German Exam

From: Mathematische Annalen, Vol 98, 1928,



# Über die Grundlagen der Quantenmechanik.


Von

D. Hilbert, J. v. Neumann und L. Nordheim in Göttingen.


## § 1.

## Einleitung.

Die neuere Entwicklung der Quantenmechanik, die an die Arbeiten von Heisenberg[1]), Born und Jordan einerseits und Schrödinger[2]) andererseits anknüpft, hat es ermöglicht, das ganze Gebiet der Atomerscheinungen unter einen einheitlichen Gesichtspunkt zusammenzufassen und die wichtigsten Beobachtungsergebnisse zu erklären, so daß man kaum noch zweifeln kann, in ihr einen Gedankenkomplex von gleicher Bedeutung wie etwa die klassische Mechanik oder Elektrodynamik gefunden zu haben. Bei dieser hohen Bedeutung der Quantenmechanik ist es ein dringendes Bedürfnis, ihre Prinzipien so klar und allgemein wie möglich zu erfassen. Hier hat nun Jordan[3]) im Anschluß an eine Idee von Pauli[4]) eine Formulierung und Begründung der Theorie gegeben, die von großer Allgemeinheit ist und dem physikalischen Charakter der zu grunde liegenden Erscheinungen sehr gut angepaßt erscheint. Unabhängig von Jordan ist Dirac[5]) zu ähnlichen Anschauungen gelangt.

Die vorliegende Abhandlung ist aus einer Vorlesung hervorgegangen, die D. Hilbert im Wintersemester 1926/27 über die neuere Entwicklung der Quantenmechanik hielt, und deren Vorbereitung mit wesentlicher Hilfe von L. Nordheim geschah. Wichtige Stücke der mathematischen Durch-

[1]) S. z. B. W. Heisenberg, Math. Annalen 95 (1926), S. 683.

[2]) E. Schrödinger, Abhandlungen zur Wellenmechanik, Leipzig 1927.

[3]) P. Jordan, Zeitschr. f. Phys. 40 (1927), S. 204 und Gött. Nachr. 1926, S. 162. Die formalen Zusammenhänge sind zum Teil auch unabhängig von F. London. Zeitschr. f. Phys. 40 (1926), S. 193 gefunden worden.

[4]) W. Pauli jr., Zeitschr. f. Phys. 41 (1927), S. 21.

[5]) P. A. M. Dirac, Proc. Royal Soc. (A), 113 (1927), S. 621.

führung rühren von J. v. Neumann her. Die endgültige Abfassung dieser Abhandlung hat L. Nordheim besorgt. Wir glauben, daß die Formulierung der Jordanschen und Diracschen Gedanken, zu der wir hier gelangt sind, erheblich einfacher und daher durchsichtiger und leichter verständlich geworden ist.

Der physikalische Grundgedanke der ganzen Theorie besteht darin, daß an Stelle von strengen funktionalen Beziehungen der gewöhnlichen Mechanik überall Wahrscheinlichkeitsrelationen treten.

Die Art dieser Relationen wird am besten durch ein besonders wichtiges Beispiel erklärt. Ist der Wert $W_n$ der Energie des Systems bekannt, und zwar gleich dem $n$-ten Eigenwert eines gequantelten Systems, so ist nach Pauli

$$|\psi_n(x)|^2$$

die Wahrscheinlichkeitsdichte dafür, daß die Koordinate des Systems einen Wert zwischen $x$ und $x + dx$ hat, wenn $\psi_n$ die zu dem Eigenwert $W_n$ gehörige Eigenfunktion bedeutet.

Allgemein wird verlangt, daß zwischen zwei verschiedenen mechanischen Größen $F_1$ und $F_2$ stets eine solche Wahrscheinlichkeitsbeziehung besteht. Unter mechanischer Größe verstehen wir hierbei Dinge wie Entfernungen, Bahnradius, Energie, Impuls, d. h. Größen, die in der gewöhnlichen Mechanik Funktionen einer mechanischen Koordinate $q$ und des zu ihr kanonisch konjugierten Impulses $p$ sind.

Wir beschränken uns dabei der Einfachheit halber auf Systeme mit einem einzigen Freiheitsgrad. Jedoch ist diese Beschränkung keineswegs wesentlich, und es lassen sich Systeme mit mehreren Freiheitsgraden ganz analog behandeln. Auf die Wahl des Koordinatensystems kommt es ferner, wie sich herausstellen wird, nicht an.

