Ist hiernach etwa $r \equiv c \pmod{\mathfrak{p}}$, so muß, da $r^m = 1$ ist, $c^m \equiv 1 \pmod{\mathfrak{p}}$, also auch $\pmod{p}$ sein, und wenn also g eine primitive Wurzel der Primzahl p ist, so muß

$$c \equiv g^{-n\frac{p-1}{m}} \pmod{p} \tag{12}$$

sein, worin n relativ prim zu m ist, weil keine niedrigere als die $m$te Potenz von r mit der Einheit kongruent ist. Lassen wir $\mathfrak{p}$ das System der konjugierten Ideale durchlaufen, so muß n in (12) die Gruppe $\mathfrak{N}$ durchlaufen, und es ist also nur Sache der Bezeichnung, wenn wir festsetzen:

$$r \equiv g^{-\frac{p-1}{m}} \pmod{\mathfrak{p}}. \tag{13}$$

Machen wir darin die Substitution (r, r^n), so wird

$$r^n \equiv g^{-\frac{p-1}{m}} \pmod{\mathfrak{p}_n}, \tag{14}$$

oder, wenn wir n' durch die Kongruenz

$$nn' \equiv 1 \pmod{m} \tag{15}$$

definieren:

$$r \equiv g^{-n'\frac{p-1}{m}} \pmod{\mathfrak{p}_n}. \tag{16}$$

Es ist also in diesem Falle, übereinstimmend mit der allgemeinen Regel, $\mathfrak{p}_n$ der größte gemeinschaftliche Teiler von

$$p \text{ und } \left(r - g^{-n'\frac{p-1}{m}}\right).$$

Durch diese Bestimmung sind die einzelnen Primideale $\mathfrak{p}_n$ genau charakterisiert. Wir erhalten also den Satz:

2. Eine Primzahl p, die nach dem Modul m mit 1 kongruent ist, zerfällt im Körper Ω_m in $\varphi(m)$ voneinander verschiedene Primfaktoren $\mathfrak{p}_n$ vom ersten Grade, die man als größte gemeinschaftliche Teiler von p mit den verschiedenen Zahlen $r - g^{-n'\frac{p-1}{m}}$ erhält, wenn g eine primitive Wurzel von p ist, und n' durch die Kongruenz $nn' \equiv 1 \pmod{m}$ bestimmt ist.

Ist $p - 1$ zwar nicht durch m, wohl aber durch eine Potenz von q, und im Falle $q = 2$ mindestens durch 4 teilbar, so bezeichnen wir mit m_1 den größten gemeinschaftlichen Teiler von m und $p - 1$ und setzen $m = m_1 m_2$. Dann ist

$$\mu = \varphi(m) = m_2 \varphi(m_1),$$

und es ist $f = m_2$ der kleinste positive Exponent, für den

$$p^f \equiv 1 \pmod{m}$$

ist. Demnach zerfällt p im Körper Ω_m in $\varphi(m_1)$ voneinander verschiedene Primfaktoren, ebenso wie im Körper Ω_{m_1}. Es ergibt sich also:

3. Ist $p - 1$ durch q und im Falle $q = 2$ durch 4 teilbar, so ist die Zerlegung von p in Primfaktoren im Körper Ω_m dieselbe, wie im Körper Ω_{m_1}, die nach dem Satze 2. gefunden wird.

SACHREGISTER.