Über die Darstellung definiter Funktionen durch Quadrate.

Von

EDMUND LANDAU in Berlin.

Erster Teil.

Es sei

$$f(x) = a_0 x^{2k} + a_1 x^{2k-1} + \cdots + a_{2k}$$

eine ganze rationale Funktion $2k^{\text{ten}}$ Grades von x mit rationalen Zahlenkoeffizienten. Diese Funktion sei definit, d. h. es sei für alle reellen x

$$f(x) \geqq 0;$$

anders ausgedrückt, es sei

$$a_0 > 0$$

und $f(x)$ habe entweder keine reelle Wurzel oder jede reelle Wurzel in gerader Vielfachheit. Dann folgt aus der Zerlegung von $f(x)$ in Linearfaktoren ohne weiteres eine Darstellung von $f(x)$ als Summe von zwei Quadraten ganzer rationaler Funktionen von x mit reellen Koeffizienten. Denn wenn $f(x)$ r Paare reeller gleicher Wurzeln*) und s Paare konjugiert-komplexer Wurzeln**) besitzt, so erhält man durch Zusammenfassung aller Linearfaktoren in drei Klassen***) eine Zerlegung

$$f(x) = f_1^2(x)\,(f_2(x) + f_3(x)i)\,(f_2(x) - f_3(x)i),$$

wo $f_1(x)$, $f_2(x)$ und $f_3(x)$ reelle Koeffizienten haben und $f_1(x)$ den Grad r, $f_2(x) + f_3(x)i$ den Grad s besitzt, und hieraus folgt weiter

$$f(x) = f_1^2(x)\,(f_2^2(x) + f_3^2(x)) = (f_1(x)\,f_2(x))^2 + (f_1(x)\,f_3(x))^2 = g_1^2(x) + g_2^2(x).$$

Die Zahlenkoeffizienten in $g_1(x)$ und $g_2(x)$ sind reell, brauchen aber nicht rational zu sein.

Eine Darstellung von $f(x)$ durch Quadrate rationalzahliger Funktionen

*) Eine 2α-fache reelle Wurzel wird dabei als α Paare gleicher Wurzeln aufgefaßt. Es kann auch $r = 0$ sein.

**) Ein Paar α-facher konjugiert-komplexer Wurzeln wird dabei als α einfache Paare angesehen. s kann auch $= 0$ sein.

***) Hierbei kommt jeder reelle Linearfaktor in die erste Klasse, und jedes Paar konjugiert-komplexer Linearfaktoren wird auf die beiden anderen Klassen verteilt.

hat zuerst Herr Hilbert*) angegeben. Er hat nämlich bewiesen, daß $f(x)$ stets als Quotient zweier Summen von Quadraten der verlangten Art darstellbar ist.

Darauf habe ich**) den weitergehenden Satz bewiesen: „Jede definite ganze rationale Funktion von x mit rationalen Zahlenkoeffizienten läßt sich als Summe von Quadraten darstellen, so daß die sämtlichen Basen dieser Quadrate ganze rationale Funktionen von x mit rationalen Koeffizienten sind."

In einer späteren Arbeit***) habe ich die Frage aufgeworfen und für die kleinsten Werte von $n = 2k$ in Angriff genommen: Für jeden Grad n diejenige Zahl $N = N(n)$ zu bestimmen, welche durch folgende beiden Eigenschaften charakterisiert ist:

1) Jede definite Funktion†) n^{ten} Grades läßt sich in N Quadrate zerlegen.

2) Nicht jede definite Funktion n^{ten} Grades läßt sich in $N-1$ Quadrate zerlegen.

Ich bewies a. a. O., daß

$$N(0) = 4,$$
$$N(2) = 5,$$
$$N(4) \leqq 6$$

ist.

Alsdann zeigte Herr Fleck††), an meine Methode zur Diskussion der biquadratischen Funktion anknüpfend, daß

$$N(4) = 5$$

ist.

Trotzdem nun $N(4)$ nicht größer ist als $N(2)$, würde man wohl erwarten, daß $N(n)$ mit n über alle Grenzen wächst. Merkwürdigerweise gilt†††) jedoch der allgemeine Satz, dessen Herleitung den Gegenstand des ersten Teiles der vorliegenden Arbeit bildet:

„Jede definite ganze Funktion n^{ten} Grades von x mit rationalen Zahlenkoeffizienten läßt sich als Summe von acht Quadraten ganzer rationalzahliger Funktionen von x darstellen."

*) „Grundlagen der Geometrie", Festschrift zur Feier der Enthüllung des Gauß-Weber-Denkmals in Göttingen, Leipzig, 1899, S. 82—85.

**) „Über die Darstellung definiter binärer Formen durch Quadrate", Mathematische Annalen, Bd. 57, 1903, S. 53—64.

***) „Über die Zerlegung definiter Funktionen in Quadrate", Archiv der Mathematik und Physik, 3te Reihe, Bd. 7, 1904, S. 271—277.

†) Im ersten Teile dieser Arbeit ist durchweg nur von ganzen rationalzahligen Funktionen die Rede, ohne daß dies immer besonders bemerkt wird.

††) „Zur Darstellung definiter binärer Formen als Summen von Quadraten ganzer rationalzahliger Formen", Archiv der Mathematik und Physik, 3te Reihe, Bd. 10, 1906, S. 23—38.

†††) Es sei gleich hier der Leser auf S. 278, Z. 17—25 aufmerksam gemacht.