Der Weg, der nun zu dieser Theorie führt, ist folgender: Man stellt gewisse physikalische Forderungen an diese Wahrscheinlichkeiten, die durch unsere bisherigen Erfahrungen und Entwicklungen nahe gelegt sind, und deren Erfüllung gewisse Relationen zwischen den Wahrscheinlichkeiten erfordern. Dann sucht man zweitens einen einfachen analytischen Apparat, in dem Größen auftreten, die genau dieselben Relationen erfüllen. Dieser analytische Apparat, und damit die in ihm auftretenden Rechengrößen, erfahren nun auf Grund der physikalischen Forderungen eine physikalische Interpretation. Das Ziel ist dabei, die physikalischen Forderungen so vollständig zu formulieren, daß der analytische Apparat gerade eindeutig festgelegt wird. Dieser Weg ist also der einer Axiomatisierung, wie sie z. B. in der Geometrie durchgeführt worden ist. Durch die Axiome werden die Relationen zwischen den geometrischen Gebilden, wie Punkt, Gerade,

ie endgültige Abfassung dieser
glauben, daß die Formulierung
zu der wir hier gelangt sind,
ger und leichter verständlich

ganzen Theorie besteht darin,
Beziehungen der gewöhnlichen
ien treten.
ten durch ein besonders wich-
Energie des Systems bekannt,
gequantelten Systems, so ist

Koordinate des Systems einen
die zu dem Eigenwert $W_n$ ge-

zwei verschiedenen mechani-
Wahrscheinlichkeitsbeziehung
hen wir hierbei Dinge wie
d. h. Größen, die in der ge-
hanischen Koordinate $q$ und
sind.
eit halber auf Systeme mit
se Beschränkung keineswegs
ehreren Freiheitsgraden ganz
tensystems kommt es ferner,

rt, ist folgender: Man stellt
rscheinlichkeiten, die durch
ngen nahe gelegt sind, und
len Wahrscheinlichkeiten er-
achen analytischen Apparat,
Relationen erfüllen. Dieser
auftretenden Rechengrößen,
lerungen eine physikalische
ikalischen Forderungen so
Apparat gerade eindeutig
r Axiomatisierung, wie sie
Durch die Axiome werden
bilden, wie Punkt, Gerade,



Ebene, beschrieben, und dann gezeigt, daß diese Relationen gerade ebenso bei einem analytischen Apparat, nämlich den linearen Gleichungen erfüllt sind. Dadurch kann man wieder umgekehrt aus den Eigenschaften der linearen Gleichungen geometrische Sätze gewinnen.

[redacted] einer bestimmten Vorschrift jeder mechanischen Größe ein mathematisches Gebilde als Repräsentanten zu, das [illegible] eine [illegible] Rechengröße ist, aus der man aber Au[illegible]n über die Repräsentanten anderer Größen, und dann durch Zurückübersetzung Aussagen über wirkliche physikalische Dinge erhalten kann.

Go to p. 4

Solche Repräsentanten sind die Matrizen in der Heisenbergschen, die $q$-Zahlen in der Diracschen und die Operatoren in der Schrödingerschen Theorie und ihrer jetzigen Weiterbildung.

Es ist also zu beachten, daß wir zwei ganz verschiedene Klassen von Dingen betrachten, nämlich einerseits die meßbaren Zahlenwerte physikalischer Größen und andererseits die ihnen zugeordneten Operatoren, mit denen lediglicch nach den Regeln der Quantenmechanik gerechnet wird.

Das oben angedeutete Verfahren der Axiomatisierung wird nun in der Physik gewöhnlich nicht genau so befolgt, sondern der Weg zur Aufstellung einer neuen Theorie ist, wie in der Regel, so auch hier, folgender.

Man mutmaßt meistens den analytischen Apparat, bevor man noch das vollständige Axiomensystem aufgestellt hat, und kommt dann erst durch die Interpretation des Formalismus zur Aufstellung der physikalischen Grundrelationen. Es ist schwer, eine solche Theorie zu verstehen, wenn man diese beiden Dinge, den Formalismus und seine physikalische Interpretation, nicht scharf genug auseinanderhält. Diese Scheidung soll hier möglichst deutlich durchgeführt werden, wenn wir auch, dem jetzigen Zustand der Theorie entsprechend, noch nicht eine vollständige Axiomatik begründen wollen. Das, was jedenfalls eindeutig festliegt, ist der analytische Apparat, der — rein mathematisch — auch keiner Abänderung fähig ist. Was dagegen modifiziert werden kann, und vorausichtlich auch noch werden wird, ist die physikalische Interpretation, bei der eine gewisse Freiheit und Willkür besteht.

Durch die Axiomatisierung verlieren die vorher etwas vagen Begriffe, wie Wahrscheinlichkeit und so weiter, ihren mystischen Charakter, da sie dann durch die Axiome implizit definiert sind.

