Randbemerkungen zu Hauptproblemen der Mathematik.

Von

H. Weyl in Zürich.

Neben solchen Arbeiten, die — in alle Richtungen sich zersplitternd und darum jeweils auch nur von wenigen mit lebhafterem Interesse verfolgt — in wissenschaftliches Neuland vorstoßen, haben wohl auch Betrachtungen wie die hier vorgelegten, in denen es sich weniger um Mehrung als um Klärung, um möglichst einfache und sachgemäße Fassung des schon Gewonnenen handelt, ihre Berechtigung, wenn sie sich auf Hauptprobleme richten, an denen alle Mathematiker, die überhaupt diesen Namen verdienen, ungefähr in gleicher Weise interessiert sind[1]).

---

[1]) Die Sammlung soll später fortgesetzt werden.

Daß ich ein Stück voranstelle, in welchem es um die symbolische Methode der Invariantentheorie geht, ist zum Teil durch Herrn Study veranlaßt, der mir — neben andern Ungenannten; ich allein bin durch ein Zitat aus meinen Schriften eindeutig gekennzeichnet — Schuld gibt, ein „reiches Kulturgebiet (nämlich die Invariantentheorie) der Verwahrlosung übergeben, ja dessen Dasein völlig ignoriert zu haben" (Einleitung in die Theorie der Invarianten linearer Transformationen, Braunschweig 1923; Einleitung). Herr Study hat Recht darin, daß ich, ein „*sonst* kenntnisreicher (?) Autor", in der Invariantentheorie geringe Literaturkenntnis und geringe eigene Erfahrung habe; aber auch dann, wenn ich hierin mit Herrn Study wetteifern könnte, würde ich in meinem Buche „Raum, Zeit, Materie" die symbolische Methode nicht angewendet und von den algebraischen Vollständigkeitssätzen der Invariantentheorie kein Sterbenswörtchen gesagt haben. Alles an seinem Platz! — Die hoch und tief gestellten Indizes sind in der Physik darum so praktisch, weil sie dieselbe Größe, z. B. die Energiedichte, mit demselben Buchstaben zu bezeichnen gestatten, ob sie nun durch ihre kovarianten oder kontravarianten Komponenten charakterisiert wird. — Was Graßmann als *Stufe* bezeichnet, trägt in der ganzen übrigen Mathematik den Namen *Dimension*; zudem stimmt ja die Bedeutung, welche ich diesem Wort in der Tensorrechnung beilegte, für den wichtigsten Spezialfall der schiefsymmetrischen („linearen") Tensoren mit dem Graßmannschen Gebrauch überein!

I. Zur Invariantentheorie.

1. Capellische Identität.

Den formalen Apparat der Invariantentheorie kann man aus der Identität von Capelli[2]) entwickeln. Für sie möchte ich hier zunächst einen durchsichtigen Beweis geben und dann einige Bemerkungen daran schließen über den auf sie sich stützenden Aufbau der Invariantentheorie für die projektive Gruppe und ihre wichtigsten Untergruppen. Neben der orthogonalen Gruppe, welche eine nicht-ausgeartete *symmetrische* Bilinearform invariant läßt, soll die „Komplexgruppe" behandelt werden, welche in dem gleichen Verhältnis zu einer *schiefsymmetrischen* Form steht.

Wir haben es zu tun mit mehreren Reihen $x, y \ldots$ („Vektoren") von je n Veränderlichen und einer ganzen rationalen Funktion f dieser Variablen. $D_{xx'}$ wie $\Delta_{xx'}$ bezeichnen den Polarenprozeß[3]):

$$f\Delta_{xx'} = f D_{xx'} = \sum_{i=1}^{n} x'_i \frac{\partial f}{\partial x_i}.$$

Dabei kann die Variablenreihe x' mit einer der Reihen $x, y, \ldots$ zusammenfallen oder auch eine weitere unabhängige Variablenreihe vorstellen. Es sollen aber Δ und D bei Zusammensetzungen sich verschieden verhalten: beispielsweise entstehe $f D_{xx'} D_{yy'} D_{zz'}$ aus f, indem man hintereinander die Operationen $D_{xx'}, D_{yy'}, D_{zz'}$ ausführt; hingegen sei

$$f\Delta_{xx'}\Delta_{yy'}\Delta_{zz'} = \sum_{ikl} x'_i y'_k z'_l \frac{\partial^3 f}{\partial x_i \partial y_k \partial z_l}.$$

Die Resultate stimmen überein, außer wenn die Variablenreihe y oder z mit x' zusammenfällt oder $z = y'$ ist. Wir benutzen das Zeichen $\delta_{xx'} = 0$ oder 1, je nachdem x, x' zwei unabhängige oder zwei identische Variablenreihen sind. Es ist zu zeigen, daß die folgenden Gleichungen bestehen:

$$(1)\quad \left\{\begin{aligned}
&\begin{vmatrix} \Delta_{xx} \\ \Delta_{xy} \\ \Delta_{xz} \\ \vdots \end{vmatrix} = \begin{vmatrix} D_{xx} \\ D_{xy} \\ D_{xz} \\ \vdots \end{vmatrix}, \qquad
\begin{vmatrix} \Delta_{xx} & \Delta_{yx} \\ \Delta_{xy} & \Delta_{yy} \\ \Delta_{xz} & \Delta_{yz} \\ \vdots & \vdots \end{vmatrix} = \begin{vmatrix} \Delta_{xx} & D_{yx} \\ \Delta_{xy} & D_{yy}+1 \\ \Delta_{xz} & D_{yz} \\ \vdots & \vdots \end{vmatrix}; \\
&\begin{vmatrix} \Delta_{xx} & \Delta_{yx} & \Delta_{zx} \\ \Delta_{xy} & \Delta_{yy} & \Delta_{zy} \\ \Delta_{xz} & \Delta_{yz} & \Delta_{zz} \\ \vdots & \vdots & \vdots \end{vmatrix} = \begin{vmatrix} \Delta_{xx} & \Delta_{yx} & D_{zx} \\ \Delta_{xy} & \Delta_{yy} & D_{zy} \\ \Delta_{xz} & \Delta_{yz} & D_{zz}+2 \\ \vdots & \vdots & \vdots \end{vmatrix}, \ldots
\end{aligned}\right.$$

[2]) Math. Annalen 29 (1887), S. 331.

[3]) Ich stelle die Operationssymbole hinter das Funktionszeichen, damit sie bei Zusammensetzung in derjenigen Reihenfolge hintereinander erscheinen, in der sie auszuführen sind.

Das ist so zu verstehen, daß in der i-ten Formel die entsprechenden i-reihigen Unterdeterminanten rechts und links übereinstimmen. Die i Faktoren jedes Determinantengliedes sind in der Reihenfolge der Spalten hinzuschreiben, denen sie entstammen. Indem man in jeder Formel das Resultat der vorhergehenden verwendet, gelangt man zu der Identität

$$\begin{vmatrix} \Delta_{xx} & \Delta_{yx} & \Delta_{zx} & \dots \\ \Delta_{xy} & \Delta_{yy} & \Delta_{zy} & \dots \\ \Delta_{xz} & \Delta_{yz} & \Delta_{zz} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} = \begin{vmatrix} D_{xx} & D_{yx} & D_{zx} & \dots \\ D_{xy} & D_{yy}+1 & D_{zy} & \dots \\ D_{xz} & D_{yz} & D_{zz}+2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}.$$

Die m-reihige Operatorendeterminante links liefert, auf f angewendet, 0, wenn die Anzahl m der zur Verwendung kommenden Variablenreihen $x, y, z, \dots$ größer als n ist. Im Falle $m = n$ aber kommt, nach dem Multiplikationssatz der Determinanten, das Produkt aus der Determinante $(x y z \dots)$ und Ωf (Ω bezeichnet den Cayleyschen Ω-Prozeß).

