

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 1

ÜBER GESCHLOSSENE KURVEN
($n+1$)-TER ORDNUNG IM \mathfrak{R}^n MIT EINER
ANWENDUNG AUF EBENE KURVEN
DER KONISCHEN ORDNUNG 5 UND 6

VON

FR. FABRICIUS-BJERRE



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1942

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

EINLEITUNG

Im folgenden soll eine Klasse von geschlossenen Kurven C_{n+1} der Ordnung $n+1$ im projektiven \mathfrak{R}^n untersucht werden. Eine Kurve dieser Klasse soll keine n linear abhängigen Punkte enthalten, d. h. keine n Punkte, die in demselben \mathfrak{R}^{n-2} gelegen sind. Wir werden bekannte Resultate über ebene Kurven 3. Ordnung ohne Doppelpunkt und über Raumkurven 4. Ordnung ohne Trisekanten verallgemeinern.

In § 1 untersuchen wir die singulären Punkte, die auf einer C_{n+1} vorkommen können. Im nächsten Paragraphen betrachten wir die Zentralprojektion einer Kurve der Ordnung $n+1$ und beweisen verschiedene Hilfssätze. § 3 enthält die Untersuchung von Kurven mit Doppelpunkt oder Spitze. Wir zeigen, dass eine C_{n+1} mit Doppelpunkt 0, 1 oder 2 (verallgemeinerte) Wendepunkte enthält, während eine C_{n+1} mit Spitze immer einen Wendepunkt enthält. In § 4 zeigen wir, dass eine geschlossene C_{n+1} immer $n+1$ Wendepunkte hat, wenn sie keine singulären Punkte anderer Art enthält. Hat die Kurve singuläre Punkte, die nicht Wendepunkte sind, so besteht eine einfache Relation zwischen den Anzahlen dieser singulären Punkte (siehe S. 13). Hat die Kurve im ganzen q singuläre Punkte, so setzt sie sich aus q Bögen n -ter Ordnung (Elementarbögen) zusammen. Endlich beweisen wir, dass die gefundene Relation zwischen den Anzahlen der singulären Punkte in dem Sinne für die betrachteten Kurven charakteristisch ist, dass eine C_{n+1} keine n linear abhängigen Punkte enthält, wenn die Relation für sie erfüllt ist. § 5 enthält einen Existenzbeweis für nicht-analytische Kurven der Ordnung $n+1$ mit vorgeschriebenen singulären Punkten; diese Kurven lassen sich als Zentralprojektionen von geschlossenen Kurven der Ordnung $n+1$ im \mathfrak{R}^{n+1} gewinnen.

In § 6 wenden wir die gefundenen Resultate auf ebene Kurven von der konischen Ordnung 5 oder 6 an, d. h. auf Kurven, die von einem beliebigen Kegelschnitt in höchstens 5 oder 6 Punkten geschnitten werden. Es erweist sich hier als bequem, die Ebene auf die Veronesesche Fläche abzubilden, wobei die ebenen Kurven in Raumkurven 5. oder 6. Ordnung in einem projektiven \mathfrak{R}^5 übergehen. Es zeigt sich hier, dass ein Oval der konischen Ordnung 6 genau 6 sextaktische Punkte enthält und aus 6 Bögen der konischen Ordnung 5 besteht. Zugleich wird dargetan, dass eine geschlossene kubische Kurve der konischen Ordnung 6 ausser den 3 Wendepunkten 3 eigentliche sextaktische Punkte enthält.

§ 1. Singuläre Punkte.

Eine stetige Kurve C_{n+1} im \mathfrak{R}^n heisst von $(n + 1)$ -ter Ordnung, wenn es einen \mathfrak{R}^{n-1} (eine Hyperebene) gibt, der $n + 1$ verschiedene Punkte mit der Kurve gemeinsam hat, und wenn keine Hyperebene mehr als $n + 1$ Punkte der Kurve enthält.

Von einer solchen Kurve hat O. HAUPT (4) gezeigt, dass sie aus einer endlichen Anzahl Elementarbögen, d. h. Bögen von n -ter Ordnung besteht. In einem inneren Punkt eines solchen Bogens existiert eindeutig sowohl eine rechtsseitige als auch eine linksseitige Tangente, ebenso eine rechtsseitige und eine linksseitige Schmiegeebene usw., und höchstens in einer abzählbaren Menge von Kurvenpunkten fallen die rechtsseitigen und linksseitigen oskulierenden Mannigfaltigkeiten nicht zusammen.

Wir wollen nun über die zu betrachtenden geschlossenen Kurven C_{n+1} folgende Voraussetzung machen:

- I Die C_{n+1} soll in jedem Punkt P eine eindeutige Tangente $\tau^1(P)$, eine eindeutige Schmiegeebene $\tau^2(P)$. . . , eine eindeutige Schmiegehyperebene $\tau^{n-1}(P)$ haben.

Aus dieser Bedingung folgt, dass die beiden Halbtangenten, die beiden Schmiegehalbebenen usw. in jedem Kurvenpunkt gleich oder entgegengesetzt gerichtet sind und ferner, dass sämtliche oskulierende Mannigfaltigkeiten stetig mit P variieren (SCHERK (12) S. 305).

Jeder innere Punkt P eines Elementarbogens heie regulrer Punkt. Ein singulrer Punkt kann auftreten, wo zwei Elementarbgen zusammentreffen. Untersuchungen ber diese singulren Punkte verdankt man F. DENK (3) und P. SCHERK ((11) und (12)). Jedem Punkt ist eine Charakteristik zugeordnet, die aus n Zahlen $(a_0, a_1, \dots, a_{n-1})$ besteht, von denen jede den Wert 1 oder 2 hat. Die Zahl a_0 hat den Wert 1 oder 2, je nachdem die Halbtangenten in P entgegengesetzt oder gleich gerichtet sind. Zur Bestimmung von a_k ($k = 1, 2, \dots, n - 1$) lege man eine Hyperebene durch $\tau^{k-1}(P)$, die nicht zugleich $\tau^k(P)$ enthlt, kurz eine Hyperebene, die genau $\tau^{k-1}(P)$ enthlt, und ferner eine Hyperebene, die genau $\tau^k(P)$ enthlt. Sind diese Hyperebenen beide Sttzebenen oder beide schneidende Ebenen, so setzt man $a_k = 2$. Haben die Ebenen verschiedenen Charakter, so setzt man $a_k = 1$. (Unter $\tau^0(P)$ verstehe man den Punkt P selbst.)

In einem regulren Punkt sind die genannten Hyperebenen bekanntlich abwechselnd schneidende Ebenen und Sttzebenen; ein solcher Punkt erhlt somit die Charakteristik $(1, 1, \dots, 1)$. Fr einen singulren Punkt wird mindestens eine der Zahlen $a_k = 2$. Im \mathfrak{R}^n gibt es somit im ganzen 2^n Punkttypen, speziell in der Ebene 4, nmlich Konvexittspunkte, Wendepunkte, Spitzen 1. und 2. Art, entsprechend den Charakteristiken $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$.

Die Charakteristik bestimmt die Ordnung ν des Punktes P durch die einfache Gleichung

$$\nu = a_0 + a_1 + \dots + a_{n-1}.$$

Hieraus folgt dann, dass es auf Kurven $(n + 1)$ -ter Ordnung nur $n + 1$ Typen von Punkten, nmlich die regulren Punkte und die n Typen singulrer Punkte gibt, wo ein einziges a den Wert 2 hat. Ist $a_0 = 2$, so heit P eine Spitze, und ist $a_{n-1} = 2$, so heit P ein Wendepunkt. Ist allgemein $a_k = 2$, so heit der Punkt $(n - k)$ -fach¹⁾ singulr. Einen s -fach singulren Punkt bezeichnet man zweckmssig mit P_s . Dem Werte $s = 0$ entspricht dann ein regulrer Punkt. — Von einem regulren Punkt P_0 kann man zu einem s -fach singulren Punkt P_s kommen, indem man den einen der beiden von P_0 ausgehenden Kurven-

¹⁾ nach SCHERK.

zweige in der oskulierenden $\tau^{n-s}(P_0)$ spiegelt, während der andere fest bleibt.

Im Fall $n = 3$ treten 3 Formen singulärer Punkte auf: Wendepunkte, doppelt-singuläre Punkte und Spitzen.

Durch Projektion eines Punktes P mit gegebener Charakteristik auf eine Hyperebene ergibt sich ein Punkt P' , dessen Charakteristik leicht zu bestimmen ist (SCHERK (12), S. 292). Von SCHERKS allgemeinen Resultaten benutzen wir die folgenden:

- 1) Projiziert man einen s -fach singulären Punkt P_s ($s = 1, 2, \dots, n$) von einem ausserhalb von $\tau^{n-1}(P_s)$ gelegenen Punkt O , so erhält man einen $(s-1)$ -fach singulären Punkt P'_{s-1} . Die Projektion eines regulären Punktes ist wieder ein regulärer Punkt.

Dies folgt unmittelbar daraus, dass die Charakteristik der Projektion $(a_0, a_1, \dots, a_{n-2})$ ist.

- 2) Projiziert man einen s -fach singulären Punkt P_s ($s = 0, 1, \dots, n-1$) von dem Punkte selbst, so erhält man wieder einen s -fach singulären Punkt P'_s . Die Projektion einer Spitze vom Punkt selbst ist ein regulärer Punkt.