## § 2.

### Die physikalischen Axiome.

Wir gehen jetzt dazu über, die physikalischen Forderungen anzugeben. Ihr Sinn wird allerdings vielleicht erst dann völlig klar werden, wenn ihr Zusammenhang mit dem Formalismus erläutert ist. Die Schwierigkeit der neuen Vorstellungen beruht zum großen Teil darauf, daß diese physikalischen Forderungen selbst sehr fremd und neuartig erscheinen. Es sei ferner gleich hier darauf hingewiesen, daß die Forderungen I—VI nur die allgemeinen Richtlinien für die Wahl des analytischen Apparates gehen sollen, aber keine eindeutige logische Festlegung.

Zunächst fassen wir unser obiges Postulat von dem Wahrscheinlichkeitszusammenhang in eine Formel.

I. Zu zwei mechanischen Größen $F_1(pq)$ und $F_2(pq)$ gibt es stets eine Funktion zweier Variablen

$$\varphi(xy;\, F_1 F_2)$$

derart, daß

$$\varphi(xy;\, F_1 F_2)\, \bar{\varphi}(xy;\, F_1 F_2) = w(xy;\, F_1 F_2)$$

die relative Wahrscheinlichkeitsdichte dafür ist, daß für gegebenen Wert $y$ von $F_2$ der Wert von $F_1$ zwischen $x$ und $x + dx$ liegt.

$\varphi$ heißt dabei die *Wahrscheinlichkeitsamplitude*. Diese Definition bedeutet also, daß die Wahrscheinlichkeit, daß für festes $y$ der Wert von $x$ im Intervall $ab$ liegt, proportional zu

$$\int_a^b w(xy;\, F_1 F_2)\, dx$$

ist. Nur Aussagen dieser Art sollen einen physikalischen Sinn haben. Da dieses Integral im allgemeinen noch von $y$ abhängt, so sieht man, daß unsere Wahrscheinlichkeiten nur relative sein können [6]).

II. Die Amplituden bzw. Wahrscheinlichkeiten sollen weiter noch folgenden Bedingungen genügen. Die Wahrscheinlichkeitsdichte, daß für einen gegebenen Wert von $F_1 = x$, $F_2$ einen Wert zwischen $y$ und $y + dy$ habe, sei durch dieselbe Funktion $w(xy;\, F_1 F_2)$ gegeben. Diese Forderung zeigt übrigens wieder, daß unsere Theorie nur *relative* Wahrscheinlichkeiten geben kann.

III. Weiter wird zu verlangen sein, daß für die Beziehung einer Größe zu sich selbst, also für $w(xy;\, F_1 F)$, eine scharfe Bestimmung eintritt, derart, daß also jeder mechanischen Größe wirklich ein bestimmter Zahlenwert zugeordnet werden kann.

[6]) Über die Normierungsmöglichkeiten siehe P. Jordan, Zeitschr. f. Phys. 41 (1927); S. 797.

IV. Eine weitere sehr wichtige Beziehung stellt die Kompositionsregel zweier Amplituden dar. Seien $F_1$, $F_2$, $F_3$ drei mechanische Größen, und $x, y, z$ die zugehörigen Zahlenwerte, so soll die Beziehung gelten

$$\varphi(xz;\, F_1 F_3) = \int \varphi(xy;\, F_1 F_2)\, \varphi(yz;\, F_2 F_3)\, dy.$$

Diese Forderung stellt offenbar ein Analogon zu dem Additions- und Multiplikationstheorem der gewöhnlichen Wahrscheinlichkeitsrechnung dar, nur daß sie hier für die Amplituden anstatt für die Wahrscheinlichkeiten selbst gelten soll.

Hierin besteht eben der charakteristische Unterschied zu der gewöhnlichen Wahrscheinlichkeitsrechnung, daß überall an Stelle der Wahrscheinlichkeiten selbst zunächst die Amplituden auftreten, die im allgemeinen komplexe Größen werden, und erst absolut genommen und quadriert gewöhnliche Wahrscheinlichkeiten ergeben.

V. Als weitere physikalische Forderung ist noch zu stellen, daß die Wahrscheinlichkeiten nur von der funktionalen Natur der Größen $F_1(pq)$ und $F_2(pq)$ abhängen, also ihrer kinematischen Verknüpfung, und nicht etwa noch von speziellen Eigenschaften des gerade untersuchten mechanischen Systems, wie etwa seiner Hamiltonschen Funktion.

VI. Als letzte Forderung ist schließlich zu stellen, daß die Wahrscheinlichkeiten unabhängig von dem speziell gewählten Koordinatensystem werden.