Mit anderen Worten, es ist stets

$$N(n) \leqq 8,$$

also

$$\limsup_{n=\infty} N(n) \leqq 8,$$

und, da offenbar*)

$$N(n) \leqq N(n+2)$$

ist, so existiert

$$\lim_{n=\infty} N(n)$$

und ist $\leqq 8$**). Welche der vier Zahlen $N = 5, 6, 7, 8$ die Eigenschaft hat, daß jedes $f(x)$ in N Quadrate, aber nicht jedes in $N-1$ Quadrate zerlegbar ist, muß vorläufig dahingestellt bleiben.

Um nun den oben ausgesprochenen Satz über die Zerlegbarkeit von $f(x)$ in acht Quadrate zu beweisen, nehme ich zunächst $f(x)$ irreduzibel im Körper der rationalen Zahlen an. Darauf wird dann der allgemeine Fall leicht zurückführbar sein. Wenn die definite Funktion $f(x)$ irreduzibel ist, so hat die Gleichung

$$f(x) = 0$$

keine mehrfachen Wurzeln, also keine reelle Wurzel, und bestimmt daher einen algebraischen Zahlkörper n^{ten} Grades, der mit sämtlichen konjugierten Körpern imaginär ausfällt. Daraus folgt, wie Herr Hilbert***) gezeigt hat, daß $f(x)$ in der Form darstellbar ist

$$(1) \qquad f(x) = \frac{1 + \{\varphi(x)\}^2 + \{\psi(x)\}^2 + \{\chi(x)\}^2 + \{\varrho(x)\}^2}{f_1(x)},$$

wo $\varphi(x)$, $\psi(x)$, $\chi(x)$, $\varrho(x)$ ganze rationalzahlige Funktionen höchstens $n-1^{\text{ten}}$ Grades sind, also $f_1(x)$ eine definite Funktion höchstens $n-2^{\text{ten}}$ Grades.

Es sei zunächst angegeben, wie Herr Hilbert zur Gleichung (1) gelangt. Er benutzt folgenden Satz ohne genauere Ausführung seines Beweises, der erhebliche Schwierigkeiten bietet und wesentlich auf Herrn Hilberts Theorie der relativquadratischen Zahlkörper†) beruht:

---

*) Denn wenn die definite Funktion n^{ten} Grades $f(x) = a_0 x^n + \cdots + a_n$ $(a_0 > 0)$ nicht in $N(n)-1$ Quadrate zerlegbar ist, so ist offenbar die definite Funktion $n+2^{\text{ten}}$ Grades $x^2 f(x) = a_0 x^{n+2} + \cdots + a_n x^2$ gleichfalls nicht in $N(n)-1$ Quadrate zerlegbar.

**) Mit anderen Worten, für alle hinreichend großen n hat $N(n)$ einen und denselben Wert, der $\leqq 8$ ist.

***) l. c., S. 84—85.

†) „Über die Theorie der relativquadratischen Zahlkörper", Jahresbericht der Deutschen Mathematiker-Vereinigung, Bd. 6, 1899, S. 88—94; „Über die Theorie des relativquadratischen Zahlkörpers", Mathematische Annalen, Bd. 51, 1899, S. 1—127; „Über die Theorie der relativ-Abel'schen Zahlkörper", Nachrichten der Königlichen

„*Jede total positive Zahl eines algebraischen Zahlkörpers*, d. h. jede Zahl des Körpers, deren konjugierte Werte in den reellen konjugierten Körpern positiv sind, *läßt sich als Summe von vier Quadraten gewisser Zahlen des Körpers darstellen*".*)

Im vorliegenden Falle sind alle durch die Gleichung

$$f(x) = 0 \tag{2}$$

bestimmten Körper imaginär; jede Zahl des Körpers $k(\vartheta)$, wo ϑ eine Wurzel von (2) ist, ist also total positiv; daher gibt es insbesondere für -1 eine Zerlegung

$$-1 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2,$$

wo $\alpha, \beta, \gamma, \delta$ ganze oder gebrochene Zahlen in $k(\vartheta)$ sind. Diese sind als ganze rationalzahlige Funktionen $\varphi(\vartheta), \psi(\vartheta), \chi(\vartheta), \varrho(\vartheta)$ von ϑ darstellbar, deren Grad $\leqq n-1$ ist. Aus

$$1 + \{\varphi(\vartheta)\}^2 + \{\psi(\vartheta)\}^2 + \{\chi(\vartheta)\}^2 + \{\varrho(\vartheta)\}^2 = 0$$

folgt, da $f(x)$ irreduzibel ist, daß die ganze rationalzahlige Funktion

$$F(x) = 1 + \{\varphi(x)\}^2 + \{\psi(x)\}^2 + \{\chi(x)\}^2 + \{\varrho(x)\}^2$$

durch $f(x)$ teilbar ist:

$$F(x) = f(x) f_1(x),$$

und dies liefert die Gleichung (1).

Nachdem nun Herr Hilbert auf diesem Wege die Gleichung (1) erhalten hat, zieht er daraus durch vollständige Induktion die Folgerung: Es sei schon bewiesen, daß jede definite Funktion $n-2^{\text{ten}}$ oder niedrigeren Grades als Quotient zweier Quadratsummen darstellbar ist. Dann gilt dies nach (1) für jede irreduzibele definite Funktion n^{ten} Grades, also für jede definite Funktion n^{ten} Grades (da, wie leicht gezeigt wird, jede reduzibele definite Funktion, abgesehen von einem konstanten Faktor, ein Quadrat ist oder als Produkt eines Quadrates**) und einer oder mehrerer irreduzibler definiter Funktionen niedrigeren Grades darstellbar ist).