Dies folgt unmittelbar daraus, dass die Charakteristik der Projektion $(a_1, a_2, \dots, a_{n-1})$ ist.

- 3) Projiziert man einen regulären Punkt P_0 von einem Punkt O , der in $\tau^{n-1}(P_0)$, aber nicht in $\tau^{n-2}(P_0)$ liegt, so erhält man einen Wendepunkt.

Dies folgt aus der allgemeinen Scherkschen Regel.

Falls zwei Kurvenzweige durch denselben Punkt P gehen, heisst P ein Doppelpunkt. Ein Doppelpunkt wird zweifach gezählt, da zwei verschiedene Parameterwerte zu dem Punkt gehören. — Auf beiden Kurvenzweigen ist P ein regulärer Punkt.

§ 2. Zentralprojektionen einer C_{n+1} .

Die Projektion einer geschlossenen C_{n+1} , die die Bedingung I erfüllt, auf eine Hyperebene α^{n-1} ist wieder eine geschlossene Kurve C' , die die Bedingung I erfüllt. Liegt das Projektions-

zentrum Q auf der Kurve selbst, so ist die Ordnung von C' gleich $n-1$ oder n . Der erste Fall trifft ein, wenn Q ein Doppelpunkt oder eine Spitze ist. Dann hat nämlich eine beliebige Hyperebene durch Q ausser Q höchstens $n-1$ Punkte mit der C_{n+1} gemeinsam. Die Projektion C'_{n-1} ist also eine geschlossene Kurve $(n-1)$ -ter Ordnung im α^{n-1} , d. h. eine geschlossene Elementarkurve. Hieraus folgt sofort, dass die C_{n+1} höchstens einen Doppelpunkt oder eine Spitze enthält.

Ist das Projektionszentrum weder Spitze noch Doppelpunkt, so hat C' die Ordnung n . Legt man nämlich durch Q und $n-1$ weitere Punkte der C_{n+1} eine Hyperebene, so hat diese entweder noch einen Punkt mit der C_{n+1} gemeinsam oder sie enthält die Tangente in einem der gegebenen Punkte. Durch die Projektion erhält man eine C' , von der n Punkte oder $n-1$ Punkte und eine Tangente in der Spur \mathfrak{R}^{n-2} der Hyperebene im α^{n-1} liegen. In keinem dieser Fälle kann C' die Ordnung $n-1$ haben.

Auf der C_{n+1} gibt es definitionsgemäss $n+1$ linear abhängige Punkte, d. h. $n+1$ Punkte, die in derselben Hyperebene liegen. Bei einigen Typen von C_{n+1} gibt es schon n linear abhängige, also in einem \mathfrak{R}^{n-2} gelegene Kurvenpunkte, während andere die folgende Bedingung erfüllen:

II Es gibt keinen \mathfrak{R}^{n-2} , der mehr als $n-1$ Punkte der Kurve enthält.

Wenn die C_{n+1} die Bedingung II nicht erfüllt, wenn also ein \mathfrak{R}^{n-2} durch n Kurvenpunkte existiert, können Doppelpunkte, Trisekanten usw. auftreten, so dass für $k = 0, 1, \dots, n-1$ ein \mathfrak{R}^k mit der Kurve $k+2$ Punkte gemein haben kann. Diese Zahl ist natürlich auch maximal, da die Ordnung der Kurve $n+1$ ist. In bezug auf die Anzahlen dieser speziellen Punkte, Geraden, Ebenen . . . haben wir gesehen, dass es höchstens einen Doppelpunkt geben kann. Was die Trisekanten angeht, so kann für $n = 3$ durch jeden Kurvenpunkt höchstens eine gehen, und für $n \geq 4$ existiert überhaupt höchstens eine Trisekante. Gäbe es nämlich zwei, so hätte ein \mathfrak{R}^3 durch diese beiden Geraden 6 Punkte mit der Kurve gemein. Hierbei ist allerdings vorauszusetzen, dass die Kurve keinen Doppelpunkt hat. Setzt man voraus, dass es keine $k+1$ linear abhängigen Kurvenpunkte gibt, so kann man auf entsprechende Weise zeigen, dass es für

$k \leq \frac{n}{2} - 1$ höchstens einen \mathfrak{R}^k gibt, der $k + 2$ Kurvenpunkte enthält. Ist k grösser, so können unendlich viele \mathfrak{R}^k durch $k + 2$ Punkte auftreten. Für $k = n - 2 \geq 1$ wollen wir den folgenden Satz beweisen:

Hilfssatz 1. Falls ein \mathfrak{R}^{n-2} existiert, der n Punkte der C_{n+1} enthält, darunter den Punkt P , der kein Doppelpunkt ist, so gibt es mindestens einen \mathfrak{R}_1^{n-2} von derselben Art durch jeden Punkt P_1 der Kurve, der genügend nahe bei P liegt.

Wir werden, mit anderen Worten, beweisen, dass diese \mathfrak{R}^{n-2} nicht isoliert sind. Hat die Kurve den Doppelpunkt D , so ist jeder \mathfrak{R}^{n-2} durch D und $n - 2$ andere Kurvenpunkte von der genannten Art (es gibt übrigens auch keine anderen), und der Satz ist offensichtlich richtig. Die Kurve habe nun also keinen Doppelpunkt. Der Satz ist bekannt für $n = 3$ (SCHERK (10)). Wir nehmen daher $n \geq 4$ an und beweisen ihn durch Induktion.

Ein \mathfrak{R}^{n-2} gehe durch n Punkte der C_{n+1} , darunter den Punkt P und einen Punkt Q , von dem man, da die Kurve höchstens eine Trisekante hat ($n \geq 4$), annehmen kann, dass er nicht auf der eventuellen Trisekante durch P liegt. Von Q projiziere man die C_{n+1} auf eine Hyperebene α^{n-1} in eine C'_n , die $n - 1$ Punkte in der Spur \mathfrak{R}^{n-3} des \mathfrak{R}^{n-2} in der Hyperebene hat. Keiner dieser $n - 1$ Punkte kann in die Spur Q' der Tangente $\tau^1(Q)$ fallen. Enthielte nämlich der \mathfrak{R}^{n-2} durch n Punkte die Tangente in einem von diesen, so würde ein \mathfrak{R}^{n-1} durch den \mathfrak{R}^{n-2} und einen weiteren Punkt $n + 1$ Punkte der Kurve und die Tangente in einem von diesen Punkten enthalten, was ausgeschlossen ist. Da die Projektion P' von P nicht Doppelpunkt der C'_n ist, geht durch jeden P' genügend benachbarten Punkt P_1 nach der Induktionsvoraussetzung ein weiterer \mathfrak{R}_1^{n-3} durch $n - 1$ Punkte der C'_n , und der Punkt P'_1 kann so nahe bei P' gewählt werden, dass keiner der anderen dem \mathfrak{R}_1^{n-3} angehörenden Kurvenpunkte in Q' fällt. Diesen neuen $n - 1$ Punkten der C'_n entsprechen $n - 1$ der C_{n+1} , darunter der P'_1 entsprechende Punkt P_1 , der zu P benachbart ist. Ein \mathfrak{R}_1^{n-2} , der durch den \mathfrak{R}_1^{n-3} und Q bestimmt ist, enthält dann n Punkte

der C_{n+1} und unter diesen einen Nachbarpunkt von P . Hiermit ist der Satz bewiesen.

Wir wollen nun eine Kurve C_{n+1} betrachten, für welche die Bedingung II erfüllt ist. Für eine solche Kurve kann kein \mathfrak{R}^k für $k = 0, 1, 2, \dots, n-2$ mehr als $k+1$ Punkte der Kurve enthalten, oder anderes ausgedrückt, k Punkte können nur dann linear abhängig sein, wenn $k = n+1$. Ein \mathfrak{R}^{k-2} durch k Punkte würde nämlich mit $n-k$ weiteren Punkten einen \mathfrak{R}^{n-2} durch n Punkte der Kurve bestimmen. Diese Kurven enthalten also weder Doppelpunkte noch Trisekanten, Ebenen durch 4 Punkte usw.

Für diese Kurven soll nun der folgende Satz bewiesen werden

Hilfssatz 2. Projiziert man eine C_{n+1} , die die Bedingung II erfüllt, von einem Punkt Q der Kurve auf eine Hyperebene α^{n-1} , so erhält man eine C' , die wieder die Bedingung II erfüllt.

Falls Q eine Spitze ist, ist der Satz einleuchtend. Ist Q keine Spitze, so wird C' von n -ter Ordnung in der α^{n-1} . Wir wollen annehmen, dass die C'_n die Bedingung II nicht erfüllt, und dass ein \mathfrak{R}^{n-3} durch $n-1$ Punkte der C'_n geht. Nach Hilfssatz 1 können wir dafür sorgen, dass keiner dieser Punkte in die Spur Q' der Tangente in Q fällt. Ein \mathfrak{R}^{n-2} durch den \mathfrak{R}^{n-3} und Q hat dann im ganzen n Punkte auf der C_{n+1} gegen die Voraussetzung.

§ 3. Die C_{n+1} mit Spitze oder Doppelpunkt.