Um z. B. die dritte der Gleichungen (1) zu gewinnen, haben wir

$$\begin{vmatrix} \Delta_{xx'} & \Delta_{yx'} & D_{zx'} \\ \Delta_{xy'} & \Delta_{yy'} & D_{zy'} \\ \Delta_{xz'} & \Delta_{yz'} & D_{zz'} \end{vmatrix} = \Sigma \pm \Delta_{xx'} \Delta_{yy'} D_{zz'}$$

zu berechnen. x, y, z und ebenso x', y', z' sind voneinander unabhängige Variablenreihen, während die gestrichenen von den ungestrichenen nicht verschieden zu sein brauchen. Die Summe rechts bezieht sich auf die $3!$ Permutationen von $x' y' z'$ mit alternierenden Vorzeichen. Es ist

$$\begin{aligned} f \Delta_{xx'} \Delta_{yy'} D_{zz'} &= \sum_l z'_l \frac{\partial}{\partial z_l} \Big(\sum_{ik} x'_i y'_k \frac{\partial^2 f}{\partial x_i \partial y_k} \Big) \\ &= f \Delta_{xx'} \Delta_{yy'} \Delta_{zz'} + \delta_{zx'} \cdot f \Delta_{xz'} \Delta_{yy'} + \delta_{zy'} \cdot f \Delta_{xx'} \Delta_{yz'}, \end{aligned}$$

also

$$\Delta_{xx'} \Delta_{yy'} D_{zz'} = \Delta_{xx'} \Delta_{yy'} \Delta_{zz'} + \Delta_{xz'} \Delta_{yy'} \delta_{zx'} + \Delta_{xx'} \Delta_{yz'} \delta_{zy'}.$$

Bei der Summation über die $3!$ Permutationen von $x' y' z'$ kann man im 2. Glied (unter Änderung des Vorzeichens) z' mit x', im 3. Glied z' mit y' vertauschen und erhält

$$\Sigma \pm \Delta_{xx'} \Delta_{yy'} D_{zz'} = \Sigma \pm \Delta_{xx'} \Delta_{yy'} \Delta_{zz'} - 2 \Sigma \pm \Delta_{xx'} \Delta_{yy'} \delta_{zz'}$$

oder, wenn das subtraktive Glied auf die andere Seite geschafft wird,

$$\begin{vmatrix} \Delta_{xx'} & \Delta_{yx'} & \Delta_{zx'} \\ \Delta_{xy'} & \Delta_{yy'} & \Delta_{zy'} \\ \Delta_{xz'} & \Delta_{yz'} & \Delta_{zz'} \end{vmatrix} = \begin{vmatrix} \Delta_{xx'} & \Delta_{yx'} & D_{zx'} + 2\delta_{zx'} \\ \Delta_{xy'} & \Delta_{yy'} & D_{zy'} + 2\delta_{zy'} \\ \Delta_{xz'} & \Delta_{yz'} & D_{zz'} + 2\delta_{zz'} \end{vmatrix}, \text{ q. e. d.}$$

2. Reduktion auf $n-1$ Vektoren.

Es sei $\mathfrak{H}$ eine Gruppe linearer homogener Transformationen in n Variablen. Unter (Vektor-) *Invariante* zu $\mathfrak{H}$ versteht man bekanntlich eine ganze rationale Funktion f mehrerer Vektoren $x, y, \ldots$, homogen in den Komponenten jedes einzelnen Vektors, die sich höchstens mit einem Faktor μ multipliziert, wenn man auf die Komponenten jedes Vektors dieselbe Transformation S von $\mathfrak{H}$ ausübt; μ soll nicht von den Vektoren, sondern nur von S abhängen. Ist μ insbesondere für alle Transformationen S von $\mathfrak{H}$ gleich 1, so spricht man von einer absoluten Invariante. Die Aufgabe ist, bestimmte Typen von Grundinvarianten ausfindig zu machen, durch welche sich alle Invarianten der Gruppe $\mathfrak{H}$ ganz rational ausdrücken. Einen wesentlichen Schritt zur Lösung dieses Problems ermöglicht die eben von neuem bewiesene Capellische Formel

$$(2) \qquad f \begin{vmatrix} D_{xx} & D_{yx} & \ldots \\ D_{xy} & D_{yy}+1 & \ldots \\ \vdots & \vdots & \vdots \end{vmatrix} = \begin{cases} 0 & (m>n) \qquad (\mathrm{C}') \\ (xy\ldots)\,\Omega f & (m=n); \qquad (\mathrm{C}) \end{cases}$$

sie reduziert es nämlich auf den Fall, daß f nur von n, bzw. $n-1$ Vektoren abhängt. Bei der ersten Reduktion, die sich auf die Formel (C′) stützt, muß nur angenommen werden, daß

(I) *durch Anwendung des Polarenprozesses aus einer Grundinvariante wiederum ein Aggregat von Grundinvarianten entsteht*;

bei der zweiten Reduktion, die sich auf (C) stützt, ist weiter zu fordern, daß

(II) *die Determinante* $(xy\ldots)$ *unter den Grundinvarianten vorkommt* (oder sich doch ganz rational durch sie darstellen läßt).

Man denke sich die in f wirklich vorkommenden Vektoren in bestimmter Reihenfolge $x, y, z, \ldots$ angeordnet, und nenne von zwei Funktionen f dieser Vektoren, wenn sie gleiche Gesamtordnung haben, diejenige die niedrigere, welche in x von geringerer Ordnung ist; wenn aber beide dieselbe Ordnung in x besitzen, soll die Ordnung in y über den Rang entscheiden; wenn die Ordnung in x und y übereinstimmt, geht die Entscheidung auf z über; usf. Das Hauptglied in der Capellischen Determinante (2) ist $= p(q+1)(r+2)\ldots f$, wenn $p, q, r, \ldots$ die Ordnungen von f in x, bzw. $y, z, \ldots$ bezeichnen. Alle andern Glieder entstehen durch Anwendung von Polarprozessen aus solchen Invarianten, die niedriger als f stehen. Wissen wir von solchen Invarianten schon, daß sie sich durch die Grundinvarianten ausdrücken, so folgt demnach aus (C′) das gleiche für f. Bei der Anwendung von (C) hat man dem Induktionsschluß noch die Annahme zugrunde zu legen, daß die Behauptung schon feststehe für Invarianten von geringerer Gesamtordnung als f, wie Ωf

eine ist. — So, wie hier geschehen, trennt man meiner Meinung nach am besten den formal-algebraischen und den mit gedanklichen Induktionsschlüssen operierenden Teil. Denn die Capellische Identität ist im Gegensatz zu den sonst daraus hergeleiteten Reihenentwicklungen eine übersichtliche, explizite anschreibbare und geschlossene Formel. Da die Weiterführung des Problems bisher nur für einzelne Gruppen $\mathfrak{H}$ gelungen ist, wird man diesem auf den beiden Formeln (2) beruhenden Konstruktionsstück, das für alle brauchbar ist, einen besonders hohen Wert zuschreiben.

3. Der Transformationsfaktor.

Will man weiterkommen, so muß man sich zunächst über den Transformationsfaktor $\mu = \mu(S)$ Klarheit verschaffen. Sind S, T irgend zwei Transformationen aus $\mathfrak{H}$, so gilt

$$(3) \qquad \mu(ST) = \mu(S)\cdot\mu(T), \quad \text{insbesondere} \quad \mu(S)\cdot\mu(S^{-1}) = 1.$$

Es ist bekannt, wie man aus der letzten Gleichung die Tatsache herleitet, daß für die volle projektive Gruppe $\mathfrak{G}$ der Faktor μ eine Potenz der Transformationsdeterminante ist. Allgemeiner: Ist $\mathfrak{H}$ dadurch definiert, daß die Transformationskoeffizienten gleich ganzen rationalen Funktionen unabhängiger Parameter gesetzt werden, und bleibt dabei die Determinante eine irreduzible Funktion dieser Parameter, so ist μ notwendig eine Potenz der Determinante. Ist insbesondere $\mathfrak{H} = \mathfrak{T}$ die Translationsgruppe im n-dimensionalen Raum, bestehend aus den Translationen

$$(4) \qquad S\colon\ \bar{x}_0 = x_0 + (c_1 x_1 + \dots + c_n x_n), \qquad \bar{x}_i = x_i \qquad (i = 1, 2, \dots, n)$$

— die x_i bedeuten hier Ebenenkoordinaten, x ist ein „kovarianter Vektor“; die c sind beliebige Konstanten —, so erhält man aus jener Gleichung

$$\mu(c_1, \dots, c_n)\cdot\mu(-c_1, \dots, -c_n) = 1.$$

Als Polynom in den Variablen $c_1, \dots, c_n$ ist darum μ eine Konstante und zwar $= 1$, weil für die Identität $\mu = 1$ ist.