Ich lege den folgenden Schlüssen auch die Gleichung (1) zugrunde, schließe jedoch folgendermaßen weiter: Es ist, wenn

$$\alpha_1(x) = 1, \quad \alpha_2(x) = \varphi(x), \quad \alpha_3(x) = \psi(x), \quad \alpha_4(x) = \chi(x), \quad \alpha_5(x) = \varrho(x),$$
$$\alpha_6(x) = 0, \quad \alpha_7(x) = 0, \quad \alpha_8(x) = 0$$

gesetzt wird,

Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, 1898, S. 370—399 und Acta mathematica, Bd. 26, 1902, S. 99—131.

*) Diesen Satz erwähnt Herr Hilbert auch in seinem Artikel „Theorie der algebraischen Zahlkörper" in der „Encyklopädie der mathematischen Wissenschaften", Bd. 1, S. 696.

**) welches eventuell $= 1$ ist.

$$(3)\quad f(x) = \frac{\alpha_1^2(x)+\alpha_2^2(x)+\alpha_3^2(x)+\alpha_4^2(x)+\alpha_5^2(x)+\alpha_6^2(x)+\alpha_7^2(x)+\alpha_8^2(x)}{f_1(x)},$$

$$(4)\quad \alpha_1^2(x)+\cdots+\alpha_8^2(x) = f(x)f_1(x),$$

wo $f_1(x)$ kleineren Grad hat als $f(x)$*).

Nun dividiere ich, wenn $f_1(x)$ nicht konstant ist, jede der acht Basen im Zähler von (3) durch $f_1(x)$; dies ergibt ein Gleichungssystem

$$(5)\quad \begin{cases} \alpha_1(x) = q_1(x)f_1(x)+\beta_1(x), \\ \cdots\cdots\cdots\cdots \\ \alpha_8(x) = q_8(x)f_1(x)+\beta_8(x), \end{cases}$$

wo $\beta_1(x), \cdots, \beta_8(x)$ geringeren Grad haben als $f_1(x)$. $\beta_1(x), \cdots, \beta_8(x)$ sind nicht sämtlich $= 0$, da sonst

$$f(x) = f_1(x)(q_1^2(x)+\cdots+q_8^2(x)),$$

also $f(x)$ reduzibel wäre.**)

Aus (4) und (5) folgt modulo f_1

$$\beta_1^2+\cdots+\beta_8^2 \equiv \alpha_1^2+\cdots+\alpha_8^2 = ff_1 \equiv 0,$$

also

$$(6)\quad \beta_1^2+\cdots+\beta_8^2 = f_1f_2,$$

wo f_2 eine ganze rationalzahlige definite Funktion ist, die geringeren Grad hat als f_1 und nicht identisch verschwindet.

Nun besteht die bekannte Identität

$$\begin{aligned}(7)\quad &(\alpha_1^2+\alpha_2^2+\alpha_3^2+\alpha_4^2+\alpha_5^2+\alpha_6^2+\alpha_7^2+\alpha_8^2)(\beta_1^2+\beta_2^2+\beta_3^2+\beta_4^2+\beta_5^2+\beta_6^2 \\ &\qquad +\beta_7^2+\beta_8^2) \\ &= (\alpha_1\beta_1+\alpha_2\beta_2+\alpha_3\beta_3+\alpha_4\beta_4+\alpha_5\beta_5+\alpha_6\beta_6+\alpha_7\beta_7+\alpha_8\beta_8)^2 \\ &+(-\alpha_1\beta_2+\alpha_2\beta_1+\alpha_3\beta_4-\alpha_4\beta_3+\alpha_5\beta_6-\alpha_6\beta_5+\alpha_7\beta_8-\alpha_8\beta_7)^2 \\ &+(-\alpha_1\beta_3-\alpha_2\beta_4+\alpha_3\beta_1+\alpha_4\beta_2+\alpha_5\beta_7-\alpha_6\beta_8-\alpha_7\beta_5+\alpha_8\beta_6)^2 \\ &+(-\alpha_1\beta_4+\alpha_2\beta_3-\alpha_3\beta_2+\alpha_4\beta_1-\alpha_5\beta_8-\alpha_6\beta_7+\alpha_7\beta_6+\alpha_8\beta_5)^2 \\ &+(-\alpha_1\beta_5-\alpha_2\beta_6-\alpha_3\beta_7+\alpha_4\beta_8+\alpha_5\beta_1+\alpha_6\beta_2+\alpha_7\beta_3-\alpha_8\beta_4)^2 \\ &+(-\alpha_1\beta_6+\alpha_2\beta_5+\alpha_3\beta_8+\alpha_4\beta_7-\alpha_5\beta_2+\alpha_6\beta_1-\alpha_7\beta_4-\alpha_8\beta_3)^2 \\ &+(-\alpha_1\beta_7-\alpha_2\beta_8+\alpha_3\beta_5-\alpha_4\beta_6-\alpha_5\beta_3+\alpha_6\beta_4+\alpha_7\beta_1+\alpha_8\beta_2)^2 \\ &+(-\alpha_1\beta_8+\alpha_2\beta_7-\alpha_3\beta_6-\alpha_4\beta_5+\alpha_5\beta_4+\alpha_6\beta_3-\alpha_7\beta_2+\alpha_8\beta_1)^2.\end{aligned}$$

Diese Identität wende ich auf die vorliegende Bedeutung der Zeichen

*) Mit anderen Worten, ich benutze gar nicht den Hilbertschen Satz

$$-1 = \alpha^2+\beta^2+\gamma^2+\delta^2$$

in vollem Umfange, sondern nur die eventuell leichter zu beweisende Tatsache, daß sich -1 im Körper $k(\vartheta)$ als Summe von *sieben* Quadratzahlen darstellen läßt.