Wegen der besonderen Verhältnisse, die eintreten, wenn eine C_{n+1} von einer Spitze oder einem Doppelpunkt projiziert wird, ist es bequem, die Kurven mit Spitze oder Doppelpunkt für sich zu betrachten.

Da die Projektion der Kurve von einer Spitze oder einem Doppelpunkt eine geschlossene Kurve $(n-1)$ -ter Ordnung im α^{n-1} , also eine Elementarkurve ist, auf der alle Punkte regulär sind, folgt aus dem unter 1), § 1 zitierten Resultat von SCHERK:

Auf einer C_{n+1} mit Spitze oder Doppelpunkt können nur Wendepunkte als singuläre Punkte auftreten.

Wir behandeln zuerst die C_{n+1} mit Spitze und beweisen durch Induktion

Satz 1. Eine C_{n+1} mit Spitze enthält ausserdem genau einen Wendepunkt.

Der Satz ist richtig für $n = 2$. Die oskulierende Hyperebene in P hat, falls P ein regulärer Punkt ist, noch einen Punkt Q mit der Kurve gemeinsam. Ist P Spitze oder Wendepunkt, so hat die Hyperebene nur P mit der Kurve gemeinsam, da ein singulärer Punkt von $(n + 1)$ -ter Ordnung ist. Q variiert stetig mit P , und wenn P sich von einer Spitze oder einem Wendepunkt aus bewegt, so bewegt sich Q in entgegengesetzter Richtung, da die Bögen, die von dem singulären Punkt ausgehen, beide von n -ter Ordnung sind. — Projiziert man die Kurve von Q aus, so erhält man eine Kurve C'_n mit Spitze und daher mit genau einem Wendepunkt. Dieser Wendepunkt muss das Bild eines Punktes P auf der C_{n+1} sein, dessen oskulierende Hyperebene den Punkt Q enthält. Die Korrespondenz P, Q ist somit eindeutig, und, da P und Q sich in der Umgebung der Spitze entgegengesetzt bewegen, ist sie überall gegenläufig. P und Q fallen daher noch genau einmal zusammen, und zwar in dem Wendepunkt der C_{n+1} (vgl. JUEL (6)).

Bei Kurven mit Doppelpunkt sind die Verhältnisse nicht so einfach. Der Doppelpunkt D teilt die Kurve in zwei Teile, die sogenannten Pseudozweige. Es gilt offenbar

Satz 2. Eine C_{n+1} mit Doppelpunkt hat für gerades n einen Pseudozweig von gerader (n -ter) Ordnung und einen von ungerader ($(n + 1)$ -ter) Ordnung. Falls n ungerade ist, haben beide Pseudozweige entweder ungerade (n -te) Ordnung oder beide gerade ($(n + 1)$ -te) Ordnung.

Eine ebene C_3 mit Doppelpunkt hat einen Wendepunkt auf dem Pseudozweig ungerader Ordnung, während eine C_4 mit Doppelpunkt keinen Wendepunkt hat, wenn beide Pseudozweige von ungerader Ordnung sind, hingegen je einen Wendepunkt auf den beiden eventuellen Pseudozweigen gerader Ordnung

(JUEL (6)). — Wie der folgende Satz zeigt, bleiben diese Verhältnisse für beliebiges gerades und ungerades n erhalten:

Satz 3. Eine C_{n+1} mit Doppelpunkt hat bei geradem n einen Wendepunkt auf dem Pseudozweig ungerader Ordnung. Bei ungeradem n gibt es keinen Wendepunkt, wenn die Pseudozweige ungerade Ordnung haben, und auf jedem der beiden Pseudozweige einen, wenn diese gerade Ordnung haben.

Der Satz ist richtig für $n = 2$. Wir nehmen zunächst n ungerade an. Wenn beide Pseudozweige ungerade, also n -te Ordnung haben, sind sie Elementarbögen und daher ohne singuläre Punkte. Nun sei P ein Punkt auf einem Pseudozweig gerader, also $(n + 1)$ -ter Ordnung. Die oskulierende Hyperebene in P schneidet denselben Zweig in einem weiteren Punkt Q . Projiziert man die Kurve von Q auf eine feste Hyperebene, so erhält man eine C'_n mit Doppelpunkt, die aus einem Zweig von gerader Ordnung ohne singuläre Punkte und einem Zweig ungerader Ordnung mit Wendepunkt besteht. Der letztere Zweig ist die Projektion desjenigen Teils der C_{n+1} , auf dem Q liegt; und dem Wendepunkt in der Projektion entspricht ein Punkt P , für welchen $\tau^{n-1}(P)$ den Punkt Q enthält. Die Korrespondenz P, Q ist dann eineindeutig, und wie oben findet man genau einen Wendepunkt auf dem betreffenden Pseudozweig gerader Ordnung.

Nehmen wir nun n gerade an, so ist der Pseudozweig gerader Ordnung singularitätenfrei (als Elementarbogen). Projiziert man die Kurve von einem Punkt auf dem Zweig ungerader Ordnung aus, so geht dieser Zweig in einen Pseudozweig gerader Ordnung der C'_n über, auf welchem sich ja gerade ein Wendepunkt befindet, und wie oben ergibt sich dann auch die Existenz eines Wendepunktes auf dem Pseudozweig ungerader Ordnung der C_{n+1} .

§ 4. Die C_{n+1} , die die Bedingung II erfüllen.

Wir gehen nun zu unserer eigentlichen Aufgabe über, die geschlossenen Kurven $(n + 1)$ -ter Ordnung im \mathfrak{R}^n , die keine n

linear abhängigen Punkte enthalten, zu untersuchen. Als Hilfsmittel zur Bestimmung der Anzahlen singulärer Punkte benutzen wir wie in § 3 die Korrespondenz P, Q , wo Q als Schnittpunkt von $\tau^{n-1}(P)$ mit der Kurve C_{n+1} bestimmt ist¹⁾. Ist P ein regulärer Punkt, so ist Q von P verschieden; ist P aber singulär, also von $(n+1)$ -ter Ordnung, so fällt Q in P . Die singulären Punkte sind also die Fixpunkte der Abbildung $P \rightarrow Q$. Die Abbildung ist stetig, und in der Umgebung eines Fixpunktes variieren P und Q gegenläufig, da die Bögen, die in dem singulären Punkt zusammenstossen, beide von n -ter Ordnung (Elementarbögen) sind.

Entscheidend bei der Anwendung dieser Korrespondenz ist die Anzahl der Punkte P , die jedem Q entsprechen, mit anderen Worten die Umlaufszahl der Abbildung. In § 3 traten nur eindeutige Abbildungen auf, während dies im folgenden nur ausnahmsweise der Fall sein wird.

Wir wollen nun zunächst annehmen, dass die betrachtete Kurve C_{n+1} keine anderen singulären Punkte als Wendepunkte enthält. Für $n = 2$ gibt es 3 Wendepunkte, für $n = 3$ gibt es 4 Wendepunkte. Dies führt uns zu dem folgenden allgemeinen

Satz 4. Eine geschlossene Kurve C_{n+1} , die die Bedingungen I und II erfüllt, und die keine anderen singulären Punkte als Wendepunkte enthält, hat $n + 1$ Wendepunkte.

Der Satz ist richtig für $n = 2$. Wir nehmen ihn als richtig für eine C_n der n -ten Ordnung im \mathfrak{R}^{n-1} an. Wir projizieren die C_{n+1} von einem ihrer Punkte Q_0 auf eine Hyperebene α^{n-1} und erhalten eine C'_n , die ebenfalls I und II erfüllt, und die daher n Wendepunkte besitzt. Diese n Wendepunkte entsprechen auf der C_{n+1} gewissen n Punkten P , für welche $\tau^{n-1}(P)$ den Punkt Q_0 enthält. Falls einer der Wendepunkte der C'_n in die Spur Q'_0 der Tangente in Q_0 fällt, so bedeutet dies nur, dass Q_0 selbst Wendepunkt der C_{n+1} ist, so dass einer der n Punkte P in Q_0 liegt. Somit entsprechen jedem Punkt Q immer n verschiedene

¹⁾ Q kann wegen der Bedingung II nicht in $\tau^{n-2}(P)$ liegen, was man leicht durch Induktion einsieht.

Punkte P . Hieraus folgt, dass die Abbildung $P \rightarrow Q$ monoton ist. Wir wollen nun annehmen, dass der Punkt Q_0 und die n Punkte P_1, P_2, \dots, P_n in der genannten Reihenfolge auf der Kurve liegen. Bewegt sich P von P_1 nach P_2 , so durchläuft Q die ganze Kurve einmal, und auf dem Bogen liegt daher mindestens ein Wendepunkt. P und Q haben in der Umgebung dieses Wendepunktes und daher überall entgegengesetzte Umlaufsrichtung. Auf jedem der Bögen $P_1P_2, P_2P_3, \dots, P_nQ_0, Q_0P_1$ gibt es also genau einen Wendepunkt, im ganzen $n + 1$. (Die Umlaufszahl der Abbildung ist $-n$.)