Diesen Forderungen, deren eventuelle wechselseitige Abhängigkeit und Vollständigkeit wir nicht weiter untersuchen wollen, wird nun genügt durch den mathematischen Formalismus einer allgemeinen Operatorenrechnung. Die Schrödingerschen Differentialgleichungen und die eingangs erwähnte Deutung der Eigenfunktionen kommen dabei als Spezialfälle heraus, was natürlich die stärkste Stütze für die Richtigkeit der hier dargestellten Anschauungen bildet.

Nach der Erklärung der logischen Struktur der Theorie kommen wir zu dem Hauptpunkt, dem analytischen Formalismus. Dabei sei besonders betont, daß wir in der vorliegenden Darstellung zugunsten einer leichteren Verständlichkeit und Durchsichtigkeit auf vollständige mathematische Exaktheit verzichten, und es insbesondere dahingestellt sein lassen, abzugrenzen, in welchem Bereich einzelne Operationen durchführbar bzw. zulässig sind.

## § 3.

## Grundformeln der Operatorrechnung.

Stop

Der Zusammenhang der Schrödingerschen Theorie mit der Heisenbergschen Matrizenmechanik wird bekanntlich durch eine Operatorenrechnung

From Mathematische Zeitschrift, Vol. 23, 1925

# Über eine paradoxe Eigenschaft gewisser bedingt konvergenter unendlicher Reihen.

Von

Konrad Knopp in Königsberg.

## I.

Ist eine unendliche Reihe

$$(1) \qquad \sum_{n=1}^{\infty} a_n$$

*absolut* konvergent, so ist bekanntlich auch jede Teilreihe derselben, insbesondere für jedes $k = 1, 2, 3, \ldots$ die Reihe

$$(2) \qquad \sum_{n=1}^{\infty} a_{kn}$$

konvergent; und wenn wir die Summe der letzteren mit $\sigma_k$ bezeichnen, so strebt mit wachsendem $k$

$$(3) \qquad \sigma_k \to 0,$$

da ja $|\sigma_k| \leqq |a_k| + |a_{k+1}| + \ldots$, also höchstens gleich dem $k$-ten Rest der nach Voraussetzung konvergenten Reihe $\Sigma |a_n|$ ist. Es liegt nahe zu prüfen, ob und inwieweit diese Tatsachen bestehen bleiben, wenn (1) nur *bedingt* konvergiert.

Es ist an und für sich nicht überraschend, daß die Dinge hier verschieden liegen. Indessen sind sie *so wesentlich* anders geartet, daß ihre besondere Feststellung Interesse beanspruchen darf.

1. Ich beginne mit dem Nachweis der *besonders paradoxen* Tatsache, *daß es konvergente Reihen* $\Sigma a_n$ *gibt, für die zwar jede der Reihen* (2) *konvergiert, für die auch* $\lim \sigma_k$ *existiert, für die aber dieser Grenzwert nicht* $= 0$, *sondern etwa* $= 1$ *ist*[1]).

[1]) Die Frage nach der Existenz solcher Reihen wurde gelegentlich einer Untersuchung über Fourierreihen von Herrn W. Rogosinski aufgeworfen.

Eine Reihe (1) mit dieser Eigenschaft kann man folgendermaßen bilden: Für

$$(4)\qquad \begin{cases} n = 1;\ \ 4,\ \ 5,\ \ 6,\ \ 7;\ \ 66,\ \ 74,\ \ 78,\ \ 82,\ \ 102,\ \ 106 \\ \text{sei} \\ a_n = 1;\ +\frac{1}{2},\ -\frac{1}{2},\ +\frac{1}{2},\ -\frac{1}{2};\ +\frac{1}{3},\ -\frac{1}{3},\ +\frac{1}{3},\ -\frac{1}{3},\ +\frac{1}{3},\ -\frac{1}{3} \end{cases}$$

und für die übrigen $n < 106$ sei $a_n = 0$. Allgemein sollen die Glieder der Reihe, von vielen zwischengeschalteten Nullen abgesehen, Gruppen der Form

$$(5)\qquad +\frac{1}{m},\ -\frac{1}{m},\ +\frac{1}{m},\ -\frac{1}{m},\ \ldots,\ +\frac{1}{m},\ -\frac{1}{m}$$

mit $2m$ von 0 verschiedenen Gliedern bilden, die abwechselnd gleich $+\frac{1}{m}$ und $-\frac{1}{m}$ sind. Die Nenner $m$ in diesen Gruppen sollen dabei monoton ins Unendliche wachsen, so daß die Konvergenz von $\Sigma a_n$ von vornherein gesichert ist und die Summe gleich 1 sein wird. Die hingeschriebenen Glieder entsprechen, von $a_1$ abgesehen, den Werten $m = 2$ und $m = 3$. Ferner sind die Nummern der Glieder $-\frac{1}{2}$, die halben Nummern der Glieder $+\frac{1}{2}$ und $-\frac{1}{3}$, sowie der sechste Teil der Nummern der Glieder $+\frac{1}{3}$ lauter verschiedene Primzahlen.