In dem komplizierteren Fall, wo die Gruppe durch Relationen zwischen den Transformationskoeffizienten gegeben ist — das ist vor allem so für die orthogonale Gruppe $\mathfrak{D}$ —, wird man am besten zur Methode der infinitesimalen Operationen greifen. Ist $\Gamma = (\gamma_{ik})$ die Abweichung der allgemeinen infinitesimalen Operation der Gruppe $\mathfrak{H}$ von der Identität und $\mu = 1 + \varphi(\Gamma)$ der zugehörige Faktor, so bestehen zwischen den γ lediglich *lineare* homogene Relationen und φ ist eine *lineare* homogene Funktion von Γ. An Stelle der Bedingung (3) tritt die, daß φ den Wert 0 hat für jede Operation der „abgeleiteten Gruppe“, d. h. für die Matrizen Γ', die aus zwei Matrizen der infinitesimalen Gruppe, etwa Γ_1, Γ_2, nach der Formel entstehen

$$\Gamma' = [\Gamma_1 \Gamma_2] = \Gamma_1 \Gamma_2 - \Gamma_2 \Gamma_1,$$

und für alle diejenigen, die sich aus solchen Γ' linear zusammensetzen. Insbesondere geht daraus hervor: *Ist die zu $\mathfrak{H}$ gehörige infinitesimale Gruppe $\mathfrak{h}$ mit ihrer abgeleiteten identisch, so ist $\mu = 1$* — zum mindesten für alle der „Hauptschicht" von $\mathfrak{H}$ angehörigen Transformationen (das sind diejenigen, welche innerhalb $\mathfrak{H}$ mit der Identität in einem gemeinsamen zusammenhängenden Kontinuum liegen). Der wichtigste Fall dieses allgemeinen Satzes ist der auf $\mathfrak{D}$ bezügliche: *Für die orthogonale Gruppe gibt es nur absolute Invarianten; ausgenommen $n = 2$.* (Die Ausnahme fällt fort, wenn der zugrunde liegende Zahlbereich $\sqrt{-1}$ nicht enthält.) Dies gilt wörtlich, wenn wir zu $\mathfrak{D}$ nur die eigentlichen orthogonalen Transformationen, die Drehungen mit der Determinante $+1$ rechnen. Wenn auch die Spiegelungen mit der Determinante -1 dazu gehören, kann für die Spiegelungen $\mu = +1$ oder $= -1$ sein; je nachdem das eine oder das andere der Fall ist, sprechen wir von einer geraden oder einer ungeraden Invariante. Ist von einer Funktion f nur bekannt, daß sie sich den Drehungen gegenüber invariant verhält, so ist sie die Summe einer geraden und einer ungeraden Invariante.

4. Aufstellung der Grundinvarianten für die wichtigsten linearen Gruppen.

Drehungsgruppe. Für die orthogonale Gruppe $\mathfrak{D}$ kommt man aus mit den Grundinvarianten

$$(\mathrm{D}_1) \quad (x\,y\,z \ldots) \quad \text{und} \quad (\mathrm{D}_2) \quad (x \mid y) = x_1 y_1 + \ldots + x_n y_n .$$

Das wurde zuerst durch eine sehr schöne, auf die Gruppe $\mathfrak{D}$ zugeschnittene Methode durch E. Study gezeigt [4]). Da das Produkt zweier Determinanten (D_1) sich durch die skalaren Produkte vom Typus (D_2) ausdrücken läßt, kann man schärfer sagen: Eine gerade Invariante kann ganz rational dargestellt werden allein durch (D_2); eine ungerade Invariante ist die Summe mehrerer Terme, deren jeder das Produkt aus einem „Klammerfaktor" (D_1) und einer geraden Invariante ist. Für dieses Theorem hat auf dem hier begonnenen Wege, nachdem Burkhardt mit $n = 3$ vorangegangen war [5]), Herr Weitzenböck in seinem Buche „Invariantentheorie" (Groningen 1923) einen Beweis versucht, der aber wegen eines Fehlschlusses — Übergang zu (17) und (18) auf S. 244 — nicht stichhaltig ist. In den Anwendungen interessieren fast ausschließlich die absoluten Invarianten. Faßt man den

---

4) Leipziger Berichte 1897, S. 443—461.

5) Math. Annalen **43** (1893), S. 197—215.

Invariantenbegriff allgemeiner, so muß man zunächst wie im vorigen Absatz zeigen, daß zur Drehungsgruppe keine andern als absolute Invarianten gehören. Dann kommt man aber durch einen Schluß von $n-1$ auf n folgendermaßen sofort zum Ziel. Die absolute Invariante f hänge von höchstens $n-1$ Vektoren $x, y, \ldots$ ab. Bei gegebenen Vektoren $x, y, \ldots$ kann man dann ein dem ursprünglichen gleichsinniges Cartesisches Koordinatensystem $e^1, e^2, \ldots, e^n$ einführen, in welchem e^n zu jenen Vektoren senkrecht ist. Wird

$$x = x_1' e^1 + x_2' e^2 + \ldots + x_{n-1}' e^{n-1}, \ldots \tag{5}$$

gesetzt, so kommt

$$f(x, y, \ldots) = f_0(x', y', \ldots), \tag{6}$$

wo f_0 diejenige Funktion bedeutet, die aus f durch die Substitution $x_n = y_n = \ldots = 0$ entsteht. War f eine ungerade Invariante, so erhält man durch Transformation auf das Koordinatensystem $e^1, \ldots, e^{n-1}, -e^n$ neben (6) die Gleichung

$$f(x, y, \ldots) = -f_0(x', y', \ldots),$$

und darum ist $f = 0$. War aber f eine gerade Invariante, so ist $f_0(x', y', \ldots)$ bei frei veränderlichen Argumenten

$$x' = (x_1', x_2', \ldots, x_{n-1}'), \ldots$$

eine gerade Orthogonalinvariante in $n-1$ Dimensionen. Setzen wir unsern Satz für $n-1$ Dimensionen schon als gültig voraus, so können wir mithin f_0 ganz rational darstellen durch die skalaren Produkte vom Typus $(x' | y') = x_1' y_1' + \ldots + x_{n-1}' y_{n-1}'$:

$$f_0(x', y', \ldots) = \Phi(\ldots, (x' | y'), \ldots). \tag{7}$$

Wenden wir das an auf die durch (5) eingeführten Größen $x', \ldots$ und beachten, daß für sie $(x' | y') = (x | y)$ ist, so erhält man aus (7) die gewünschte Darstellung

$$f(x, y, \ldots) = \Phi(\ldots, (x | y), \ldots).$$

Am besten ordnet man danach den Beweis so an. f_n bedeute eine gerade oder ungerade Orthogonalinvariante in n Dimensionen, die von höchstens n Vektoren $x, y, \ldots$ abhängt; Φ bezeichne ein Aggregat der skalaren Produkte dieser Vektoren. Durch einen auf die Dimensionszahl n bezogenen Induktionsschluß beweist man T_n: „Jedes gerade f_n ist $= \Phi$, jedes ungerade $f_n = (x y \ldots)\, \Phi$". T_0 ist trivial; aus T_{n-1} folgt T_n. T_n folgt nämlich zunächst, wie oben gezeigt, für solche f_n, die höchstens $n-1$ Argumentvektoren enthalten. Die Übertragung auf n Argumente geschieht mit Hilfe der Capellischen Formel (C). Der Polarenprozeß verwandelt gerade in gerade, ungerade in ungerade Invarianten, während der Ω-Prozeß gerade und ungerade vertauscht. Bei Anwendung von (C) auf ein gerades f_n hat man noch, wenn auf der rechten Seite Ωf_n schon in der Form $(xy\ldots)\,\Phi$ dargestellt ist, den Ausdruck von $(xy\ldots)^2$ durch die skalaren Produkte zu benutzen. *Nachdem* T_n bewiesen ist, leitet man aus (C′), ohne den n dimensionalen Raum zu verlassen, den allgemeinen Satz her: Jede gerade Orthogonalinvariante ist $=\Phi$, jede ungerade gleich einer Summe von Termen des Typus $(xy\ldots)\,\Phi$. — Weder bei dem Studyschen noch bei diesem Verfahren braucht der von Burkhardt verwendete raffinierte Kunstgriff [6]) herangezogen zu werden.

Erweiterungssatz. Die *Bewegungsgruppe* $\mathfrak{B}$ ist für Ebenenkoordinaten $x_0 \,|\, x_1 x_2 \ldots x_n$ dadurch gegeben, daß der vektorielle Bestandteil $x_1 x_2 \ldots x_n$ einer willkürlichen orthogonalen Transformation unterliegt, der skalare x_0 ersetzt wird durch

$$(8) \qquad \bar{x}_0 = x_0 + (c_1 x_1 + \ldots + c_n x_n) \qquad \text{(die } c \text{ beliebige Konstante).}$$

Nach dem, was über die orthogonale Gruppe und die Translationen (4) bemerkt wurde, existieren auch zu $\mathfrak{B}$ nur absolute Invarianten. Bedeutet l das Wertsystem $1\,|\,0\,0\ldots 0$, so erscheinen als Grundinvarianten die drei Typen:

$$\text{(B)} \qquad (x\,|\,y) = x_1 y_1 + \ldots + x_n y_n, \qquad (x\,y\,z\ldots), \qquad (l\,x\,y\ldots)$$

(die Klammerfaktoren enthalten $n+1$ Glieder). Hier führt der gleiche Induktionsschluß — als Ersatz für das von Weitzenböck eingeschlagene Verfahren [7]) — zum Ziel. Hängt die Invariante f von höchstens n „Ebenen“ $x, y, \ldots$ ab, so kann man (zunächst unter der Einschränkung, daß ihre vektoriellen Bestandteile linear unabhängig sind) durch eine Translation (4), bei welcher f sich nicht ändert, erzwingen, daß ihre skalaren Bestandteile $x_0, y_0, \ldots$ alle verschwinden, — der Anfangspunkt des Koordinatensystems wird in den Schnitt der Ebenen verlegt. Nach dem Fundamentalsatz für $\mathfrak{D}$ drückt sich dann f ganz und rational aus durch $(l\,x\,y\ldots)$ und Grundinvarianten vom Typus $(x\,|\,y)$. Das gleiche Verfahren empfiehlt sich beim Übergang von der projektiven zur *affinen Gruppe* [8]), bei den *Semi- und Schiebungsinvarianten* [9]). Der allgemeine Satz, der hier sich ergibt (Erweiterungssatz), lautet: *Einer Gruppe* $\mathfrak{H}$ *von projektiven Trans-*

---

[6]) A. a. O., S. 201. Vgl. Weitzenböck, Invariantentheorie, S. 240.