Der obige Satz 4 ist ein Spezialfall des folgenden Satzes über Kurven mit Punkten S_1, S_2, \dots, S_n , die bzw. 1-fach, 2-fach, \dots n -fach singularär sind:

Satz 5. Für eine geschlossene C_{n+1} , die die Bedingungen I und II erfüllt, und die p_1 Punkte S_1, p_2 Punkte S_2, \dots, p_n Punkte S_n enthält, gilt die Gleichung

$$(1) \quad s_n = p_1 + 2p_2 + \dots + np_n = n + 1.$$

Falls nur p_1 von 0 verschieden ist, geht Satz 5 in Satz 4 über. Ist speziell $p_n = 1$, hat die Kurve also eine Spitze, so hat sie nach Satz 1 genau einen Wendepunkt, d. h. $p_1 = 1$, während die übrigen p_k gleich 0 sind. Satz 5 umfasst auch Satz 1 und ist also für $p_n = 1$ bewiesen.

Satz 5 ist richtig für $n = 2$, wo nur die beiden folgenden Möglichkeiten vorliegen: eine C_3 mit 3 Wendepunkten und eine C_3 mit Spitze und einem Wendepunkt. Der Satz sei nun richtig für jede C_n , die I und II befriedigt. Es genügt, eine C_{n+1} ohne Spitze zu betrachten. Wir nehmen nun an, es gäbe p_2 Punkte S_2, p_3 Punkte $S_3, \dots, p_n (= 0)$ Punkte S_n und wollen zeigen, dass man die Anzahl p_1 der Wendepunkte aus Gleichung (1) finden kann. Man projiziere die C_{n+1} von einem regulären Punkt Q aus auf eine Hyperebene α^{n-1} in eine Kurve C'_n , die I und II erfüllt. Nach § 1 erhält die C'_n hierbei p_2 Wendepunkte S'_1, p_3 Punkte S'_2, \dots, p_n Punkte S'_{n-1} . Ausserdem entsteht eine Anzahl Wendepunkte, die von Punkten auf der C_{n+1} herrühren, deren oskulierende Hyperebenen Q enthalten. Diese Anzahl sei p'_1 .

Zwischen diesen Anzahlen besteht nun nach Induktionsvoraussetzung die Relation

$$(2) \quad s_{n-1} = (p'_1 + p_2) + 2 p_3 + \dots + (n-1) p_n = n.$$

Die Anzahl p'_1 ändert sich nicht, wenn Q in einen singulären Punkt der C_{n+1} , z. B. in S_k fällt. Die Anzahl p_{k+1} der Punkte S'_k wird um 1 erhöht, während die Anzahl p_k der Punkte S'_{k-1} zugleich um 1 vermindert wird, wodurch die Anzahl der Punkte P auf der C_{n+1} , deren oskulierende Hyperebene Q enthält, scheinbar um 1 vermindert wird. Da aber S_k gleichzeitig ein Fixpunkt der Abbildung $P \rightarrow Q$ ist, muss auch dieser Punkt zu den P -Punkten gerechnet werden, d. h. die Anzahl bleibt doch erhalten.

Die Kurve C_{n+1} enthält singuläre Punkte; gibt es nämlich keine Punkte S_k , $k \geq 2$, so gibt es Satz 4 zufolge $n+1$ Wendepunkte S_1 . Die Punkte P und Q haben dann überall entgegengesetzte Umlaufsrichtung, und es gibt infolgedessen $p'_1 + 1$ Fixpunkte in der Korrespondenz, also $p'_1 + 1$ singuläre Punkte auf der C_{n+1} . Von diesen singulären Punkten sind $p_2 + p_3 + \dots + p_n$ nicht Wendepunkte. Die Anzahl p_1 der Wendepunkte wird dann

$$(3) \quad p_1 = (p'_1 + 1) - (p_2 + p_3 + \dots + p_n).$$

Setzt man hierin den aus Gleichung (2) gefundenen Wert von p'_1 ein und ordnet die Gleichung um, so erhält man gerade Gleichung (1). — Wir werden später eine geometrische Deutung der Relation (1) angeben.

Durch diesen Satz haben wir Klarheit über die mögliche Verteilung der singulären Punkte auf einer C_{n+1} ohne n linear abhängige Punkte gewonnen. Zwischen zwei auf einander folgenden singulären Punkten ist die Kurve singularitätenfrei und folglich monoton. Wir wollen zeigen, dass ein solcher Bogen nicht nur monoton, sondern sogar ein Elementarbogen (von n -ter Ordnung) ist.

Die C_{n+1} habe $q = p_1 + p_2 + \dots + p_n$ singuläre Punkte, die in der durch die Anordnung auf der Kurve gegebenen Reihe-

folge mit $S', S'', \dots, S^{(q)}$ bezeichnet seien. Wenn P sich auf dem monotonen Bogen $S'S''$ bewegt, durchläuft der entsprechende Punkt Q den komplementären Bogen von S' nach S'' und analog, wenn sich P von S'' nach S''' bewegt, usw. Hieraus folgt, dass die oskulierende Hyperebene in einem Punkt P , der auf einem der monotonen Bögen zwischen zwei auf einander folgenden singulären Punkten gelegen ist, diesen Bogen nicht wieder schneidet. Mit anderen Worten: Zwischen zwei korrespondierenden Punkten P und Q gibt es immer mindestens einen singulären Punkt. Dieser Satz kann als Spezialfall des folgenden Satzes aufgefasst werden:

Satz 6. Enthält eine Hyperebene $n + 1$ Punkte P_1, \dots, P_{n+1} der Kurve C_{n+1} und liegen diese in der angegebenen Reihenfolge auf der Kurve, so gibt es auf dem Bogen $P_1P_2 \dots P_{n+1}$ stets mindestens einen singulären Punkt.

Der Satz ist richtig für $n = 2$. Die Kurve C_3 ist nämlich aus höchstens 3 konvexen Bögen zusammengesetzt. Wir nehmen an, der Satz sei richtig für jede C_n und projizieren die C_{n+1} vom Punkte P_{n+1} aus auf eine Hyperebene in eine Kurve C'_n , die von n -ter Ordnung ist, da P_{n+1} keine Spitze sein kann. Durch Projektion der n Punkte P_1, P_2, \dots, P_n entstehen n Punkte P'_1, P'_2, \dots, P'_n auf der C'_n , die auf dieser Kurve in der angegebenen Reihenfolge liegen. Nach Induktionsvoraussetzung gibt es auf dem Bogen $P'_1P'_2 \dots P'_n$ mindestens einen singulären Punkt S' . Wenn dieser mindestens zweifach singulär ist, so ist S selbst ein singulärer Punkt auf dem Bogen $P_1P_2 \dots P_n$, also umso mehr auf dem Bogen $P_1P_2 \dots P_{n+1}$. Ist S' dagegen Wendepunkt, so enthält $\tau^{n-1}(S)$ den Punkt P_{n+1} , und dann gibt es auf dem Bogen SP , der in dem Bogen $P_1P_2 \dots P_{n+1}$ enthalten ist, einen singulären Punkt.

Da eine Hyperebene also nicht $n + 1$ Schnittpunkte mit der C_{n+1} haben kann, die auf einem und demselben Bogen zwischen zwei singulären Punkten liegen, muss jeder derartige monotone Bogen von n -ter Ordnung sein. Hieraus folgt

Satz 7. Eine C_{n+1} mit q singulären Punkten ist aus q Elementarbögen zusammengesetzt.

Speziell sieht man, dass eine C_{n+1} mit $n + 1$ Wendepunkten aus $n + 1$ Elementarbögen besteht. Für $n = 3$ ist dies ohne Beweis von COURTAND (2) behauptet worden.

Für $n = 3$ heben wir hervor

Satz 8. Raumkurven 4. Ordnung ohne Trisekanten können in 4 Typen eingeteilt werden. Es können 4 Wendepunkte, 2 Wendepunkte und ein doppelt-singulärer Punkt, 2 doppelt-singuläre Punkte oder ein Wendepunkt und eine Spitze auftreten. Die Kurve besteht aus 4, 3, 2, bzw. 2 Elementarbögen.

Besonders interessant ist der dritte Typus. Die Kurve ist hier aus 2 Bögen 3. Ordnung zusammengesetzt, die im selben Sinne gewunden sind. Die Kurve ist also, abgesehen von den beiden singulären Punkten, monoton. — Eine algebraische Kurve dieser Art erhält man nicht als Schnitt zweier Flächen zweiten Grades, sondern als Schnitt einer kubischen Fläche mit einer Fläche zweiten Grades. Legt man durch zwei Erzeugende derselben Schar einer Fläche zweiten Grades eine kubische Fläche, für welche zwei andere Erzeugende derselben Schar Wendetangenten sind, so ist die Schnittkurve der Flächen eine solche C_4 . — Man kann sie auch durch Projektion einer algebraischen C_4 im \mathbb{R}^4 von einem Schnittpunkt zweier Schmiegeebenen erhalten (siehe § 5).

Wir haben oben gezeigt, dass eine C_{n+1} , die den Bedingungen I und II genügt, singuläre Punkte verschiedener Art hat, deren Anzahlen die Gleichung (1) befriedigen. Umgekehrt wollen wir nun zeigen, dass diese Relation (1) für Kurven $(n + 1)$ -ter Ordnung, die keine n linear abhängigen Punkte enthalten, charakteristisch ist. Wir haben zu beweisen:

Satz 9. Wenn eine C_{n+1} der Bedingung I genügt und singuläre Punkte hat, die nach Anzahl und Art die Gleichung

$$(1) \quad s_n = p_1 + 2p_2 + \dots + np_n = n + 1$$

befriedigen, so genügt die C_{n+1} auch der Bedingung II.