**) Übrigens ist hier $\beta_1(x) = 1$; doch habe ich für die in der Folge notwendige Wiederholung des Schlußverfahrens die Begründung des Textes angegeben.

$\alpha_1, \cdots, \beta_8$ an, d. h. ich multipliziere die Gleichungen (4) und (6) und stelle das Produkt nach (7) als Summe von acht Quadraten dar. Die Betrachtung der Basen dieser acht Quadrate zeigt, daß jede derselben durch f_1 teilbar ist; z. B. ist modulo f_1

$$\begin{gathered}\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 + \alpha_4\beta_4 + \alpha_5\beta_5 + \alpha_6\beta_6 + \alpha_7\beta_7 + \alpha_8\beta_8 \\ \equiv \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 + \alpha_5^2 + \alpha_6^2 + \alpha_7^2 + \alpha_8^2 = ff_1 \equiv 0, \\ -\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_3\beta_4 - \alpha_4\beta_3 + \alpha_5\beta_6 - \alpha_6\beta_5 + \alpha_7\beta_8 - \alpha_8\beta_7 \\ \equiv -\alpha_1\alpha_2 + \alpha_2\alpha_1 + \alpha_3\alpha_4 - \alpha_4\alpha_3 + \alpha_5\alpha_6 - \alpha_6\alpha_5 + \alpha_7\alpha_8 - \alpha_8\alpha_7 = 0.\end{gathered}$$

Wenn daher die Basen mit $\gamma_1 f_1, \cdots, \gamma_8 f_1$ bezeichnet werden, so sind $\gamma_1, \cdots, \gamma_8$ ganze rationalzahlige Funktionen, und man erhält

$$\begin{gathered}(\gamma_1 f_1)^2 + \cdots + (\gamma_8 f_1)^2 = ff_1^2 f_2, \\ \gamma_1^2 + \cdots + \gamma_8^2 \quad = ff_2, \\ f = \frac{\gamma_1^2 + \cdots + \gamma_8^2}{f_2},\end{gathered}$$

wo f_2 kleineren Grad hat als f_1.

Ist dieser Grad noch nicht 0, so wiederhole man dasselbe Verfahren. Man gelangt alsdann schließlich zu einer Gleichung

$$f = \frac{\delta_1^2 + \cdots + \delta_8^2}{c},$$

wo c eine positive Konstante ist, $\delta_1(x), \cdots, \delta_8(x)$ ganze rationale Funktionen vom Grade $\leqq \frac{n}{2}$.*) Daraus folgt, wenn die ganzen rationalzahligen Funktionen $\frac{\delta_1}{c}, \cdots, \frac{\delta_8}{c}$ mit $\varepsilon_1, \cdots, \varepsilon_8$ bezeichnet werden,

$$f = c\left(\left(\frac{\delta_1}{c}\right)^2 + \cdots + \left(\frac{\delta_8}{c}\right)^2\right) = c(\varepsilon_1^2 + \cdots + \varepsilon_8^2).$$

Da nun die Zahl c in vier Quadrate zerlegbar ist**), so ergibt sich

$$f(x) = (b_1^2 + b_2^2 + b_3^2 + b_4^2 + 0^2 + 0^2 + 0^2 + 0^2)\,(\varepsilon_1^2(x) + \cdots + \varepsilon_8^2(x)),$$

also durch Anwendung von (7)

$$f(x) = g_1^2(x) + g_2^2(x) + g_3^2(x) + g_4^2(x) + g_5^2(x) + g_6^2(x) + g_7^2(x) + g_8^2(x).$$

Nachdem nun die Zerlegbarkeit in acht Quadrate für irreduzibele definite Funktionen bewiesen ist, so folgt sie leicht für alle definiten Funktionen. Denn jede definite reduzible Funktion ist von der Gestalt

---

*) Mindestens eine derselben hat natürlich den Grad $\frac{n}{2}$.

**) Mit dem bekannten Beweise dieses Bachet-Lagrangeschen Satzes hat das ganze Verfahren große Ähnlichkeit.

$$f(x) = c(F(x))^2 F_1(x) F_2(x) \cdots F_\varrho(x),$$

wo $F_1(x), \cdots, F_\varrho(x)$ definit und irreduzibel sind; da nach dem Obigen die Faktoren $F_1(x), \cdots, F_\varrho(x)$ in je acht Quadrate zerlegbar sind, so ergibt die wiederholte Anwendung der Identität (7), daß $f(x)$ als Summe von acht Quadraten darstellbar ist.

Bekanntlich hat Herr Hurwitz*) bewiesen, daß für $\nu > 8$ (und $\nu = 3, 5, 6, 7$) das Produkt zweier Summen von ν Quadraten nicht als Summe von ν Quadraten darstellbar ist. Das Gelingen des Nachweises, daß $f(x)$ in eine feste (von n unabhängige) Anzahl von Quadraten ganzer rationalzahliger Funktionen zerlegbar ist, ist also nur dem glücklichen Umstand zu verdanken, daß durch den Hilbertschen Satz die Zerlegbarkeit von -1 in sieben Quadrate gewährleistet ist. Würde man z. B. nur beweisen können, daß -1 stets in acht Quadrate zerlegbar ist, so würde das oben eingeschlagene Verfahren überhaupt keine von n unabhängige obere Schranke für die Anzahl der zur Darstellung von $f(x)$ erforderlichen Quadrate ergeben**).

Es ist nun sehr zu wünschen, daß Herr Hilbert den Beweis seines Satzes veröffentlicht. Der Satz scheint mir dadurch noch an Bedeutung gewonnen zu haben, daß er die Erledigung unseres nur auf rationale Zahlen bezüglichen Problems nach sich zieht.