[7]) Wiener Berichte, Abt. IIa, **122** (1913), S. 1255; Invariantentheorie, S. 277.

[8]) Weitzenböck, Jahresber. d. Deutsch. Math.-Vereinig. **22** (1913), S. 192—209; Invariantentheorie, S. 223—232.

[9]) J. Deruyts, Essai d'une théorie générale des formes algébriques. Lüttich 1890.

formationen in n Variablen entspricht die Gruppe $\mathfrak{H}_\nu$ *in* $\nu+n$ *Variablen* $x=(\xi_1,\ldots,\xi_\nu,\ x_1,\ldots,x_n)$, *welche definiert ist durch die Gleichungen*

$$\begin{aligned}\bar{\xi}_1 &= (\gamma_{11}\xi_1+\ldots+\gamma_{1\nu}\xi_\nu)+(c_{11}x_1+\ldots+c_{1n}x_n),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \bar{\xi}_\nu &= (\gamma_{\nu 1}\xi_1+\ldots+\gamma_{\nu\nu}\xi_\nu)+(c_{\nu 1}x_1+\ldots+c_{\nu n}x_n)\end{aligned}$$

zusammen mit der Festsetzung, daß $x_1,\ldots,x_n$ *unter sich eine beliebige Transformation aus* $\mathfrak{H}$ *erleiden können; die c und* γ *sind willkürlich mit der einen Einschränkung, daß die Determinante der* γ *gleich* 1 (*bzw.* $\neq 0$) *ist. Ein System von Grundinvarianten für* $\mathfrak{H}_\nu$ *wird alsdann gebildet aus den Grundinvarianten von* $\mathfrak{H}$ *zusammen mit der* $(\nu+n)$*-gliedrigen Determinante* $(x\,y\,z\ldots)$.

Beweis. Wieder können wir uns beschränken auf Invarianten f der Gruppe $\mathfrak{H}_\nu$, die von weniger als $\nu+n$ Vektoren $x, y, \ldots$ abhängen. Durch eine Translation

(9) $\bar{\xi}_1=\xi_1+\gamma_{12}\xi_2+\ldots+c_{1n}x_n$ (alle übrigen Variablen gehen in sich über)

können wir dann bei gegebenen $x, y, \ldots$ bewirken, daß die erste griechische Komponente aller dieser Vektoren verschwindet. Weil die Translationen (9) zu $\mathfrak{H}_\nu$ gehören und nach einer obigen Bemerkung der zugehörige Faktor $\mu=1$ ist, ändert sich f also nicht, wenn in ihm $\xi_1, \eta_1, \ldots$ durch 0 ersetzt werden. Das heißt: in f kommt die erste griechische Komponente von $x, y, \ldots$ gar nicht vor; ebensowenig die 2-te bis ν-te. f enthält demnach nur die lateinischen Variablen und ist als Invariante der Gruppe $\mathfrak{H}$ durch deren Grundinvarianten darstellbar.

Die *volle projektive Gruppe* $\mathfrak{G}$, ebenso die Gruppe aller homogenen linearen Transformationen *von der Determinante* 1 fällt unter diesen Beweis. Der Erweiterungssatz liefert ferner die Grundinvarianten für die *Gruppe der „gestuften" Transformationen*; man erhält sie, wenn man die Variablenreihe in eine Anzahl Abschnitte, sagen wir in drei Abschnitte x_1, x_2, x_3 teilt und nun ansetzt

$$\begin{aligned}\bar{x}_1 &= c_{11}x_1+c_{12}x_2+c_{13}x_3,\\ \bar{x}_2 &= \qquad\quad c_{22}x_2+c_{23}x_3,\\ \bar{x}_3 &= \qquad\qquad\qquad c_{33}x_3\end{aligned}$$

mit willkürlichen Matrizen c, die nur der einen Einschränkung genügen, daß die in der Hauptdiagonale stehenden die Determinante 1 (bzw. eine von 0 verschiedene Determinante) besitzen. Geometrisch handelt es sich hier um diejenigen projektiven Transformationen, welche ein System inzidenter linearer Elemente verschiedener Stufenzahl festlassen (Verallgemeinerung der Semiinvarianten; die Stufenzahlen brauchen nicht die lückenlose Reihe $0, 1, \ldots, n-1$ zu bilden).

Wie die Tabelle der Grundinvarianten zu ergänzen ist, wenn neben den kovarianten auch kontravariante Vektoren in Betracht gezogen werden, lehrt die Weitzenböcksche Methode der Komplexsymbole. Insbesondere hat Herr Weitzenböck das große Verdienst, in diesem allgemeineren Fall für die Gruppe $\mathfrak{B}$ der Elementargeometrie bei beliebiger Dimensionszahl jene Tabelle vollständig aufgestellt zu haben[10]).

Komplexgruppe. Von den wichtigsten linearen Gruppen hat bisher meines Wissens diejenige noch keine Bearbeitung gefunden, welche einen nicht-singulären linearen Geradenkomplex in sich überführt. Dennoch ist es mit den vorliegenden Mitteln leicht möglich, auch für sie die Grundinvarianten aufzustellen; in gewissem Sinne ist hier die Theorie noch einfacher als bei der orthogonalen Gruppe. Die Komponenten eines Vektors x im Raume von $n = 2h$ Dimensionen wollen wir mit $x_1, y_1, x_2, y_2, \ldots, x_h, y_h$ bezeichnen. Zugrunde liegt eine nicht-ausgeartete schiefsymmetrische Bilinearform in normaler Darstellung

$$(10) \qquad [x x'] = (x_1 y_1' - y_1 x_1') + \ldots + (x_h y_h' - y_h x_h')$$

(„schiefes Produkt" von x und x'). Die Gruppe $\mathfrak{C}$ besteht aus allen homogenen linearen Transformationen, welche, auf x und x' simultan angewendet, die Form $[x x']$ ungeändert lassen. *Als einzige Grundinvariante von $\mathfrak{C}$ tritt diese Form selber auf.* Der Beweis stützt sich auf die folgenden Tatsachen.

1. Die zugehörige Gruppe $\mathfrak{c}$ der infinitesimalen Operationen stimmt mit ihrer abgeleiteten $\mathfrak{c}'$ überein. Außerdem besteht $\mathfrak{C}$ aus einer einzigen zusammenhängenden Schicht. Infolgedessen können keine anderen als absolute Invarianten existieren.

2. Ist $b = b^1$ ein beliebiger von 0 verschiedener Vektor, so läßt sich dazu ein Koordinatensystem konstruieren: $a^1, b^1; a^2, b^2; \ldots; a^h, b^h$, in welchem das schiefe Produkt die Normalform (10) besitzt. — Hängt die Invariante f von weniger als n Vektoren $x, x', \ldots$ ab, so können wir bei gegebenen $x, x', \ldots$ hier b so wählen, daß $[b x] = [b x'] = \ldots = 0$ ist. Setzen wir dann

$$(11) \qquad x = (\xi_1 a^1 + \eta_1 b^1) + \ldots + (\xi_h a^h + \eta_h b^h),$$
entsprechend für $x', \ldots$,

so wird $\xi_1 = \xi_1' = \ldots = 0$; und es gilt

$$(12) \qquad f(x, x', \ldots) = f(\xi, \xi', \ldots).$$

[10]) Über Bewegungsinvarianten, 1. bis 15. Mitteilung, Wiener Berichte, Abt. IIa, 122 ff. (ab 1913); ders., Invariantentheorie, S. 281—301.