Wir beweisen den Satz unter der scheinbar schwächeren Voraussetzung $s_n \geq n + 1$, wobei jedoch zu bemerken ist, dass das Ungleichheitszeichen von selbst fortfällt, wenn der Satz unter dieser Voraussetzung gilt. Denn dann erfüllt die C_{n+1} ja die Bedingungen I und II.

Der Satz ist richtig für $n = 2$, $s_2 \geq 3$, da eine Kurve mit Spitze oder mit 3 Wendepunkten doppeltpunktfrei ist, und für eine Kurve mit Doppelpunkt ist $s_2 = 1$.

Nehmen wir nun an, die Bedingung II sei für jede C_n , für welche $s_{n-1} \geq n$ ist, erfüllt ($n \geq 3$).

Die C_{n+1} muss doppeltpunktfrei sein. Sonst wäre nämlich $s_n = 0, 1$ oder 2 , während s_n mindestens 3 sein sollte. — Die C_{n+1} habe nun $q = p_1 + p_2 + \dots + p_n$ singuläre Punkte. Die Korrespondenz P, Q hat dann q Fixpunkte, und da P und Q von jedem singulären Punkt aus gegenläufig variieren, muss Q die Kurve $(q - 1)$ -mal durchlaufen, wenn P die Kurve einmal durchläuft. Durch jeden Punkt Q gehen folglich $r \geq q - 1$ oskulierende Hyperebenen der Kurve. Wir projizieren nun die C_{n+1} von dem regulären Punkt Q aus auf eine feste Hyperebene α^{n-1} , wodurch eine C'_n entsteht, die der Bedingung I genügt. Auf dieser Kurve gibt es $r + p_2$ Wendepunkte, p_3 singuläre Punkte S'_2, \dots, p_n singuläre Punkte S'_{n-1} . Für die Grösse s_{n-1} ergibt sich

$$s_{n-1} = (r + p_2) + 2p_3 + \dots + (n - 1) p_n.$$

Da nun

$$r \geq q - 1 = p_1 + p_2 + \dots + p_n - 1$$

ist, wird

$$s_{n-1} \geq p_1 + 2p_2 + \dots + np_n - 1$$

oder

$$s_{n-1} \geq s_n - 1 \geq n.$$

Mit anderen Worten: Die durch Projektion entstandene Kurve C'_n befriedigt dieselbe Bedingung $s_k \geq k + 1$ wie die ursprüngliche Kurve. Hieraus folgt z. B., dass die C'_n keinen Doppelpunkt, also die C_{n+1} keine Trisekante besitzt. Es folgt aber überhaupt, dass die Bedingung II für die C_{n+1} erfüllt ist. Man nehme nämlich an, es gäbe einen \mathfrak{R}^{n-2} durch n Punkte der gegebenen Kurve. Einer dieser Punkte kann als regulär vorausgesetzt werden, da man gegebenenfalls zu einem benachbarten

\mathfrak{R}^{n-2} durch n Punkte, von denen wenigstens einer regulär ist, übergehen könnte. Durch Projektion der C_{n+1} von diesem Punkt aus erhielte man eine C'_n , für welche $s_{n-1} \geq n$ ist, und die $n-1$ Punkte mit einem \mathfrak{R}^{n-3} gemein hat, was der Induktionsvoraussetzung widerspräche.

Hiermit ist der Satz bewiesen, und zugleich ergibt sich:

Satz 10. Wenn eine C_{n+1} die Bedingung I, aber nicht die Bedingung II erfüllt, so besteht zwischen den Anzahlen der singulären Punkte die Ungleichung

$$s_n = p_1 + 2p_2 + \dots + np_n \leq n - 1.$$

Dass die rechte Seite sich von $n+1$ auf $n-1$ erniedrigt, ist eine Folge davon, dass die Anzahl der Wendepunkte immer dieselbe Parität haben muss.

Die obige Ungleichung hat sich für Kurven mit Doppelpunkt als richtig erwiesen, und sie gilt bekanntlich auch für Raumkurven 4. Ordnung mit Trisekanten. Da notwendigerweise $p_n = 0$ ist, kann die Ungleichung übrigens zu

$$s_n = p_1 + 2p_2 + \dots + (n-1)p_{n-1} \leq n - 1$$

vereinfacht werden.

§ 5. Die Existenz der Kurven C_{n+1} .

Wir wollen nun zeigen, dass sämtliche verschiedenen Typen von Kurven $(n+1)$ -ter Ordnung, die wir in § 4 betrachtet haben, und für die die Anzahlen der Singularitäten Gleichung (1) befriedigen, durch Projektion einer und derselben Kurve erzeugt werden können, nämlich einer Kurve $(n+1)$ -ter Ordnung im \mathfrak{R}^{n+1} .

Es sei \bar{C}_{n+1} eine solche geschlossene Elementarkurve im \mathfrak{R}^{n+1} . Wir wählen ein Projektionszentrum O ausserhalb der Kurve und projizieren auf eine Hyperebene (einen \mathfrak{R}^n), wodurch eine C_{n+1} in diesem \mathfrak{R}^n entsteht. Falls O in einem oskulierenden $\tau^{n+1-k}(P)$ der Kurve \bar{C}_{n+1} , aber nicht in einem oskulierenden $\tau^{n-k}(P)$ liegt, ist die Projektion des Punktes P nach den allgemeinen Regeln ein k -fach singulärer Punkt auf der C_{n+1} .

Projizieren wir nun vom Schnittpunkt von $n + 1$ oskulierenden Hyperebenen der \bar{C}_{n+1} aus, so entsteht eine C_{n+1} mit $n + 1$ Wendepunkten, und diese Anzahl ist — wie in § 4 gezeigt wurde — gerade kennzeichnend für einen der Kurventypen, die die Bedingung II erfüllen. Allgemeiner können wir die \bar{C}_{n+1} vom Schnittpunkt von p_1 oskulierenden τ^n , p_2 oskulierenden τ^{n-1} , . . . , p_n oskulierenden τ^1 , d. h. Tangenten, aus projizieren, wo die Grössen p_ν der Relation

$$p_1 + 2p_2 + \dots + np_n = n + 1$$

genügen müssen. Wir erhalten dann eine C_{n+1} mit p_1 Wendepunkten, p_2 doppelt-singulären Punkten usw. Es gilt also

Satz 11. Projiziert man eine geschlossene Kurve $(n + 1)$ -ter Ordnung im \mathfrak{R}^{n+1} von einem Punkt O aus, durch den p_1 oskulierende τ^n der Kurve, p_2 oskulierende τ^{n-1} , . . . , p_n Tangenten gehen, so entsteht eine C_{n+1} im \mathfrak{R}_n mit p_1 Wendepunkten, p_2 doppelt-singulären Punkten, . . . , p_n Spitzen, wobei zwischen den Anzahlen p_ν die Relation

$$p_1 + 2p_2 + \dots + np_n = n + 1$$

besteht.

Bei dieser Projektion entstehen also alle wesentlichen Typen. Berücksichtigt man jedoch auch die Klasse der Kurve, so sind nicht alle möglichen Typen der C_{n+1} repräsentiert. So hat eine ebene C_3 , die durch Projektion einer Raumkurve 3. Ordnung erzeugt wird, die Klasse 4, während auch ebene C_3 der Klasse 6 existieren.

Durch diese Projektionsmethode ist die Existenz von Kurven der betrachteten Art, die nicht algebraisch, nicht einmal analytisch zu sein brauchen, gesichert. Eine allgemeine Bestimmung der monotonen Bögen in Räumen von beliebig vielen Dimensionen hat J. HJELMSLEV (5) durchgeführt, und man kann, wie FRL. SAUTER (9) gezeigt hat, diese Bögen ordnungsfest schliessen, wodurch nicht-analytische geschlossene Kurven $(n + 1)$ -ter Ordnung im \mathfrak{R}^{n+1} entstehen. Auch die in § 3 betrachteten Kurven $(n + 1)$ -ter Ordnung

mit Doppelpunkt können durch Projektion einer \bar{C}_{n+1} im \mathfrak{R}^{n+1} erhalten werden. Ist n gerade, sind die Kurven also selbst von ungerader Ordnung, so hat man nur das Projektionszentrum auf einer Doppelsekante der \bar{C}_{n+1} zu wählen. — Bei ungeradem n , also Kurven gerader Ordnung, lagen zwei Möglichkeiten vor. Man kann in diesem Fall erreichen, dass die Kurve \bar{C}_{n+1} des \mathfrak{R}^{n+1} ganz im Endlichen liegt. Wir konstruieren die konvexe Hülle der Kurve und betrachten eine Doppelsekante PQ der Kurve. Die Strecke PQ liegt ganz innerhalb der konvexen Hülle, während die Verlängerungen ausserhalb liegen. Jede Hyper-ebene durch einen Punkt auf der Strecke PQ schneidet die Kurve in mindestens 2 Punkten, während durch die Punkte der Verlängerungen der Strecke PQ Hyperebenen gelegt werden können, die keinen Punkt mit der Kurve gemeinsam haben. Im ersten Fall wird der Index der Projektion 2, im zweiten Fall 0. Projiziert man daher von einem Punkt auf der Strecke PQ , so erhält man eine C_{n+1} mit Doppelpunkt und ungeraden Pseudozweigen, während die Projektion von einem Punkt ausserhalb der Strecke Pseudozweige von gerader Ordnung ergibt.