Die im Vorangehenden auseinandergesetzte Methode gestattet also, für jeden Körper, in welchem -1 als Summe von sieben Zahlenquadraten darstellbar ist, das zugehörige $f(x)$ als Summe von acht Quadraten ganzer rationalzahliger Funktionen darzustellen; nach Herrn Hilbert gilt dies also für jedes definite $f(x)$. Ich mache aber besonders darauf aufmerksam, daß die obige Reduktionsmethode in allen Fällen, für welche -1 in drei Quadrate zerlegt werden kann, sogar die Zerlegbarkeit von $f(x)$ in vier Quadrate ergibt; man braucht ja nur auf die Gleichung

$$f(x) = \frac{\alpha_1^2(x) + \alpha_2^2(x) + \alpha_3^2(x) + \alpha_4^2(x)}{f_1(x)}$$

dieselben Schlüsse anzuwenden wie oben auf (3) und dabei fortwährend von der Identität

---

*) „Über die Komposition der quadratischen Formen von beliebig vielen Variabeln", Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, 1898, S. 309—316.

**) Der Hilbertsche Satz, daß -1 stets in vier Quadrate zerlegbar ist, ergibt nicht mehr als die bloße Kenntnis, daß -1 in sieben Quadrate zerlegbar ist; das Produkt zweier Summen von je fünf Quadraten kann ja mit Sicherheit nur als Summe von acht Quadraten dargestellt werden.

$$(8)\quad (\alpha_1^2+\alpha_2^2+\alpha_3^2+\alpha_4^2)(\beta_1^2+\beta_2^2+\beta_3^2+\beta_4^2) = (\alpha_1\beta_1+\alpha_2\beta_2+\alpha_3\beta_3+\alpha_4\beta_4)^2 + (-\alpha_1\beta_2+\alpha_2\beta_1+\alpha_3\beta_4-\alpha_4\beta_3)^2 + (-\alpha_1\beta_3-\alpha_2\beta_4+\alpha_3\beta_1+\alpha_4\beta_2)^2 + (-\alpha_1\beta_4+\alpha_2\beta_3-\alpha_3\beta_2+\alpha_4\beta_1)^2$$

Gebrauch zu machen*).

Es ist vielleicht von Interesse, für eine spezielle Klasse von Körpern in elementarer Weise die Zerlegbarkeit von -1 in vier Quadrate darzutun. Ich verdanke meinem Freunde J. Schur einen solchen Nachweis für den Kreisteilungskörper der m^{ten} Einheitswurzeln ($m \geqq 3$) und teile hier diesen Beweis in unwesentlich modifizierter Form mit.

Ohne Beschränkung der Allgemeinheit kann m als ungerade Primzahl oder $=4$ angenommen werden. Denn wenn der Nachweis hierfür erledigt ist, so ergibt sich für jedes $m \geqq 3$, falls p einen ungeraden Primfaktor von m bezw. die Zahl 4 (für $m = 2^\varrho$) bezeichnet, im Körper der p^{ten} Einheitswurzeln eine Zerlegung

$$-1 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2,$$

und diese ist gleichzeitig eine Zerlegung im Körper der m^{ten} Einheitswurzeln.

Es handelt sich also, da offenbar im Körper der 4^{ten} Einheitswurzeln

$$-1 = i^2 = \alpha^2$$

ist, nur um den durch die Gleichung

$$f(x) = 1 + x + \cdots + x^{p-1} = 0$$

bestimmten Körper, wo p eine ungerade Primzahl ist.

1) p habe die Form $8\nu+3$ oder $8\nu+5$. Dann ist

$$2^{\frac{p-1}{2}} \equiv \left(\frac{2}{p}\right) = -1 \pmod p,$$

also

$$2^{\frac{p-1}{2}} + 1 = hp.$$

Nun ist

$$\begin{aligned}&(1+x+\cdots+x^{p-1})(1+x^p+x^{2p}+\cdots+x^{(h-1)p})\\ &= 1+x+\cdots+x^{p-1}+x^p+\cdots+x^{hp-1}\\ &= 1+x+\cdots+x^{2^{\frac{p-1}{2}}-1}+x^{hp-1}\\ &= (1+x)(1+x^2)(1+x^4)\cdots\left(1+x^{2^{\frac{p-1}{2}-1}}\right)+x^{hp-1},\end{aligned}$$

also, wenn für x eine Wurzel ϑ von $f(x)=0$ eingesetzt wird,

*) Daß analog für jeden Körper, in welchem -1 Quadratzahl ist, d. h. welcher die Zahl i enthält, $f(x)$ in zwei Quadrate zerlegt werden kann, ist trivial.

$$0 = (1+\vartheta)(1+\vartheta^2)(1+\vartheta^4)\cdots\left(1+\vartheta^{2^{\frac{p-1}{2}-1}}\right)+\vartheta^{-1},$$

also, da

$$\vartheta = \vartheta^{p+1} = \left(\vartheta^{\frac{p+1}{2}}\right)^2$$

ist und jeder der $\frac{p-1}{2}$ Faktoren des Produktes eine Summe von zwei Quadraten ist,

$$0 = \alpha^2+\beta^2+\vartheta^{-1},$$
$$-\vartheta^{-1} = \alpha^2+\beta^2,$$
$$-1 = (\alpha^2+\beta^2)\vartheta = (\alpha^2+\beta^2)\left(\vartheta^{\frac{p+1}{2}}\right)^2 = \alpha_1^2+\beta_1^2;$$

-1 ist also in zwei Quadrate zerlegbar und folglich nach meinen allgemeinen Bemerkungen auf S. 278—279 $f(x)$ in vier Quadrate*).