3. Zu $\mathfrak{C}$ gehören die unimodularen Transformationen des Variablenpaares x_1, y_1 (bei welchen die übrigen Variablen ungeändert bleiben):

$$\begin{aligned} \bar{x}_1 &= \alpha x_1 + \beta y_1 \\ \bar{y}_1 &= \gamma x_1 + \delta y_1 \end{aligned} \qquad (\alpha\delta - \beta\gamma = 1). \tag{13}$$

Nach dem Fundamentalsatz für diese zweidimensionale Gruppe ist f eine ganze rationale Funktion der Determinanten vom Typus $(xx')_1 = x_1 y_1' - y_1 x_1'$. Die Funktion f_0, welche aus f durch die Substitution $x_1 = x_1' = \ldots = 0$ entsteht, enthält also auch die Variablen $y_1, y_1', \ldots$ nicht mehr. Oder in mehr elementarer Schlußweise: Aus der Invarianz von f gegenüber (13) folgt

$$f(0, y_1;\ 0, y_1';\ \ldots) = f(\beta y_1, \delta y_1;\ \beta y_1', \delta y_1';\ \ldots)$$

(wobei nur das erste Variablenpaar in Evidenz gesetzt wurde). Diese Gleichung ist numerisch richtig für beliebige Zahlen $(\beta, \delta) \neq (0, 0)$. Infolgedessen liegt eine algebraische Identität in den Variablen β, δ vor, und ich darf in ihr auch $\beta = \delta = 0$ setzen. — Die numerische Gleichung (12) kann daraufhin, mit der abgeänderten Bezeichnung

$$\xi = (\xi_2 \eta_2, \ldots, \xi_h \eta_h), \qquad \xi' = (\xi_2' \eta_2', \ldots, \xi_h' \eta_h'), \qquad \ldots, \tag{14}$$

so geschrieben werden:

$$f(x, x', \ldots) = f_0(\xi, \xi', \ldots). \tag{15}$$

$f_0(\xi, \xi', \ldots)$ ist bei frei veränderlichen Argumenten (14) eine Komplexinvariante in $2(h-1)$ Dimensionen. Ist für diese Dimensionszahl unser Satz schon als richtig erkannt, so läßt sich f_0 also ganz rational ausdrücken durch die schiefen Produkte vom Typus

$$[\xi\xi'] = (\xi_2\eta_2' - \eta_2\xi_2') + \ldots + (\xi_h\eta_h' - \eta_h\xi_h'):$$
$$f_0(\xi, \xi', \ldots) = \Phi(\ldots, [\xi\xi'], \ldots).$$

Für die besonderen durch (11) eingeführten Wertsysteme folgt dann aus (15) und $[xx'] = [\xi\xi']$ die behauptete Darstellung von f.

4. Um nun aber den Induktionsschluß zu beenden und in n Dimensionen mit Hilfe der Capellischen Identität von weniger als n zu einer beliebigen Anzahl von Argumenten überzugehen, brauchen wir noch den Satz: Die Determinante $(x'x'' \ldots x^{(n)})$ von n Vektoren $x', x'', \ldots, x^{(n)}$ drückt sich ganz rational durch ihre schiefen Produkte zu je zweien aus (als ein sog. Pfaffsches Aggregat):

$$(x'x'' \ldots x^{(n)}) = \frac{1}{2 \cdot 4 \ldots 2h} \sum \pm [x'x''][x'''x''''] \ldots [x^{(n-1)}x^{(n)}].$$

In dem Produkt, das rechts unter dem Summenzeichen auftritt, sollen die n Vektoren sämtlichen Permutationen unterworfen werden; alle $n!$ Terme sind darauf mit alternierenden Vorzeichen zu addieren.

Zieht man neben kovarianten auch kontravariante Vektoren

$$u = (u^1, v^1, u^2, v^2, \ldots, u^h, v^h)$$

in Betracht, so erledigt sich die Aufstellung der Grundinvarianten durch die einfache Bemerkung, daß mit jedem solchen Vektor u ein kovarianter verbunden ist durch die Gleichungen

$$x_i = v^i, \quad y_i = -u^i \qquad (i = 1, 2, \ldots, h).$$

Die sich ergebenden Typen sind

$$[xx'], \quad (xu), \quad [uu'].$$

II. Fundamentalsatz der Algebra und Grundlagen der Mathematik.

Der Fundamentalsatz der Algebra ist oft genug behandelt worden; keiner der bekannten Beweise genügt aber den Anforderungen, welche die Brouwersche Analysis an einen strengen Existenzbeweis stellt[11]). Ich möchte zunächst an diesem Beispiel zeigen, wie jene von seiten der Logik geltend gemachten Ansprüche mit denen der Praxis genau zusammenfallen, und dann einige allgemeine Bemerkungen über die Grundlagen der Mathematik anschließen, welche durch die anders gerichteten Hilbertschen Untersuchungen auf diesem Gebiet hervorgerufen sind.

Die Aufgabe ist: wenn die Koeffizienten der Gleichung

$$(1) \qquad y = f(x) = x^n + a_1 x^{n-1} + \ldots + a_n = 0$$

approximativ gegeben sind, die Wurzeln angenähert zu bestimmen — in solcher Weise, daß die Genauigkeit der Wurzelbestimmung mit unbegrenzt wachsender Genauigkeit der Koeffizienten gleichfalls schließlich jeden Grad überschreitet. Bei solcher Fassung schließt die Existenz der Wurzeln offenbar ihre stetige Abhängigkeit von den Koeffizienten ein.

Wir bedecken die komplexe x-Ebene mit dem quadratischen Raster von der Maschenweite $\frac{1}{2^h}$; die Teilungslinien sind diejenigen, auf denen der Realteil oder der Imaginärteil von x ein ganzzahliges Vielfaches von $\frac{1}{2^h}$ ist. Die Quadrate des Rasters nennen wir die Dualquadrate h-ter Ordnung; ein „Dualstern“ h-ter Ordnung ist das Quadrat von der Seitenlänge $\frac{1}{2^{h-1}}$, das gebildet wird von den vier Dualquadraten h-ter Ordnung, die um einen Eckpunkt des Rasters herumliegen. Entsprechend dem Übergang zu immer verschärfter Genauigkeit durchlaufe h die Folge der Werte

[11]) Vgl. darüber Weyl, Über die neue Grundlagenkrise der Mathematik. Math. Zeitschr. 10 (1921), S. 39.

$0, 1, 2, \ldots$. Auf der h-ten Stufe seien die Koeffizienten mit solcher Genauigkeit bekannt, daß für jeden von ihnen mit Sicherheit ein Dualstern h-ter Ordnung in der komplexen Ebene angewiesen werden kann, in welchem er liegt. Man kann auch statt der einzelnen Koeffizienten gleich das ganze System $(a_1, a_2, \ldots, a_n)$ als Punkt in einem $2n$-dimensionalen Zahlenraum ins Auge fassen; dann ist in ihm ein Dualstern h-ter Ordnung $A = A_h$ bekannt, welchem der Koeffizientenpunkt sicher angehört. — Schon auf der 0-ten Stufe der Approximation können wir eine ganze Zahl m angeben, daß alle Wurzeln der Gleichung innerhalb des im Nullpunkt zentrierten Quadrats $\mathfrak{Q}$ von der Seitenlänge $2m$ liegen. Wir gehen nun zunächst genau nach dem Muster desjenigen Beweises vor, der mit Weierstraß in dem von Stufe zu Stufe feiner werdenden Quadratraster der x-Ebene ein Minimum des absoluten Betrages $|f(x)|$ abzufangen sucht.

Bei der Genauigkeit, mit welcher die Koeffizienten auf der h-ten Stufe bekannt sind, kann ich zu jedem Dualquadrat h-ter Ordnung $\mathfrak{q}_x$ der x-Ebene, welches zu $\mathfrak{Q}$ gehört, ein Quadrat $\mathfrak{q}_y$ in der y-Ebene angeben, in welchem die Werte y meines Polynoms gelegen sind, wenn x in $\mathfrak{q}_x$ variiert; $\mathfrak{q}_y$ soll ein Quadrat sein, das aus lauter Dualquadraten h-ter Ordnung zusammengesetzt ist. Enthält $\mathfrak{q}_y$ den Nullpunkt der y-Ebene (eines der vier um den Nullpunkt herumliegenden Dualquadrate h-ter Ordnung), so nennen wir $\mathfrak{q}_x$ ein *nullstellen-verdächtiges Quadrat.* Diese Prüfung kann für jedes der endlichvielen Quadrate $\mathfrak{q}_x$, aus denen $\mathfrak{Q}$ besteht, vorgenommen werden. Ein verdächtiges Quadrat $\mathfrak{q}_x$ wird man als approximative Bestimmung einer Wurzel betrachten wollen. Geht man nun aber zur nächsten Genauigkeitsstufe $(h+1)$ über, so kann es sich natürlich ereignen, daß sich von den vier Dualquadraten $(h+1)$-ter Ordnung, in welche $\mathfrak{q}_x$ zerfallen ist, keines mehr als nullstellen-verdächtig ergibt: wir werden aus $\mathfrak{q}_x$ vertrieben und müssen anderswo unser Heil versuchen. So können wir von Stufe zu Stufe in dem ganzen Quadrat $\mathfrak{Q}$ herumgetrieben werden, ohne irgendwo zur Ruhe zu kommen und uns einem festen Wurzelwert anzunähern. Dies ist die Hauptschwierigkeit, welche überwunden werden muß; bei der üblichen Gedankenführung wird hier der Sprung ins Jenseits durch den Existentialabsolutismus vollzogen.