§ 6. Die Veronesesche Fläche. Ebene Kurven der konischen Ordnung 5 oder 6.

Wir wollen nun einige der vorstehenden Resultate auf die Untersuchung von ebenen Kurven der konischen Ordnung 5 oder 6, d. h. Kurven, die von beliebigen Kegelschnitten in höchstens 5 bzw. 6 Punkten geschnitten werden, anwenden. Die ebenen Kurven werden auf Kurven abgebildet, die auf der Veroneseschen Fläche liegen, über die wir einige Bemerkungen vorausschicken.¹⁾

Wir bezeichnen mit (y_1, y_2, y_3) projektive Punktkoordinaten in der Ebene II und mit $(x_1, x_2, x_3, x_4, x_5, x_6)$ projektive Punktkoordinaten in einem 5-dimensionalen Raum \mathfrak{R}^5 . Die Veronesesche Fläche wird in Parameterform durch

$$x_1 : x_2 : x_3 : x_4 : x_5 : x_6 = y_1^2 : y_2^2 : y_3^2 : y_2 y_3 : y_3 y_1 : y_1 y_2$$

dargestellt. Wenn der Punkt Y die projektive Ebene II durchläuft, so durchläuft der Punkt X die Veronesesche Fläche V .

¹⁾ Ähnliche Abbildungen sind von J. HJELMSLEV herangezogen worden (5).

Von den wohlbekanntenen Eigenschaften der Fläche heben wir einige hervor. Einem Kegelschnitt in der Ebene Π entspricht offenbar die Schnittkurve von V mit einer Hyperebene im \mathfrak{R}^5 . Zwei Kegelschnitte in der Ebene haben im allgemeinen 4 Punkte gemeinsam, woraus folgt, dass die Veronesesche Fläche eine algebraische Fläche 4. Grades ist, so dass die Flächenkurve, die einem ebenen Kegelschnitt entspricht, eine C_4 in einem \mathfrak{R}^4 , also eine algebraische Elementarkurve (Normalkurve) im \mathfrak{R}^4 ist. Falls der \mathfrak{R}^4 die Fläche V berührt, hat die C_4 einen Doppelpunkt und zerfällt in zwei Kegelschnitte, die somit Bilder von zwei Geraden in Π sind, deren Schnittpunkt dem Berührungspunkt auf V entspricht. Einer Geraden in Π entspricht also ein Kegelschnitt auf V . Die Veronesesche Fläche enthält also ∞^2 Kegelschnitte, so dass zwei Punkte einen und nur einen Kegelschnitt bestimmen.

Während ein \mathfrak{R}^3 die Fläche im allgemeinen in höchstens 4 reellen Punkten schneidet, kann es vorkommen, dass ein solcher Raum einen ganzen Kegelschnitt mit V gemeinsam hat, nämlich dann, wenn die Hyperebenen, die den \mathfrak{R}^3 bestimmen, diesen Kegelschnitt enthalten. Dem entspricht in der Ebene Π , dass man zwei Kegelschnitte hat, die in zwei Geradenpaare mit einer gemeinsamen Geraden zerfallen. Die beiden anderen Geraden haben noch einen Schnittpunkt, der jedoch ebenfalls auf die gemeinsame Gerade fallen kann. Hieraus folgt, dass ein \mathfrak{R}^3 durch einen Kegelschnitt höchstens noch einen Punkt mit V gemeinsam hat.

Als dem obigen geht unmittelbar hervor, dass einer ebenen (geschlossenen) Kurve der konischen Ordnung p auf der Veroneseschen Fläche eine (geschlossene) Kurve p -ter (linearer) Ordnung entspricht.

Nehmen wir zunächst $p = 5$ an. Die ebene Kurve Γ muss dann ein konvexer Bogen sein. Gäbe es nämlich Geraden, die Γ in 3 oder mehr Punkten schneiden, so würden zwei dieser Geraden einen ausgearteten Kegelschnitt mit 6 Punkten auf Γ bestimmen. — Die entsprechende Kurve Γ' im \mathfrak{R}^5 ist von 5. Ordnung, also ein Elementarbogen. Einem Schmiegekegelschnitt von Γ entspricht auf der Veroneseschen Fläche eine Kurve C_4 , die mit Γ gemeinsame oskulierende τ^4 hat. Aus bekannten Eigenschaften der Elementarbögen erhält man somit für Γ :

Ein (konvexer) Bogen der konischen Ordnung 5 hat in jedem Punkt sowohl einen vorderen als auch einen hinteren Schmiegekegelschnitt, die in einer höchstens abzählbaren Punktmenge von einander verschieden sind. Fallen die Kegelschnitte in jedem Punkt zusammen, so variieren sie stetig mit dem Punkt.

Ein Bogen der konischen Ordnung 5 heiße selbst ein (koni-scher) Elementarbogen. Wenn die Schmiegekegelschnitte stetig variieren, so ist Γ die einhüllende Kurve dieses Kegelschnitt-systems. Man sieht hieraus, dass die Schmiegekegelschnitte eines konischen Elementarbogens in derselben Weise in einander liegen wie die Schmiegekreise eines ebenen konvexen Bogens der zyklischen Ordnung 3 (vgl. BLASCHKE (1), S. 34, Aufgabe 4).

Wir betrachten nun Kurven der konischen Ordnung 6, die also in Kurven 6. Ordnung auf V abgebildet werden. Auf diese Kurven können frühere Ergebnisse mit $n = 5$ angewendet werden. Aus dem Seite 4 genannten Satz von Haupt ergibt sich, dass eine Kurve der konischen Ordnung 6 aus einer endlichen Anzahl konischer Elementarbögen besteht. Dieses Ergebnis folgt jedoch auch sofort aus Haupts allgemeiner Ordnungstheorie.

Eine Γ_6 kann höchstens die lineare Ordnung 3 haben, andererseits können Ovale und kubische Kurven wirklich vor-kommen. Wir nehmen im folgenden an, dass die Γ_6 geschlossen ist und die folgende Bedingung erfüllt:

III. Wenn ein Kegelschnitt durch 5 Punkte von Γ in der Weise variiert, dass 2 oder mehrere der Punkte gegen denselben Grenzpunkt konvergieren, so soll der Kegelschnitt einer bestimmten Grenzlage zustreben.

Hieraus folgt dann unmittelbar, dass das Bild Γ' auf der Veroneseschen Fläche in jedem Punkt eine eindeutige Tangente, eine eindeutige Schmiegeebene, eindeutige τ^3 und τ^4 hat, so dass Bedingung I erfüllt ist.

Ferner ist für Γ' die Bedingung II erfüllt. Ein \mathfrak{R}^3 hat nämlich im allgemeinen höchstens 4 Punkte mit der Veroneseschen

Fläche und also umso mehr höchstens 4 Punkte mit I' gemein. Enthält der \mathfrak{R}^3 einen Kegelschnitt der Fläche, so hat er höchstens noch einen Punkt mit V gemein. Einem Kegelschnitt auf V entspricht eine Gerade in der Ebene Π , die I in höchstens 3 Punkten schneidet. Die totale Anzahl von Schnittpunkten des \mathfrak{R}^3 mit I' ist also höchstens 4.

Endlich enthält die Kurve I' keine anderen singulären Punkte als Wendepunkte. Soll nämlich eine Hyperebene durch 5 Kurvenpunkte einer bestimmten Grenzlage zustreben (Bedingung III), wenn die 5 Punkte gegen denselben Kurvenpunkt konvergieren, so können nach einem Satz von SCHERK ((12), Satz 4, 5) keine anderen Singularitäten als Wendepunkte auftreten.

Einem Wendepunkt auf I' entspricht offenbar ein sextaktischer Punkt auf der Kurve I in der Ebene, während einem regulären Punkt auf I' ein regulärer Punkt auf I entspricht. Wenden wir Satz 4, S. 12 auf Ovale an, so erhalten wir

Satz 12. Ein Oval der konischen Ordnung 6, für welches die Bedingung III erfüllt ist, enthält 6 sextaktische Punkte und besteht aus 6 Bögen der konischen Ordnung 5.

Mit Hilfe von Satz 12 können wir einige weitere Schlüsse ziehen. MUKHOPADHYAYA (8) hat gezeigt, dass ein Oval, das von einem Kegelschnitt in $2n$ Punkten geschnitten wird, mindestens $2n$ sextaktische Punkte enthält. (Hierin ist der Satz enthalten, dass jedes Oval mindestens 6 sextaktische Punkte besitzt; dieser Satz wurde von MUKHOPADHYAYA 1909 bewiesen.) Wir schliessen hieraus

Satz 13. Hat ein Oval 8 oder mehr sextaktische Punkte, so gibt es immer einen Kegelschnitt, der mindestens 8 Punkte mit dem Oval gemeinsam hat.

Wenn ein Oval einen Mittelpunkt hat, gibt es nach BLASCHKE ((1), S. 46) mindestens 8 sextaktische Punkte. Hieraus folgt

Satz 14. Die konische Ordnung eines Ovals mit Mittelpunkt ist mindestens 8.