Allgemein lege ich nun die zu verwendenden Nenner $m$ und die Indizes $n$, welche den von 0 verschiedenen Gliedern gegeben werden sollen, folgendermaßen fest:

Es sei $v_h$ das kleinste gemeinsame Vielfache der Zahlen $1, 2, \ldots, h$. Dann soll $m$ diejenigen natürlichen Zahlen $\geqq 2$ durchlaufen, für die $v_m > v_{m-1}$ (und folglich $v_{m-1} \mid v_m$) ist, d. h. die Zahlen

$$(6)\qquad m = 2,\ 3,\ 4,\ 5,\ 7,\ 8,\ 9,\ 11,\ 13,\ 16,\ 17,\ 19,\ \ldots .$$

Es seien nun schon bis zu einer bestimmten von diesen Zahlen, etwa bis $m_0$, die Gruppen (5) hingeschrieben und ihren Gliedern bestimmte (wachsende) Nummern $n$ gegeben, deren höchste $n_0$ sei, — wie dies unter (4) für $m_0 = 3$ und $n_0 = 106$ geschehen. Es sei überdies die Nummer $n'$ eines Gliedes $+\frac{1}{m'}$ stets durch $v_{m'}$, die Nummer $n''$ eines Gliedes $-\frac{1}{m''}$ stets durch $v_{m''-1}$ teilbar, und die Quotienten seien dabei *lauter verschiedene* Primzahlen $p'$ bzw. $p''$, wie dies bei (4) für die Glieder $\pm\frac{1}{2}$ und $\pm\frac{1}{3}$ der Fall ist, denen die Primzahlen $p' = 2, 3, 11, 13, 17$ und $p'' = 5,\ 7,\ 37,\ 41,\ 53$ entsprechen. Endlich sei jedes $p''$ größer als das entsprechende $v_{m''}$. Die hierbei bis $n_0$ benutzten Nummern $n$ sowie die bis dahin aufgetretenen Primzahlen $p'$ und $p''$ mögen als „verbraucht" gelten.

Es sei dann $m$ die auf $m_0$ folgende Zahl der Folge (6), und $r$ so gewählt, daß $v_m^r > n_0$ ist. Dann wähle man $2m$ natürliche Zahlen $x_1, x_2, \ldots, x_m;\ y_1, y_2, \ldots, y_m$, so daß

1. $$x_\lambda v_m < y_\lambda v_{m-1} \quad \text{für } \lambda = 1, 2, \ldots, m \quad \text{und}$$
$$y_\lambda v_{m-1} < x_{\lambda+1} v_m \quad \text{für } \lambda = 1, 2, \ldots, m-1$$

ist, und daß

2. die $2m$ Zahlen
$$v_m^{r-1} + x_\lambda \quad \text{und} \quad \frac{v_m}{v_{m-1}} v_m^{r-1} + y_\lambda \qquad (\lambda = 1, 2, \ldots, m)$$

lauter Primzahlen sind, die voneinander und von den bisher verbrauchten Primzahlen verschieden sind, sowie oberhalb $v_m$ liegen.

Dann geben wir

$$\text{den } m \text{ Gliedern } +\frac{1}{m} \text{ die Nummern } n = v_m^r + x_\lambda v_m,$$
$$\text{,,}\ m \quad \text{,,} \quad -\frac{1}{m} \quad \text{,,} \quad \text{,,} \quad n = v_m^r + y_\lambda v_{m-1}$$
$$(\lambda = 1, 2, \ldots, m)$$

und setzen für die übrigen $n$, soweit $n_0 < n < v_m^r + y_m v_{m-1}$ ist, stets $a_n = 0$.