2) Die ungerade Primzahl p sef beliebig. Dann läßt sich jedenfalls eine Primzahl q so bestimmen, daß q die Form $8\nu+3$ hat und daß q Nichtrest modulo p ist. Dann ist

$$q^{\frac{p-1}{2}} \equiv \left(\frac{q}{p}\right) = -1 \pmod{p},$$

also

$$q^{\frac{p-1}{2}}+1 = hp,$$

$$(9)\quad (1+x+\cdots+x^{p-1})(1+x^p+\cdots+x^{(h-1)p}) = 1+x+\cdots+x^{hp-1}$$
$$= 1+x+\cdots+x^{q^{\frac{p-1}{2}}-1}+x^{hp-1}$$
$$= (1+x+x^2+\cdots+x^{q-1})(1+x^q+x^{2q}+\cdots+x^{(q-1)q})\cdot\cdot$$
$$\cdot\cdot\left(1+x^{q^{\frac{p-1}{2}-1}}+x^{2q^{\frac{p-1}{2}-1}}+\cdots+x^{(q-1)q^{\frac{p-1}{2}-1}}\right)+x^{hp-1}.$$

Nun ist nach 1)

$$1+x+x^2+\cdots+x^{q-1} = F_1^2(x)+\cdots+F_4^2(x),$$

---

*) Übrigens folgt für $p=8\nu+3$ die Zerlegbarkeit von $f(x)$ in vier Quadrate und ebenso für $p=8\nu+7$ die Zerlegbarkeit von $f(x)$ in fünf Quadrate direkt daraus, daß bekanntlich für $p\equiv 3 \pmod 4$

$$4f(x) = 4\frac{x^p-1}{x-1} = X_1^2+pX_2^2$$

ist, wo X_1 und X_2 ganze ganzzahlige Funktionen von x sind. Denn, da für $p=8\nu+3$ bezw. $p=8\nu+7$ eine Gleichung $p=a^2+b^2+c^2$ bezw. $p=a^2+b^2+c^2+d^2$ besteht, so ist $f(x)$ in vier bezw. fünf Quadrate zerlegbar. Nur der Fall $p=8\nu+1$ ist schwieriger zu behandeln, etwa wie oben unter 2) im Anschluß an die Mitteilungen von Herrn Schur, wobei der Dirichletsche Satz von der arithmetischen Progression angewendet wird.

also

$$1 + x^q + x^{2q} + \cdots + x^{(q-1)q} = F_1^2(x^q) + \cdots + F_4^2(x^q) = G_1^2(x) + \cdots + G_4^2(x),$$

. . . . . .

$$1 + x^{q\frac{p-1}{2}-1} + x^{2q\frac{p-1}{2}-1} + \cdots + x^{(q-1)q\frac{p-1}{2}-1} = F_1^2\left(x^{q\frac{p-1}{2}-1}\right) + \cdots + F_4^2\left(x^{q\frac{p-1}{2}-1}\right)$$
$$= K_1^2(x) + \cdots + K_4^2(x).$$

Daher hat nach (8) die rechte Seite von (9) die Gestalt einer Summe von vier Quadraten plus x^{hp-1}. Setzt man nun in (9) für x eine Wurzel ϑ der Gleichung

$$f(x) = 0$$

ein, so ergibt sich

$$0 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 + \vartheta^{-1},$$
$$-1 = (\alpha^2 + \beta^2 + \gamma^2 + \delta^2)\left(\vartheta^{\frac{p+1}{2}}\right)^2 = \alpha_1^2 + \beta_1^2 + \gamma_1^2 + \delta_1^2.$$

Zweiter Teil.

Die im ersten Teil angewendete Reduktionsmethode führt auch in der Theorie der definiten ganzen rationalen Funktionen zweier Variabeln

$$f(x, y) = a_0 + a_{10}x + a_{01}y + a_{20}x^2 + \cdots + a_{0n}y^n.$$

mit beliebigen reellen Zahlenkoeffizienten zu einem neuen Resultat. Bisher ist darüber folgendes bekannt.

Herr Hilbert*) hat zuerst den Nachweis der merkwürdigen Tatsache geführt, daß $f(x, y)$ im allgemeinen nicht als Summe von Quadraten ganzer rationaler Funktionen von x und y mit reellen Koeffizienten darstellbar ist. Später hat Herr Hilbert**) bewiesen, daß sich $f(x, y)$ stets als Quotient zweier solcher Quadratsummen darstellen läßt. Er beweist zu diesem Zwecke — unter Anwendung der Theorie der Abelschen Funktionen — zunächst den Satz***): „Jede beliebige ternäre definite Form†) F von der n^{ten} Ordnung ist in der Gestalt darstellbar

$$F = \frac{\Phi^2 + \Psi^2 + X^2}{H}, \tag{10}$$

wo Φ, Ψ, X Formen mit reellen Koeffizienten von der $n - 2^{\text{ten}}$ Ordnung sind und H die $n - 4^{\text{te}}$ Ordnung besitzt."

---

*) „Über die Darstellung definiter Formen als Summe von Formenquadraten", Mathematische Annalen, Bd. 32, 1888, S. 342—350.

**) „Über ternäre definite Formen", Acta mathematica, Bd. 17, 1893, S. 169—197.

***) l. c., S. 196.

†) D. h. jede definite ganze rationale homogene Funktion dreier Variabeln.