Nötig ist offenbar ein Kriterium, das schon auf der h-ten Stufe darüber entscheidet, ob man bei fortgesetzter Teilung und beliebig fortschreitender genauerer Festlegung der Koeffizienten in $\mathfrak{q}_x$ immer wieder nullstellen-verdächtige Quadrate antreffen wird oder nicht. Es ist eine wunderbare Tatsache, daß ein solches Kriterium vorhanden ist — nicht freilich für das einzelne Quadrat $\mathfrak{q}_x$, sondern für die einzelnen zusammenhängenden, aber gegenseitig voneinander getrennten Gebiete $\mathfrak{g}$, zu denen sich die verdächtigen Quadrate zusammenfügen: die auf h-ter Stufe vorliegende Genauigkeit genügt, um zu berechnen, welchen Zuwachs

$$\frac{1}{2\pi}\cdot \arg y = \frac{1}{2\pi i}\int \frac{f'(x)\,dx}{f(x)} \tag{2}$$

erfährt, wenn x das äußere Umrißpolygon Π von $\mathfrak{g}$ durchläuft. Dieser Zuwachs ist eine ganze nicht-negative Zahl $n(\mathfrak{g})$, welche angibt, wie viele Nullstellen in $\mathfrak{g}$ liegen. Jede Seite σ von Π ist nämlich Seite eines nicht nullstellen-verdächtigen Dualquadrats h-ter Ordnung $\mathfrak{q}_x$, welchem in der y-Ebene ein den Nullpunkt nicht enthaltendes Quadrat $\mathfrak{q}_y$ entspricht; infolgedessen ist für die Werte von $\frac{1}{2\pi}\arg y$ auf σ ein Intervall bekannt, dem sie angehören und dessen Länge $< \frac{1}{2}$ ist, und das genügt, um jene ganze Zahl mit Sicherheit zu bestimmen. Indem man jede Seite σ durch jenes zugehörige Quadrat $\mathfrak{q}_y$ in der y-Ebene ersetzt, hat man nur darauf zu achten, wie oft die dem Umrißpolygon Π korrespondierende Kette von Quadraten $\mathfrak{q}_y$ den Nullpunkt umschließt. Die über die verschiedenen Gebiete $\mathfrak{g}$ erstreckte Summe $\Sigma n(\mathfrak{g})$ ist gleich dem Grad n des Polynoms. Beizubehalten sind nur diejenigen Gebiete $\mathfrak{g}$, für welche $n(\mathfrak{g}) \neq 0$ ist. Bei der Durchführung des Beweises werden wir noch jedes der Gebiete $\mathfrak{g}$ um der Einfachheit willen durch ein $\mathfrak{g}$ einschließendes Rechteck $\mathfrak{r}$ ersetzen.

Es ist weiter erforderlich, daß die Gebiete $\mathfrak{g}$ bei unbegrenzt fortgesetzter Teilung schließlich beliebig klein werden und sich auf einzelne Punkte zusammenziehen. Dies beruht offenbar darauf, daß die Dichtigkeit der Wurzelverteilung eine von vornherein angebbare Schranke nicht übersteigt, weil nur eine beschränkte Anzahl von Wurzeln, nämlich höchstens n, existieren. Der Satz muß folgendermaßen verschärft werden: Hat das Polynom (1) an $n+1$ Stellen $x = \alpha_0, \alpha_1, \ldots, \alpha_n$ Werte, welche absolut $\leqq \varepsilon$ sind, so können nicht alle Abstände $|\alpha_i - \alpha_k|$ $(i \neq k)$ dieser $n+1$ Werte voneinander $> \sqrt[n]{n\varepsilon}$ sein. An der Schranke $\sqrt[n]{n\varepsilon}$ ist natürlich allein das von Wichtigkeit, daß sie mit ε gegen 0 konvergiert. Der Beweis ergibt sich sogleich aus der L a g r a n g eschen Interpolationsformel: Hat $f(x)$ an der Stelle $x = \alpha_i$ den Wert β_i, so ist nach ihr der höchste Koeffizient von f:

$$1 = \frac{\beta_0}{(\alpha_0 - \alpha_1)\ldots(\alpha_0 - \alpha_n)} + \cdots.$$

Sind alle Differenzen $|\alpha_i - \alpha_k| \geqq \delta$ $(i \neq k)$, alle $|\beta_i|$ hingegen $\leqq \varepsilon$, so folgt daraus

$$\frac{n\varepsilon}{\delta^n} \geqq 1, \quad \text{d. i.} \quad \delta \leqq \sqrt[n]{n\varepsilon}.$$

Die Konstruktion verläuft demnach folgendermaßen: $\mathfrak{q} = \mathfrak{q}_x$ bedeute jetzt ausschließlich die nullstellen-verdächtigen Dualquadrate h-ter Ordnung: ihnen entsprechen in der y-Ebene die Quadrate $\mathfrak{q}_y$, welche den Nullpunkt enthalten. Die größte unter den Seitenlängen dieser endlichvielen $\mathfrak{q}_y$ — ein Dualbruch mit dem Nenner 2^h — sei ε_h, δ_h aber der kleinste Dualbruch mit dem Nenner 2^{h+1} und ungeradem Zähler, für welchen $\delta_h^n \geqq 2\, n\, \varepsilon_h$ ist. Es gilt $\lim\limits_{h=\infty} \delta_h = 0$. Wir beginnen mit einem Quadrat $\mathfrak{q}_x = \mathfrak{q}_1$ und vergrößern es unter Festhaltung seines Mittelpunktes zu einem Quadrat $\mathfrak{q}_1^*$ von der Seitenlänge $2\delta_h$. Gibt es noch nullstellen-verdächtige $\mathfrak{q}$ außerhalb dieses größeren $\mathfrak{q}_1^*$, so wählen wir eines von ihnen und vergrößern es in der gleichen Art zu $\mathfrak{q}_2^*$. Existieren weiter Quadrate $\mathfrak{q}$, welche weder zu $\mathfrak{q}_1^*$ noch zu $\mathfrak{q}_2^*$ gehören, so verfällt eines von ihnen abermals dem gleichen Vergrößerungsprozeß: $\mathfrak{q}_3^*$; und so fort. Nach dem Hilfssatz kann sich das nicht öfter als n-mal wiederholen: wir bekommen höchstens n Quadrate $\mathfrak{q}_1^*, \mathfrak{q}_2^*, \ldots$, welche alle $\mathfrak{q}$ enthalten. Wünscht man eine bestimmte Vorschrift für die Reihenfolge, in der die Quadrate vorgenommen werden, so setze man etwa fest, daß von je zwei Dualquadraten dasjenige den Vorrang hat, welches weiter links liegt; konkurrieren aber zwei solche Quadrate, welche demselben Vertikalstollen angehören, soll das tiefer gelegene den Vorrang haben. Die Quadrate $\mathfrak{q}_1^*, \mathfrak{q}_2^*, \ldots$ brauchen nicht völlig getrennt zu liegen. Wir konstruieren das kleinste (aus Dualquadraten h-ter Ordnung bestehende) Rechteck $\mathfrak{r}_1$, das 1. $\mathfrak{q}_1^*$ enthält und an welches 2. keines der Quadrate $\mathfrak{q}^*$ stößt, ohne ganz in ihm enthalten zu sein. So bekommen wir an Stelle der Quadrate $\mathfrak{q}^*$ höchstens n Rechtecke $\mathfrak{r}$, die alle nullstellen-verdächtigen $\mathfrak{q}$ enthalten und welche sich weder überdecken noch aneinander grenzen. Die Seitenlängen dieser Rechtecke sind höchstens gleich $2\, n\, \delta_h$. Für jedes $\mathfrak{r}$ bestimmen wir nach oben gegebener Anweisung die Zahl $n(\mathfrak{r})$ und scheiden diejenigen aus, für welche $n(\mathfrak{r}) = 0$ ist. Bleiben z. B. drei Rechtecke $\mathfrak{r}$ übrig: $\mathfrak{r}', \mathfrak{r}'', \mathfrak{r}'''$ mit den zugehörigen Anzahlen $n(\mathfrak{r}) = 3, 1, 2$ (die Gleichung sei also vom Grade $3 + 1 + 2 = 6$), so definieren wir in dem Wurzelraum $(\alpha_1, \alpha_2, \ldots, \alpha_6)$ ein 12-dimensionales Rechteck $\mathfrak{R}$ durch die Bedingungen: $\alpha_1, \alpha_2, \alpha_3$ in $\mathfrak{r}'$, α_4 in $\mathfrak{r}''$, α_5, α_6 in $\mathfrak{r}'''$. $\mathfrak{R}$ möge noch ersetzt werden durch das kleinste Quadrat A_h, das aus Dualquadraten h-ter Ordnung zusammengesetzt ist, denselben Mittelpunkt wie $\mathfrak{R}$ hat und $\mathfrak{R}$ ganz im Innern enthält. Beim Übergang zur nächsten Teilungsstufe brauchen wir von vornherein nur diejenigen Dualquadrate $(h+1)$-ter Ordnung in der x-Ebene zu berücksichtigen, welche durch Teilung aus nullstellen-verdächtigen Dualquadraten h-ter Ordnung hervorgehen. Es ergibt sich im Wurzelraum ein ganz im Innern von A_h enthaltenes $2n$-dimensionales Quadrat A_{h+1}. *Mit unbegrenzt wachsendem h konvergiert die Seitenlänge von A_h gegen 0.*