Was die Existenz von Ovalen der konischen Ordnung 6 angeht, so kann z. B. auf die Ovale hingewiesen werden, die bei den algebraischen Kurven 3. Ordnung vorkommen. Ein solches Oval hat die konische Ordnung 6 und besitzt daher genau 6 sextaktische Punkte. Ein Oval mit genau 6 sextaktischen Punkten ist nach MUKHOPADHYAYA von der konischen Ordnung 6. Eine solche Kurve hat BLASCHKE konstruiert (l. c. S. 45).

Eine kubische Kurve der konischen Ordnung 6 hat weder Doppelpunkt noch Spitze, also 3 Wendepunkte. Diese sind auch zu den sextaktischen Punkten zu rechnen, da eine doppelt gezählte Wendetangente mit T im Wendepunkt 6 Punkte gemeinsam hat. Hieraus folgt

Satz 15. Eine geschlossene ebene Kurve der linearen Ordnung 3 und der konischen Ordnung 6, die die Bedingung III erfüllt, hat ausser den Wendepunkten 3 sextaktische Punkte und besteht aus 6 konischen Elementarbögen.

Dies stimmt überein mit der Anzahl der sextaktischen Punkte auf den algebraischen Kurven 3. Ordnung. Diese haben nämlich 27 reelle oder imaginäre sextaktische Punkte, nämlich die Berührungspunkte der Tangenten, die man von den 9 Wendepunkten an die Kurve legen kann. Da 6 Wendepunkte imaginär sind, bleiben 9 möglicherweise reelle sextaktische Punkte; 6 von diesen liegen auf dem eventuellen reellen Oval oder sind imaginär. Es gibt also genau 3 eigentliche sextaktische Punkte auf dem Zweig 3. Ordnung.

Literaturverzeichnis.

- (1) BLASCHKE, W. Vorlesungen über Differentialgeometrie II (Berlin 1923).
- (2) COURTAND, M. Sur les courbes gauches du quatrième ordre (C. R. de l'Ac. des Sc. B. 204, 1937 S. 215).
- (3) DENK, F. Über elementare Punkte höherer Ordnung auf Kurven im \mathfrak{R}^n . (Sitzungsberichte d. phys.-med. Soz. zu Erlangen B. 67, 1935, S. 1).
- (4) HAUPT, O. Ein Satz über die reellen Raumkurven vierter Ordnung und seine Verallgemeinerung (Math. Ann. B. 108, 1933, S. 126).
- (5) HJELMSLEV, J. Introduction à la théorie des suites monotones (D. Kgl. Danske Vidensk. Selskab, Oversigt 1911. S. 343).
- (6) JUEL, C. Om ikke-analytiske Kurver (D. Kgl. Danske Vidensk. Selskab, Skrifter 7. I. 6. 1906. S. 297).
- (7) — Einleitung in die Theorie der ebenen Elementarkurven dritter und vierter Ordnung (D. Kgl. Danske Vidensk. Selskab, Skrifter 7. XI. 2. 1914. S. 113).
- (8) MUKHOPADHYAYA, S. Extended Minimum Number Theorems of Cyclic and Sextactic Points on a Plane Convex Oval (Math. Zeitschr. B. 33, 1931, S. 648).
- (9) SAUTER, I. Zur Theorie der Bogen n -ter (Realitäts)-Ordnung im projektiven \mathfrak{R}^n (Math. Zeitschr. B. 41, 1936, S. 507).
- (10) SCHERK, P. Über reelle geschlossene Raumkurven vierter Ordnung (Math. Ann. B. 112, 1936, S. 743).
- (11) — Über differenzierbare Kurven und Bögen (Journ. Tschechoslov. de Math. et de Phys. B. 66, 1937, S. 165).
- (12) — Über Punkte $(n + 1)$ -ter Ordnung auf Bögen im \mathfrak{R}^n . (Annali di Mat. Ser. IV, B. 17, 1938, S. 289).

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 2

ÜBER DAS
KOEFFIZIENTENDARSTELLUNGS-
PROBLEM DIRICHLETSCHER
REIHEN

VON

HARALD BOHR



KØBENHAVN

I KOMMISSION HOS EINAR MUNKSGAARD

1942

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

Einleitung.

In der Theorie der Dirichletschen Reihen $f(s) = f(\sigma + it)$ $= \sum \frac{a_n}{n^s}$, besonders ihre Anwendung auf Probleme der analytischen Zahlentheorie betreffend, spielt bekanntlich die folgende Formel, welche die Summe der Koeffizienten durch die Funktion ausdrückt, eine wichtige Rolle

$$\frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \frac{w^s}{s} f(s) ds = \sum_{n < w} a_n.$$

Hierbei ist die Gerade $\sigma = \alpha > 0$ im Inneren des Konvergenzbereiches gelegen, und w ist eine beliebige positive Grösse. Sollte diese gerade eine ganze Zahl $w = m$ sein, soll jedoch das entsprechende Reihenglied $\frac{a_m}{m^s}$ fehlen, d. h. $a_m = 0$ sein. (Sonst bedarf die Formel nur einer geringfügigen Modifikation.)

Formal folgt diese Darstellungsformel natürlich aus der bekannten Formel

$$\frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \frac{w^s}{s} ds = \begin{cases} 1, & w > 1 \\ 0, & 0 < w < 1. \end{cases}$$

Die Richtigkeit dieser klassischen Darstellungsformel wurde zuerst von PERRON in einer bedeutsamen Arbeit »Zur Theorie der Dirichletschen Reihen«, Crelles Journal 1908, dargetan. Für verschiedene Zwecke ist es bekanntlich bequem, die allgemeinere Darstellungsformel

$$\frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \frac{w^s}{s^q} f(s) ds = \frac{1}{\Gamma(q)} \sum_{n < w} a_n \left(\log \frac{w}{n} \right)^{q-1}$$

zur Verfügung zu haben, wo die positive Zahl ϱ nicht gerade gleich 1 ist. Diese Formel folgt ihrerseits formal aus der entsprechenden Relation

$$\frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} \frac{w^s}{s^\varrho} ds = \begin{cases} \frac{1}{\Gamma(\varrho)} (\log w)^{\varrho-1}, & w > 1 \\ 0, & 0 < w < 1. \end{cases}$$

Über die Gültigkeit dieser allgemeineren Formel bezüglich der Werte von ϱ besteht aber heute noch ein ungelöstes Problem, das auf PERRON zurückgeht. In seiner oben zitierten Arbeit beweist er, dass die Formel für jedes $\varrho \geq 1$ gültig ist. Dagegen betont er, dass seine Beweismethode für $0 < \varrho < 1$ versagt, und dass es daher unentschieden bleibt, ob die Formel für solche Werte von ϱ überhaupt gültig bleibt, d. h. ob auch für solche ϱ die gliedweise Integration der auftretenden unendlichen Reihe gestattet ist.

In der vorliegenden Arbeit, welche mit älteren Untersuchungen des Verfassers in seiner Habilitationsschrift »Bidrag til de Dirichlet'ske Rækkers Teori«, Kopenhagen 1910, eng zusammenhängt, wird die vollständige Lösung des genannten Problems gegeben und zwar wird gezeigt, dass die betreffende Formel für keinen einzigen Wert von ϱ im Intervall $0 < \varrho < 1$ allgemein gültig ist. Dies geschieht dadurch, dass zu jedem $0 < \varrho < 1$ eine Dirichletsche Reihe $f(s) = \sum \frac{a_n}{n^s}$ konstruiert wird, für welche die obige Formel nicht gilt. Übrigens wird die Konstruktion eines einzigen Gegenbeispiels geliefert, das für sämtliche ϱ des Intervalls $0 < \varrho < 1$ verwendbar ist. Hierbei wird es bequem sein, die Nicht-Gültigkeit gerade für $w = 1$ (und also $\alpha_1 = 0$) darzutun, wo die Formel einfach

$$\int_{\alpha-i\infty}^{\alpha+i\infty} \frac{f(s)}{s^\varrho} ds = 0$$

heisst.

Vorbereitungen und Angabe des zu beweisenden Satzes.

Im folgenden werden nur Dirichletsche Reihen betrachtet, deren Konvergenzabszisse gleich 0 ist, die also in der ganzen

Halbebene $\sigma > 0$ konvergieren. Wir werden eine Dirichletsche Reihe (mit $a_1 = 0$) konstruieren, für welche es zu jedem $0 < \rho < 1$ eine positive Abszisse $\alpha = \alpha(\rho)$ gibt, für welche

$$\int_{\alpha - i\infty}^{\alpha + i\infty} \frac{f(s)}{s^\rho} ds$$

übrigens sogar

$$\int_{\alpha}^{\alpha + i\infty} \frac{f(s)}{s^\rho} ds$$

divergiert. Damit wird natürlich die Nicht-Gültigkeit der obigen Formel bewiesen. Der Nachweis der Divergenz des letztgenannten Integrals wird dadurch erbracht, dass bei festem $0 < \rho < 1$ zwei Abszissen σ_1 und σ_2 mit $0 < \sigma_1 < \sigma_2 < 1$ angegeben werden,

so dass das Integral $\int_{\sigma_1 + iT}^{\sigma_2 + iT} \frac{f(s)}{s^\rho} ds$ für gewisse ins Unendliche wach-

sende Werte von T dem Betrage nach über alle Grenzen wächst. Dies widerspricht nämlich nach dem Cauchyschen Integralsatz, auf das Rechteck $\sigma_1, \sigma_2, \sigma_2 + iT, \sigma_1 + iT$ angewendet, der Mög-

lichkeit, dass sowohl das Integral $\int_{\sigma_1}^{\sigma_1 + i\infty} \frac{f(s)}{s^\rho} ds$ als auch das Integral $\int_{\sigma_2}^{\sigma_2 + i\infty} \frac{f(s)}{s^\rho} ds$ beide existieren können, so dass mindestens eine

der beiden Abszissen σ_1 und σ_2 als unser $\alpha = \alpha(\rho)$ verwendet werden kann.