Aus (10) schließt nun Herr Hilbert weiter: H ist eine definite Form, läßt sich also in der Gestalt

$$\frac{\Phi_1^2+\Psi_1^2+X_1^2}{H_1}$$

darstellen, wo $n-8$ der Grad von H_1 ist. Die Wiederholung dieser Schlußweise führt schließlich zu einem Bruch, dessen Nenner eine positive Konstante oder eine quadratische definite Form ist. Da letztere eine Summe von Formenquadraten ist, so ergibt sich durch Ausführung der Multiplikationen für F eine Darstellung als Quotient von zwei Quadratsummen

$$f=\frac{\Phi_1^2+\cdots+\Phi_P^2}{\varphi_1^2+\cdots+\varphi_Q^2}. \tag{11}$$

Jede definite ternäre Form oder, was dasselbe besagt, jede definite Funktion zweier Variabeln ist also als Quotient von zwei Quadratsummen darstellbar.

Dies Hilbertsche Resultat läßt sich ohne Mühe dahin ergänzen, daß $f(x, y)$ als Quotient zweier Summen von je vier Quadraten darstellbar ist; denn es werden stets Summen von je drei, also von je vier Quadraten miteinander multipliziert, und es greift die Identität (8) Platz. Aus

$$f(x,y)=\frac{\Phi_1^2+\cdots+\Phi_4^2}{\varphi_1^2+\cdots+\varphi_4^2} \tag{12}$$

kann man noch schließen:

$$f(x,y)=\frac{(\Phi_1^2+\cdots+\Phi_4^2)(\varphi_1^2+\cdots+\varphi_4^2)}{(\varphi_1^2+\cdots+\varphi_4^2)^2}=\frac{\Psi_1^2+\cdots+\Psi_4^2}{\Psi^2}. \tag{13}$$

Aber damit ist nicht viel gewonnen. Sowohl in (11), als auch in (12) und (13) wachsen die Grade der auftretenden Funktionen in bezug auf jede Variable sehr an; sie haben die Größenordnung n^2.

Nun schließe ich aber, von (10) ausgehend, anders weiter und werde — unter der Voraussetzung, daß in der definiten Funktion n^{ten} Grades $f(x, y)$ das Glied mit y^n nicht den Koeffizienten Null hat*) — zu folgendem Satz gelangen:

„*$f(x, y)$ läßt sich als Summe von vier Quadraten*

$$f=\psi_1^2+\cdots+\psi_4^2$$

darstellen, wo $\psi_1, \cdots, \psi_4$ ganze rationale Funktionen von y mit rationalen reellen Funktionen von x als Koeffizienten sind.“

*) Durch eine ganze lineare Transformation der Variabeln läßt sich bekanntlich bei einer Funktion mehrerer Variabeln stets erreichen, daß der Grad in jeder Variabeln gleich dem Grade der Funktion ist.

Dieser Satz ist offenbar richtig für Funktionen 0^{ten} und 2^{ten} Grades. Denn für

$$f(x, y) = c$$

ist

$$f(x, y) = (\sqrt{c})^2;$$

für

$$f(x, y) = A_0(x)y^2 + A_1(x)y + A_2(x),$$

wo nach Voraussetzung A_0 eine positive Konstante und für alle reellen x

$$A_1^2(x) - 4A_0(x)A_2(x) \leqq 0$$

ist, ergibt sich

$$\begin{aligned} f(x, y) &= A_0(x)\left(y + \frac{A_1(x)}{2A_0(x)}\right)^2 + \frac{4A_0(x)A_2(x) - A_1^2(x)}{4A_0(x)} \\ &= (ay + \varphi_1(x))^2 + \varphi_2^2(x) + \varphi_3^2(x) \\ &= \psi_1^2(x, y) + \psi_2^2(x, y) + \psi_3^2(x, y). \end{aligned}$$

Der Satz möge für Funktionen $0^{\text{ten}}, 2^{\text{ten}}, \cdots, n-2^{\text{ten}}$ Grades als bewiesen angenommen werden.

Nach (10) ist die Funktion n^{ten} Grades $f(x, y)$*) in der Gestalt darstellbar

$$f = \frac{\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2}{f_1}, \tag{14}$$

wo $\alpha_1, \alpha_2, \alpha_3, \alpha_4, f_1$ ganze rationale Funktionen von x und y mit reellen Koeffizienten sind und f_1 in x und y, also auch in y allein, geringeren Grad hat als f**). (14) läßt sich auch so schreiben:

$$ff_1 = \alpha_1^2 + \cdots + \alpha_4^2. \tag{15}$$

Ich dividiere nun, wenn $f_1(x, y)$ nicht schon von y unabhängig ist, unter Bevorzugung der Variabeln y die vier Funktionen $\alpha_1, \cdots, \alpha_4$ durch f_1. Dies gibt

$$\begin{cases} \alpha_1 = q_1 f_1 + \beta_1, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \alpha_4 = q_4 f_1 + \beta_4, \end{cases} \tag{16}$$

wo $\beta_1, \cdots, \beta_4$ ganze rationale Funktionen von y sind, in denen die Koeffizienten der Potenzen von y (ganze oder gebrochene) rationale reelle Funktionen von x sind. Hierbei haben $\beta_1, \cdots, \beta_4$ in y kleineren Grad als f_1, also a fortiori als f.

1) Wenn die vier Reste $\beta_1, \beta_2, \beta_3, \beta_4$ identisch Null sind, ergibt sich aus (15)

$$f = f_1(q_1^2 + q_2^2 + q_3^2 + q_4^2) = f_1 f_2,$$

*) $f(x, y)$ kann ja als definite ternäre Form $F(x_1, x_2, x_3)$ geschrieben werden.