Die beschriebene Konstruktion ist eine solche, die mit einer durch freie Wahl werdenden Schachtelfolge von Dualsternen A_h im Koeffizientenraum ($h = 0, 1, 2, \ldots$) gesetzmäßig eine werdende Schachtelfolge von Quadraten A_h im Wurzelraum verknüpft. Sie benutzt lauter längst bekannte Elemente; diese sind aber auch alle für einen konstruktiv durchführbaren Existenzbeweis wirklich erforderlich. Der Satz, zu dem wir gelangen, besagt: *Jedes Polynom* (1) *läßt sich in Linearfaktoren zerlegen:*

$$f(x) = (x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n).$$

Dabei kann noch eine Zusatzbedingung wie

$$\alpha_1 \leqq \alpha_2 \leqq \ldots \leqq \alpha_n$$

gefordert werden, wo $\prec$ das Zeichen für die Rangordnung gemäß einer oben getroffenen Festsetzung ist. Unzulässig ist hingegen die Behauptung, $f(x)$ lasse sich stets als Potenzprodukt von lauter *verschiedenen* Linearfaktoren darstellen.

In seiner ersten Mitteilung zur „Neubegründung der Mathematik"[12]) hat sich Hilbert in heftiger Polemik gegen die von Brouwer und mir vertretene Auffassung gewendet. Mir scheint, selbst von seinem Standpunkt mit geringem Recht; denn soviel ich sehe, stimmen wir in dem entscheidendsten Punkte miteinander überein. Auch für Hilbert reicht die Kraft des inhaltlichen Denkens nicht weiter als für Brouwer; es ist für ihn ganz selbstverständlich, daß sie die „transfiniten" Schlußweisen der Mathematik nicht trägt[13]), daß es keine Rechtfertigung für alle die transfiniten Aussagen der Mathematik als *inhaltlicher Wahrheiten* gibt. Er wird nicht leugnen wollen, daß Brouwer hier im „Axiom des ausgeschlossenen Dritten" den wesentlichen Punkt getroffen hat; dieser neuen Einsicht zufolge kann Hilbert jetzt auch das eigentlich Fragwürdige und zu Begründende, dem inhaltlichen Denken schlechterdings Unzugängliche, was vor ihm als logisch selbstverständlich nicht in die Axiome aufgenommen wurde, sondern unerkannt „zwischen den Zeilen" sein Wesen trieb, der Formalisierung unterwerfen und zum Kernstück seiner Axiomatik machen. Ein weiterer großer Fortschritt von Hilbert über die Ansätze seines bekannten Heidelberger Vortrages[14]) hinaus liegt in der Anerkennung der Tatsache, daß ohne ein anschaulich-sachhaltiges Denken, das sich nicht auf Axiome gründen läßt und in welchem sich insbesondere ein finites Prinzip der vollständigen Induktion betätigt[15]), nicht auszukommen ist.

[12]) Abhandlungen aus dem Mathem. Seminar Hamburg 1 (1922), S. 157.

[13]) Vgl. hierzu Hilbert, Die logischen Grundlagen der Mathematik, Math. Ann. 88 (1922), S. 151; namentlich S. 155, 156.

[14]) Verhandlungen des 3. internationalen Mathematiker-Kongresses (1904), S. 174.

[15]) Dem wird kein Abbruch dadurch getan, daß das Prinzip auch als eine Formel im Axiomensystem erscheint. Der Syllogismus kommt bei Hilbert sogar in dreierlei Bedeutung vor: erstens als naiv gehandhabtes Prinzip des anschaulichen Denkens,

Mit Hilfe dieses anschaulichen Denkens ist nach Hilbert eine Begründung der elementaren Arithmetik in der Weise möglich, daß ihre Sätze nicht als „Formeln", sondern als inhaltliche Wahrheit verstanden werden[16]. Brouwer zeigt, vor allem vermöge seines Begriffs der werdenden Folge, daß das inhaltliche Denken wesentlich weiter reicht; es ist unter allen Umständen wichtig, zu sehen, *wie* weit man damit kommt. Vielleicht nimmt Hilbert in dieser Frage einen noch radikaleren Standpunkt ein als Brouwer; ich glaube aber, daß solche durchgeführten Beweise wie der eben angegebene für den funktionentheoretischen Fundamentalsatz der Algebra erkennen lassen, daß die Grenzen des finiten Denkens in der Brouwerschen Analysis — wenigstens in der von mir gegebenen Interpretation[17] — nicht überschritten werden. Auf dem Hintergrund der vorangehenden Epochen der modernen Grundlagenforschung — der ersten Epoche des naiven Existentialabsolutismus (Dedekind, Cantor), der zweiten, welche das „nicht-Prädikative" durch eine Typentheorie zu überwinden strebt (Russell) — heben sich Brouwer und Hilbert deutlich als zusammengehörig zu einer neuen dritten Epoche ab (der Übergang von der 1. zur 2. Epoche wird durch Frege, ein Übergang von der 2. zur 3. durch meine Schrift „Das Kontinuum" aus dem Jahre 1918 bezeichnet).

Die neue, Hilbert eigentümliche Wendung ist die, daß er an den Sätzen der Mathematik ihre inhaltliche Bedeutung fahren läßt und sie zu einem Formelspiel entleert. Gelingt ihm der Beweis der Widerspruchslosigkeit — und er hat den Weg, der zu diesem Ziele führen soll, bereits deutlich und überzeugend abgesteckt[18] —, so wird ein Resultat gewonnen sein, dem auch ich große Bedeutung zuschreibe, die Vollendung und Krönung von Hilberts axiomatischem Lebenswerk. Die Hilbertsche Mathematik ist aber, wenn man seine Erklärungen ernst nimmt, in der Tat nur ein *Formelspiel*. Die Analogie mit dem Schachspiel sei in einer kleinen Tabelle durchgeführt.

	Schachspiel	Hilbertsche Mathematik
1.	Stein.	Zeichen.
2.	Stellung der Steine auf dem Brett.	Formel.
3.	Ausgangsstellung.	System der Axiome.
4.	Zugregeln.	Regeln, nach denen aus Formeln Formeln „deduziert" werden.

zweitens als axiomatische Formel und drittens als Operationsregel, nach welcher aus „beweisbaren Formeln" beweisbare Formeln entspringen.

[16]) a. a. O. [12]), S. 164.

[17]) Vgl. die unter [11]) zitierte Arbeit.

[18]) In der unter [13]) zitierten Arbeit.

Schachspiel	Hilbertsche Mathematik
5. „Spielgerechte“ Stellung (die auf Grund der Zugregeln aus der Ausgangsstellung hervorgegangen ist).	„Beweisbare“ Formel (welche auf Grund der Operationsregeln aus den Axiomen hervorgegangen ist).
6. Stellung, in welcher 10 Damen der gleichen Farbe auftreten.	Formel des Widerspruchs.

Durch die sich im finiten anschaulichen Denken betätigende vollständige Induktion kann man beim Schachspiel ebenso einsehen, daß 10 Damen einer Farbe in einer spielgerechten Stellung unmöglich sind, wie Hilbert beweist (oder zu beweisen sich anschickt), daß die Formel des Widerspruchs keine beweisbare Formel ist. (Man geht dabei natürlich von der Bemerkung aus, daß durch einen spielgerechten Zug die Summe der Anzahlen der Damen und Bauern einer Farbe niemals zunehmen kann.) Das ist *Erkenntnis* und nicht mehr *Spiel*; und nur zur Gewinnung dieser einen Erkenntnis wird von Hilbert das inhaltliche Denken benötigt. Auf diesem Standpunkt darf man nicht nach einem tieferen Grund für die angenommenen Axiome und Operationsregeln fragen; auch ist nicht abzusehen, warum man gerade Wert darauf legt, daß das Formelspiel „widerspruchsfrei“ ist oder warum man das inhaltliche Denken sich nicht noch mit andern aus dem Spiel entspringenden Fragen beschäftigen läßt. Solange man auf diesem Standpunkt beharrt, ist man in der Tat aller Philosophie überhoben; und es gibt keine andere „Einwendung“ dagegen als die Erklärung: Ich spiele nicht mit!