*Die so gebildete Reihe* (1) *leistet das Verlangte.* Denn sie ist, wie schon betont, sicher konvergent (und ihre Summe $= 1$). Ist dann $k \geqq 2$ beliebig gegeben, so sei $m = m(k)$ die kleinste Zahl der Folge (6), für die $k \mid v_m$. Zur Summe der Reihe (2) trägt dann die diesem $m$ entsprechende Gruppe offenbar $m$-mal $+\frac{1}{m}$, also $+1$, bei. Die Glieder aller folgenden Gruppen haben Nummern, die *sämtlich* durch $k$ teilbar sind. Die Reihe (2) ist also jedenfalls konvergent und ihre Summe $\sigma_k$ ist $= +1$, vermehrt um die Summe derjenigen Glieder $a_n$, die der eben genannten $m$-Gruppe vorangehen. Unter diesen ist aber *höchstens ein* von 0 verschiedenes Glied. Denn deren Nummern haben die Form $v_h \cdot p$, bei der $h < m$ ist und $p$ eine Primzahl bedeutet. Es kann aber $k$ nicht Teiler zweier verschiedener solcher Nummern sein. Denn wäre

$$\text{neben} \quad k \mid v_h \cdot p \quad \text{zugleich noch} \quad k \mid v_{h'} \cdot p',$$

so müßte, wenn wir $h' \leqq h$ annehmen und den größten gemeinsamen Teiler $(v_h, k) = d$ und $k : d = k'$ setzen, auch neben $k' \mid p$ noch $k' \mid p'$ sein. Das ist aber unmöglich, weil $k' > 1$ und $p$ und $p'$ *verschiedene* Primzahlen sind. Es ist also

$$\text{entweder} \quad \sigma_k = 1 \quad \text{oder} \quad \sigma_k = 1 + \varepsilon_k,$$

wenn mit $\varepsilon_k$ ein bestimmtes von 0 verschiedenes Reihenglied bezeichnet wird, dessen Nummer $\geqq k$ ist. Wegen $a_n \to 0$ strebt also auch $\varepsilon_k \to 0$ und daher, wie behauptet, $\sigma_k \to 1$.

2. Addiert man zu der eben konstruierten Reihe gliedweise irgendeine absolut konvergente Reihe, so erhält man offenbar wieder eine Reihe (1), für die $\sigma_k \to 1$ strebt.

Stop

3. Multipliziert man irgendeine dieser Reihen gliedweise mit der Konstanten $\alpha$, so erhält man eine Reihe (1), für die $\sigma_k \to \alpha$, also gegen einen beliebig wählbaren (auch komplexen) Wert strebt.

4. Eine leichte Modifikation liefert uns nun auch (reelle) *konvergente Reihen* (1), *für die wieder alle* $\sigma_k$ *existieren, für die nun aber* $\sigma_k \to +\infty$ *strebt.* Dazu braucht man nur den Gruppen (5) nicht wie bisher $2m$, sondern etwa $2m^2$ Glieder zu geben und alle übrigen Vorschriften sinngemäß anzupassen. Dann ist offenbar $\sigma_k = m$ oder $= m + \varepsilon_k$ und es strebt, da $m$ zugleich mit $k$ ins Unendliche wächst, nun $\sigma_k \to +\infty$.

Kehrt man alle Zeichen um, so erhält man eine Reihe (1), für die $\sigma_k \to -\infty$ strebt.

5. Ebenso leicht lassen sich nun (reelle) Reihen bilden, für die $\lim \sigma_k$ überhaupt nicht existiert[2]), sondern etwa — um gleich ein sehr allgemeines Verhalten anzugeben — die Häufungsgrenzen

$$\underline{\lim}\, \sigma_k = \varkappa \quad \text{und} \quad \overline{\lim}\, \sigma_k = \mu$$

[2]) Herr G. Szegö, dem ich den Inhalt dieser Note mitteilte, gab mir (8. 12. 24) ein in gewisser Hinsicht natürlicheres Beispiel einer Reihe $\Sigma a_n$ an, bei der alle $\sigma_k$ existieren, aber $\lim \sigma_k$ nicht vorhanden ist. Dies leistet nämlich für geeignetes $x$ schon die Reihe $\sum \frac{\cos 2\pi n x}{n}$, bei der für jedes irrationale $x$ alle $\sigma_k$ existieren und die Werte

$$\sigma_k = -\frac{1}{k} \log |\, 2 \sin k \pi x \,|$$

haben. Wird nun $x$ in $0 \ldots 1$ so gewählt, daß bei seiner dyadischen Entwicklung

$$x = \frac{z_1}{2} + \frac{z_2}{2^2} + \frac{z_3}{2^3} + \ldots \qquad (z_\nu = 0 \text{ oder } 1),$$

für unendlich viele $\nu$

$$z_\nu = 1, \quad z_{\nu+1} = z_{\nu+2} = \ldots = z_{\nu+2^\nu} = 0$$

ist, so ist für diese $\nu$, wenn $2^\nu = m$ gesetzt wird,

$$|\sin m \pi x| < \frac{\pi}{2^m},$$

$$\sigma_m > \frac{1}{m} \log \frac{2^m}{2\pi} = \log 2 - \frac{\log 2\pi}{m},$$

so daß $\overline{\lim}\, \sigma_k \geqq \log 2$ ist. Und da hier offenbar $\underline{\lim}\, \sigma_k = 0$ ist, so ist $\lim \sigma_k$ sicher nicht vorhanden.