**) Für die spätere Wiederholung des folgenden Schlußverfahrens ist es wichtig zu bemerken, daß nur von der Tatsache Gebrauch gemacht wird, daß f_1 in y geringeren Grad hat als f, nicht davon, daß dies auch für den Grad in x und y gilt.

wo f_1 und f_2 ganze Funktionen von y sind, mit rationalen Funktionen von x als Koeffizienten. Der Grad k von f_1 in y liegt zwischen 0 (exkl.) und n (exkl.), der Grad $n-k$ von f_2 also gleichfalls. Nach der bekannten Verallgemeinerung eines Gaußschen Satzes ist also die ganze rationale Funktion $f(x, y)$ in zwei Faktoren $F_1(x, y)$ und $F_2(x, y)$ zerlegbar, die in x und y ganz sind und in y die Grade k und $n-k$ haben. In

$$f(x, y) = F_1(x, y)\, F_2(x, y)$$

sind auch die Grade von $F_1(x, y)$ und $F_2(x, y)$ in beiden Variabeln beziehlich k und $n-k$, da ja $f(x, y)$ in beiden Variabeln auch den Grad n hat. Die Koeffizienten von y^k bezw. y^{n-k} in $F_1(x, y)$ bezw. $F_2(x, y)$ sind Konstanten, von denen erstere ohne Beschränkung der Allgemeinheit $= 1$ angenommen werden kann. $F_1(x, y)$ und $F_2(x, y)$ sind ferner definit, da $F_1(x, y)$ aus einer definiten Funktion $f_1(x, y)$ durch Division mit einer definiten rationalen Funktion von x (nämlich dem Koeffizienten von y^k in $f_1(x, y)$) entsteht. $F_1(x, y)$ und $F_2(x, y)$ sind also Funktionen von x und y, welche genau dieselben Voraussetzungen erfüllen wie $f(x, y)$, aber geringeren Grad haben als $f(x, y)$; für solche Funktionen war der Satz als bewiesen angenommen worden, so daß sich für $f(x, y)$ eine Zerlegung

$$f(x, y) = (\varphi_1^2 + \cdots + \varphi_4^2)\,(\chi_1^2 + \cdots + \chi_4^2) = \psi_1^2 + \cdots + \psi_4^2$$

ohne Nenner in y ergibt.

2) Wenn β_1, β_2, β_3, β_4 nicht sämtlich identisch 0 sind, so ist $\beta_1^2 + \beta_2^2 + \beta_3^2 + \beta_4^2$ nicht identisch 0, und es folgt aus (15) und (16) modulo f_1

$$\beta_1^2 + \cdots + \beta_4^2 \equiv \alpha_1^2 + \cdots + \alpha_4^2 = f f_1 \equiv 0,$$

d. h. der Quotient

$$\frac{\beta_1^2 + \cdots + \beta_4^2}{f_1} = f_2$$

ist in y ganz und nicht identisch 0, in x rational. f_2 hat in y kleineren Grad als f_1. Aus (15) und

$$f_1 f_2 = \beta_1^2 + \cdots + \beta_4^2$$

folgt durch Multiplikation unter Benutzung der Identität (8)

$$\begin{aligned} f f_1^2 f_2 = (\ & \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 + \alpha_4\beta_4)^2 + (-\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_3\beta_4 - \alpha_4\beta_3)^2 \\ & + (-\alpha_1\beta_3 - \alpha_2\beta_4 + \alpha_3\beta_1 + \alpha_4\beta_2)^2 + (-\alpha_1\beta_4 + \alpha_2\beta_3 - \alpha_3\beta_2 + \alpha_4\beta_1)^2. \end{aligned}$$

Jede der vier Basen rechts ist durch f_1 teilbar, und man erhält

$$f f_2 = \gamma_1^2 + \cdots + \gamma_4^2,$$

$$f = \frac{\gamma_1^2 + \cdots + \gamma_4^2}{f_2}, \tag{17}$$

wo $\gamma_1, \cdots, \gamma_4$, f_2 in x rational, in y ganz sind, und wo f_2 in y geringeren Grad hat als f_1.

Durch Erweiterung des Bruches (17) mit einer gewissen ganzen rationalen Funktion von x nimmt (17) die Form an

$$f = \frac{\delta_1^2 + \cdots + \delta_4^2}{F_2},$$

wo $\delta_1, \cdots, \delta_4, F_2$ ganze rationale reelle Funktionen von x und y sind, und wo F_2 in y kleineren Grad hat als f_1 und a fortiori als f.

Ist dieser Grad von F_2 in bezug auf y noch nicht 0, so läßt sich das Verfahren wiederholen, bis man schließlich zu einer Gleichung kommt

$$f(x, y) = \frac{\varphi_1^2(x, y) + \cdots + \varphi_4^2(x, y)}{\psi(x)}, \tag{18}$$

wo im Nenner eine (definite ganze reelle) Funktion von x allein steht.

$f(x, y)$ läßt sich also mit einer solchen definiten ganzen Funktion von x allein multiplizieren, daß das Produkt als Summe von vier Quadraten ganzer reeller Funktionen von x und y darstellbar ist.

Aus (18) folgt schließlich wegen

$$\psi(x) = \chi_1^2(x) + \chi_2^2(x)$$

$$f(x, y) = \frac{(\chi_1^2(x) + \chi_2^2(x))(\varphi_1^2(x, y) + \cdots + \varphi_4^2(x, y))}{\psi^2(x)}$$

$$= \psi_1^2(x, y) + \cdots + \psi_4^2(x, y),$$

wo $\psi_1(x, y), \cdots, \psi_4(x, y)$ ganze rationale Funktionen von y sind, deren Koeffizienten rationale reelle Funktionen von x sind.

Dies war die auf S. 282 ausgesprochene Behauptung.

Den 31. August 1905.