Soll aber Mathematik eine ernsthafte Kulturangelegenheit bleiben, so muß sich nun doch mit diesem Formelspiel irgend ein *Sinn* verknüpfen. — Zunächst kann es sich in den Dienst der inhaltlichen Analysis, insbesondere der inhaltlichen Arithmetik stellen. Als das berühmteste Beispiel, in welchem heute transfinite Schlüsse zur Begründung eines finiten Theorems herangezogen werden, nenne ich Dirichlets Klassenzahlbestimmung quadratischer Formen. Definiert man die zur Diskriminante D gehörige Klassenzahl h mit Hilfe einer finiten Reduktionsmethode, so hat man zwei Wege zur Konstruktion von h: der eine ist durch dieses Reduktionsverfahren gegeben, der andere durch Dirichlets explizite Formel. Ist die Behauptung, daß aus einer beliebigen Zahl D sich auf beiden Wegen die gleiche Zahl h ergibt, durch Hilberts formale Analysis transfinit begründet, so ist sie inhaltlich richtig. Ergäbe sich nämlich für einen bestimmten Wert, z. B. $D = 129$, auf dem einen Wege $h = 5$, auf dem andern $h = 7$, so wäre $5 = 7$ und damit $0 = 2$ eine beweisbare Formel im Hilbertschen System, die mit $2 \neq 0$ zusammen zu der Formel des Widerspruchs $0 \neq 0$ führt. So oft also eine beweisbare Hilbertsche Formel, inhaltlich interpretiert, einen Sinn gibt — dazu ist insbesondere erforderlich, daß sie die transfiniten Funktionen nicht enthält —, spricht sie einen richtigen Satz der inhaltlichen Analysis aus: der Formalismus ist auf Grund seiner Widerspruchslosigkeit ein legitimes Hilfsmittel zur Gewinnung derartiger Sätze. Immerhin wird man den Wert dieses formalen Hilfsmittels nicht allzu hoch anschlagen. Denn nicht im *Beweis* bei gegebener Konstruktion, sondern in der Erfindung der *Konstruktion* liegt in den meisten Fällen die eigentliche Schwierigkeit. *Nachdem* einmal die vorhin besprochene Konstruktion gefunden ist, welche aus der durch freie Wahl werdenden Schachtelfolge von Dualsternen im Koeffizientenraum die werdende Schachtelfolge von Quadraten im Wurzelraum erzeugt, gelingt der Nachweis mühelos, daß sie die Wurzeln der vorgelegten willkürlichen Gleichung liefert. Und so liegt die Sache fast immer.

Mit dieser untergeordneten Rolle seiner Formeln im Dienste der Brouwerschen Analysis wird sich Hilbert schwerlich zufrieden geben. Ich sehe nur *eine* Möglichkeit, ihnen einschließlich ihrer transfiniten Bestandteile eine selbständige geistige Bedeutung beizulegen. In der theoretischen Physik haben wir das große Beispiel einer Erkenntnis von ganz anderem Gepräge vor uns als die gewöhnliche intuitive oder phänomenale Erkenntnis, welche das in der Anschauung[19]) Gegebene rein ausspricht. Während hier jedes Urteil seinen eigenen, restlos in der Anschauung vollziehbaren Sinn hat, ist dies mit den einzelnen Aussagen der theoretischen Physik keineswegs der Fall; sondern dort steht, wenn es mit der Erfahrung konfrontiert wird, nur das System als Ganzes in Frage. In der *Theorie* gelingt es dem Bewußtsein, „über den eigenen Schatten zu springen", den Stoff des Gegebenen hinter sich zu lassen, das Transzendente darzustellen; aber, wie sich von selbst versteht, nur im *Symbol*. Die Beziehung der symbolischen Konstruktion zum unmittelbar Erlebten muß, wenn nicht explizite beschrieben, so doch irgendwie innerlich verstanden sein; aber diese Beziehung allein kann niemals die theoretische Deutung rechtfertigen. Hier walten Vernunftprinzipien, von denen wir vorerst nur das der Widerspruchslosigkeit klar erfassen; es ist aber gewiß nicht der einzige Leitfaden bei der Ausbildung der theoretischen Physik (der Sinn der *theoretischen* ist uns im Grunde ebenso dunkel wie der Sinn der *künstlerischen* Gestaltung). Wenn ich die phänomenale Erkenntnis als *Wissen* bezeichne, so ruht die theoretische auf dem *Glauben*[20]) — dem Glauben an die

[19]) „Anschauung" wird hier natürlich nicht aufs Sinnliche beschränkt, sondern bezeichnet jeden gebenden Akt.

[20]) Daß der Glaube, d. i. die transzendente Vernunftthesis, ohne welche alles Wissen tot und völlig gleichgültig ist, nicht erst bei „Gott, Freiheit und Unsterblichkeit" auf den Plan tritt, war eine der ersten und wichtigsten Erkenntnisse, welche Fichte über Kant hinausführten.

Realität des eigenen und fremden Ich oder die Realität der Außenwelt oder die Realität Gottes. Ist das Organ jener das „Sehen“ im weitesten Sinne, so ist das Organ der Theorie „das Schöpferische“. Wenn Hilbert nicht ein bloßes Formelspiel treibt, so will er eine theoretische im Gegensatz zu Brouwers intuitiver Mathematik. Aber wo ist jenes vom Glauben getragene Jenseits, auf das sich ihre Symbole richten? Ich finde es nicht, wenn ich nicht die Mathematik sich völlig mit der Physik verschmelzen lasse und annehme, daß die mathematischen Begriffe von Zahl, Funktion usw. (oder die Hilbertschen Symbole) prinzipiell in der gleichen Art an der theoretischen Konstruktion der wirklichen Welt teilnehmen wie die Begriffe Energie, Gravitation, Elektron u. dgl. Dann aber ist das System kaum durch seine Widerspruchslosigkeit schon ausreichend gerechtfertigt, und es steht ferner zu erwarten, daß es das Schicksal aller andern theoretischen Erkenntnisse teilen wird: im Gegensatz zum phänomenalen Wissen, das wohl dem Irrtum menschlich unterworfen, aber seinem Wesen nach unwandelbar ist, bleiben sie, wie ich glaube, getragen von dem an uns sich vollziehenden Lebensprozeß des Geistes und werden von ihm niemals als ein totes, „endgültiges“ Resultat abgesetzt. Vielleicht ist es aber ja doch so, wie Hilbert zu meinen scheint, daß für den mathematischen Teil der theoretischen Weltkonstruktion das Prinzip der Widerspruchslosigkeit zusammen mit der Forderung, die vom naiven Existentialabsolutismus über das Unendliche nach Analogie des Endlichen aufgestellten Behauptungen so weitgehend wie möglich symbolisch zu rechtfertigen, als einziger Leitfaden genügt. In der Geschichte der Physik zeigt sich, daß Anschauung und Theorie beständig Hand in Hand gehen müssen. Auf der einen Seite ist z. B. nicht zu leugnen, daß der Machsche Phänomenalismus der Atomtheorie unterlegen ist; auf der andern Seite aber lehrte Einsteins Relativitätstheorie, wie wichtig der Rückgang auf die anschauliche Bedeutung der theoretischen Konstruktion und die Ausscheidung allzu willkürlicher Elemente (Geometrie, absoluter Raum) sein kann. So ist es sicherlich von großem Nutzen, daß uns Brouwer in der Mathematik wieder den Sinn für das anschaulich Gegebene gestärkt hat. Seine Analysis spricht den Gehalt der mathematischen Urintuition rein aus und ist darum von rätselloser Klarheit durchleuchtet. Aber neben dem Brouwerschen wird man den Hilbertschen Weg verfolgen müssen; denn es ist nicht zu leugnen, daß in uns ein vom bloß phänomenalen Standpunkt schlechterdings unverständliches theoretisches Bedürfnis lebendig ist, dessen auf symbolische Darstellung des Transzendenten gerichteter Schaffensdrang Befriedigung verlangt. Aber hier beginnt das Rätsel, und so sind auch diese Schlußgedanken nur ein zaghaftes Hinaustasten ins Dunkel.

(Eingegangen am 25. Oktober 1923.)