German Exam Spring 02 (8)

# 2.

## Approximation algebraischer Zahlen

Mathematische Zeitschrift 10 (1921), 173—213

### Einleitung.

1. Im Jahre 1851 zeigte Liouville (Liouville 1): Zu jeder algebraischen Zahl $\xi$ vom Grade $n \geqq 2$ gibt es ein positives $a_1 = a_1(\xi)$ derart, daß für alle Paare $x, y$ ganzer rationaler Zahlen mit $y > 0$ die Ungleichung

$$\text{(A)} \qquad \left|\xi - \frac{x}{y}\right| > \frac{a_1}{y^n}$$

gilt.

2. Diese fast triviale Abschätzung verbesserte Thue (Thue 3) 1908 zu folgender: Zu jedem $\varepsilon > 0$ gibt es ein positives $a_2 = a_2(\xi, \varepsilon)$, so daß (A) durch die schärfere Ungleichung

$$\text{(B)} \qquad \left|\xi - \frac{x}{y}\right| > \frac{a_2}{y^{\frac{n}{2}+1+\varepsilon}}$$

ersetzt werden kann. Mit Hilfe hiervon bewies dann Thue seinen bekannten Satz: Bedeutet $U(x,y)$ eine homogene binäre Form mit rationalen Koeffizienten, die nicht Potenz einer linearen oder indefiniten quadratischen Form ist, so hat die Diophantische Gleichung $U(x,y) = k$ für jedes rationale $k \neq 0$ nur endlich viele Lösungen in ganzen rationalen $x, y$.

Bereits Thue wies auf eine naheliegende Erweiterung dieses Satzes hin, die dann von Maillet (Maillet 2) 1916 ausgeführt wurde. Ist nämlich $n$ der Grad von $U(x,y)$ und $V(x,y)$ irgendein zu $U(x,y)$ teilerfremdes Polynom, dessen Dimension $< \frac{n}{2} - 1$ ist, so bleibt der Satz von Thue richtig, wenn $k$ durch $V(x,y)$ ersetzt wird.

Thue zeigte ferner mit Hilfe seines Satzes, daß der größte Primteiler des Polynoms $f(x) = (ax+b)(cx+d)$, wo $a, b, c, d$ ganze rationale Zahlen mit $ad - bc \neq 0$, $a \neq 0$, $c \neq 0$ bedeuten, über alle Grenzen wächst, wenn $x$ die Reihe der natürlichen Zahlen wachsend durchläuft. Für spezielle

Werte der Koeffizienten (sowie für $f(x) = x^2 + 1$) war dies bereits durch Størmer bekannt (Størmer 1, 2). Pólya hat 1918 gezeigt (Pólya 2), daß diese Aussage bestehen bleibt, wenn $f(x)$ irgendeine quadratische Funktion von nichtverschwindender Diskriminante mit ganzen rationalen Koeffizienten bedeutet.

3. Durch Thues Untersuchungen ist nahegelegt worden, den Exponenten $\frac{n}{2} + 1 + \varepsilon$ weiter zu verkleinern zu suchen. Welches diejenige Funktion von $n$ ist, die den kleinstmöglichen Wert des Exponenten angibt, ist unbekannt. Ich beweise aber in dieser Arbeit, welche ein Abdruck meiner Inaugural-Dissertation (Göttingen 1920) ist: Es gibt ein positives $a_3 = a_3(\xi)$, so daß die Ungleichung

$$\left|\xi - \frac{x}{y}\right| > \frac{a_3}{y^{2\sqrt{n}}}$$

besteht; dies gilt sogar (mit $a_3 = a_3(\xi, \varepsilon)$), wenn $2\sqrt{n}$ durch den besseren Wert $\min\limits_{\lambda = 1, \ldots n} \left(\frac{n}{\lambda + 1} + \lambda\right) + \varepsilon$ ersetzt wird.

Im folgenden erscheint dies als spezieller Fall eines Satzes über die Stärke der Approximation einer algebraischen Zahl durch eine andere Ist nämlich ein algebraischer Zahlkörper $K$ gegeben, in bezug auf welchen $\xi$ vom Grade $d \geqq 2$ ist, bedeutet $\zeta$ eine primitive Zahl von $K$ und $l$ das Maximum der absolut genommenen teilerfremden ganzen rationalen Koeffizienten der irreduziblen Gleichung für $\zeta$, so gilt für ein gewisses $a_4 = a_4(\xi, K) > 0$

$$\text{(C)} \qquad |\xi - \zeta| > \frac{a_4}{l^{2\sqrt{d}}};$$

dies bleibt sogar richtig (mit $a_4 = a_4(\xi, K, \varepsilon)$), wenn $2\sqrt{d}$ durch $\min\limits_{\lambda = 1, \ldots d} \left(\frac{d}{\lambda + 1} + \lambda\right) + \varepsilon$ ersetzt wird.

Beschränkt man sich bei der Approximation von $\xi$ nicht auf Zahlen eines festen Körpers $K$, sondern läßt für $\zeta$ beliebige algebraische Zahlen eines festen Grades $h < n$ zu, so gilt Ungleichung (C), wenn darin $2\sqrt{d}$ durch $2h\sqrt{n}$ $\left(\text{resp. durch } \min\limits_{\lambda = 1, \ldots n} \left(\frac{n}{\lambda + 1} + \lambda\right)h + \varepsilon\right)$ ersetzt wird.

§ 1 enthält den Beweis dieser beiden Sätze.

§ 2 gibt einige Anwendungen derselben. Insbesondere (Satz 5) verallgemeinere ich den Thueschen Satz über Diophantische Gleichungen auf beliebige algebraische Zahlkörper. Satz 7 dehnt den Thue-Pólyaschen Satz über die Primteiler quadratischer Funktionen auf beliebige Polynome in beliebigen Zahlkörpern aus. Satz 9 ist die entsprechende

Aussage über gebrochene rationale Funktionen. Satz 8 liefert eine arithmetische Eigenschaft der Koeffizienten in der Potenzreihenentwicklung gewisser gebrochener rationaler Funktionen.

## § 1.

Hilfssatz I. Es seien $\lambda_1, \ldots \lambda_h$ irgend welche Zahlen. Es bedeute $A$ den größten der absolut genommenen Koeffizienten des Polynoms $\prod_{\nu=1}^{h}(z-\lambda_\nu)$. Dann ist

$$\prod_{\nu=1}^{h}(1+|\lambda_\nu|) \leqq 6^h A.$$

Beweis[1]). Ohne Beschränkung der Allgemeinheit seien $\lambda_1, \ldots \lambda_\tau$ diejenigen von den $h$ Zahlen, deren absoluter Betrag $\leqq 2$ ist (wenn es solche gibt). Das Polynom $f(z) = \prod_{\nu=1}^{\tau}(z-\lambda_\nu)$ ist in einem Punkte $z_0$ des Einheitskreises absolut $\geqq 1$; bedeuten nämlich $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_\tau$ die sämtlichen $\tau+1$-ten Einheitswurzeln, so ist

$$\sum_{\nu=0}^{\tau}\varepsilon_\nu f(\varepsilon_\nu) = \tau + 1,$$

so daß $z_0$ sogar als $\tau+1$-te Einheitswurzel gewählt werden kann. Dann ist

Stop

$$(1) \qquad \prod_{\nu=1}^{\tau}(1+|\lambda_\nu|) \leqq (1+2)^\tau = 3^\tau \leqq 3^\tau \left|\prod_{\nu=1}^{\tau}(z_0-\lambda_\nu)\right|.$$

Ist $\tau < h$, so ist für $\nu = \tau+1, \ldots h$[2]) wegen $|\lambda_\nu| > 2$

$$\frac{1+|\lambda_\nu|}{|z_0-\lambda_\nu|} \leqq \frac{1+|\lambda_\nu|}{|\lambda_\nu|-|z_0|} = \frac{|\lambda_\nu|+1}{|\lambda_\nu|-1} = 1 + \frac{2}{|\lambda_\nu|-1} < 1 + \frac{2}{2-1} = 3,$$

$$\prod_{\nu=\tau+1}^{h}(1+|\lambda_\nu|) < 3^{h-\tau}\left|\prod_{\nu=\tau+1}^{h}(z_0-\lambda_\nu)\right|.$$

Wegen (1) ergibt sich

$$\prod_{\nu=1}^{h}(1+|\lambda_\nu|) \leqq 3^h\left|\prod_{\nu=1}^{h}(z_0-\lambda_\nu)\right| \leqq 3^h A\,(|z_0|^h + \ldots + 1) = 3^h(h+1)A \leqq 6^h A^{3}).$$

---

[1]) Herrn Ostrowski verdanke ich Vereinfachungen beim Beweis der Hilfsätze I und II.

[2]) bzw., wenn $\lambda_1, \ldots \lambda_h$ sämtlich absolut $> 2$ sind, für $\nu = 1, \ldots h$.

[3]) Für $h \geqq 1$ ist $1+h \leqq 1 + \binom{h}{1} + \ldots + \binom{h}{h} = (1+1)^h = 2^h$.