Bei einem solchen Nachweis werden wir offenbar dazu geführt, die Grössenordnung von $f(s)$ für ins Unendliche wachsende t zu untersuchen. Bekanntlich gilt für jede für $\sigma > 0$ konvergente Dirichletsche Reihe $f(s) = \sum \frac{a_n}{n^s}$ bei jedem $0 < \sigma < 1$ für $t \rightarrow \infty$ die Abschätzung

$$f(s) = O(t^{1-\sigma+\epsilon})$$

für beliebiges $\epsilon > 0$. Zur Vorbereitung der später vorzunehmenden Konstruktion wird es nützlich sein, kurz an die klassische

Herleitung dieser Abschätzung zu erinnern, wobei wir für unsere Zwecke die Voraussetzung ein bischen einschränken, indem wir nicht nur Konvergenz der Reihe für $\sigma > 0$ annehmen, sondern darüber hinaus Beschränktheit der Koeffizientensumme $S_m = \sum_{\nu=1}^m a_\nu$ voraussetzen, etwa

$$|S_m| \leq 1 \text{ für alle } m,$$

dafür aber auch eine etwas weitergehende Aussage beweisen können, nämlich

$$f(s) = f(\sigma + it) = O(t^{1-\sigma}),$$

wobei ε weggefallen ist. Zur Herleitung dieser Abschätzung stellen wir üblicherweise die Funktion $f(s)$ in der Halbebene $\sigma > 0$ durch die (absolut konvergente) Reihe

$$f(s) = \sum_{m=1}^{\infty} S_m \left(\frac{1}{m^s} - \frac{1}{(m+1)^s} \right)$$

dar, welche durch partielle Summation aus der ursprünglichen Reihe sofort hervorgeht. Aus der Voraussetzung $|S_m| \leq 1$ für alle m folgt daher in der ganzen Halbebene $\sigma > 0$ die Abschätzung

$$|f(s)| \leq \sum_{m=1}^{\infty} \left| \frac{1}{m^s} - \frac{1}{(m+1)^s} \right|.$$

Wir werden diese Reihe für $0 < \sigma < 1$ und etwa $t > 1$ abschätzen. Für die auftretende Differenz $\frac{1}{m^s} - \frac{1}{(m+1)^s}$ bieten sich von selbst zwei verschiedenartige Abschätzungen dar, nämlich einerseits

$$(A) \quad \left| \frac{1}{m^s} - \frac{1}{(m+1)^s} \right| \leq \frac{1}{m^\sigma} + \frac{1}{(m+1)^\sigma} < \frac{2}{m^\sigma}$$

und andererseits

$$\begin{aligned} \left| \frac{1}{m^s} - \frac{1}{(m+1)^s} \right| &= \left| s \int_m^{m+1} \frac{du}{u^{1+s}} \right| \\ &\leq |s| \int_m^{m+1} \frac{du}{u^{1+\sigma}} < \frac{|s|}{m^{1+\sigma}} \end{aligned}$$

also (da $|s| < 2t$) die Abschätzung

$$(B) \quad \left| \frac{1}{m^s} - \frac{1}{(m+1)^s} \right| < \frac{2t}{m^{1+\sigma}}.$$

Vergleichen wir die beiden Ausdrücke $\frac{2}{m^\sigma}$ und $\frac{2t}{m^{1+\sigma}}$, die für $t = m$ zusammenfallen, so sehen wir sofort, dass der erste der bessere (d. h. kleinere) für $m < t$ und der zweite der bessere für $m > t$ ist. Wir schätzen daher für ein festes σ in $0 < \sigma < 1$ und $t \rightarrow \infty$ folgendermassen ab:

$$\begin{aligned} |f(s)| &\leq \sum_{m=1}^{\infty} \left| \frac{1}{m^s} - \frac{1}{(m+1)^s} \right| \\ &\leq \sum_{m \leq t} \frac{2}{m^\sigma} + \sum_{m > t} \frac{2t}{m^{1+\sigma}} \\ &= O\left(2 \int_1^t \frac{du}{u^\sigma}\right) + O\left(2t \int_t^\infty \frac{du}{u^{1+\sigma}}\right) \\ &= O(t^{1-\sigma}). \end{aligned}$$

Es wird sich im folgenden darum handeln, eine Dirichletsche Reihe $f(s) = \sum \frac{a_n}{n^s}$ mit der Konvergenzabszisse 0 (und welche im Punkte 0 das obige Verhalten $|S_m| \leq 1$ aufweist) zu konstruieren, welche grob gesagt auf jeder Geraden $\sigma = \sigma_0$ ($0 < \sigma_0 < 1$) in gewissen Punkten $\sigma_0 + it_\nu$ (wo die Folge der t_ν von σ_0 unabhängig ist) mit $t_\nu \rightarrow \infty$ so gross wird, wie nach der obigen Abschätzung überhaupt möglich ist, d. h. für welche

$$|f(\sigma_0 + it_\nu)| > ct_\nu^{1-\sigma_0} \quad (\nu = 1, 2, \dots)$$

wird, wo $c = c(\sigma_0)$ eine positive Konstante ist. Überdies müssen wir gewisse Forderungen (z. B. in bezug auf Gleichmässigkeit betreffs der auftretenden Konstanten $c = c(\sigma_0)$) stellen.

Zunächst einige allgemeine Bemerkungen. Wie ein Blick auf die obige Abschätzung lehrt, ist es offenbar nötig, wenn wir erreichen wollen, dass $f(\sigma + it)$ für ein grosses t tatsächlich von der Grössenordnung $t^{1-\sigma}$ ausfallen soll, das Augenmerk auf die-

jenigen Glieder $S_m \left(\frac{1}{m^s} - \frac{1}{(m+1)^s} \right)$ zu richten, für welche m die Grössenordnung des momentan betrachteten t hat, denn sowohl diejenigen Glieder, für welche m wesentlich kleiner als t ist, etwa $m < t^{1-\delta}$, als auch diejenigen, für welche m wesentlich grösser als t ist, etwa $m > t^{1+\delta}$, geben offenbar zusammen einen Beitrag von kleinerer Grössenordnung als $t^{1-\sigma}$. Bei der Konstruktion haben wir sowohl die Möglichkeit, die Aufmerksamkeit auf diejenigen Glieder zu richten, für welche m in der Nähe von t unterhalb t liegt (und wo wir oben die Abschätzung $\frac{2}{m^\sigma}$ verwendet haben) als auch auf diejenigen Glieder in der Nähe von t , für welche m oberhalb t liegt, (und für welche wir oben die Abschätzung $\frac{2t}{m^{1+\sigma}}$ verwendet haben). Bei der früheren Untersuchung in meiner oben zitierten Arbeit, die einen ähnlichen Zweck verfolgt, habe ich an der betreffenden Stelle die Glieder mit $m > t$ bevorzugt. Es wird aber hier, wo ein etwas weitergehendes Resultat bezüglich der Grössenordnung zu erzielen ist, einfacher sein, umgekehrt die Glieder mit $m < t$ besonders in Auge zu fassen.

Wir werden den folgenden Satz beweisen:

Satz: *Es gibt eine Dirichletsche Reihe $f(s) = \sum \frac{a_n}{n^s}$ mit der Konvergenzabszisse 0, zu der eine Ordinatenfolge $t_1 < t_2 < \dots < t_\nu \rightarrow \infty$ derart existiert, dass es zu jedem Wertepaare $0 < \sigma_1 < \sigma_2 < 1$ eine positive Konstante $c = c(\sigma_1, \sigma_2)$ und ein $N = N(\sigma_1, \sigma_2)$ gibt, so dass für jedes $\nu > N$ auf der ganzen Strecke $\sigma_1 \leq \sigma \leq \sigma_2$, $t = t_\nu$ der Betrag von $f(\sigma + it_\nu)$ grösser als $ct_\nu^{1-\sigma}$ ist, während ihr Argument in einem festen Winkelraum kleiner als π etwa um die positive Achse gelegen ist.*

Der Beweis dieses Satzes wird im nächsten Paragraphen gegeben werden. Hier soll aber dargetan werden, dass mit diesem Satz tatsächlich das PERRONSche Problem gelöst wird und zwar, wie erwähnt, im negativen Sinne. In der Tat, bei fest gegebenem $0 < \varrho < 1$ seien die Abszissen σ_1 und σ_2 so gewählt, dass $0 < \sigma_1 < \sigma_2 < 1 - \varrho$. Wir betrachten die Funktion $\frac{f(s)}{s^\varrho}$ auf der Strecke $\sigma_1 \leq \sigma \leq \sigma_2$, $t = t_\nu$ ($\nu > N$). Hierbei verhält sich der Nenner $s^\varrho = e^{\varrho \log s}$ für grosse ν dem Betrage nach wie

t_ν^σ , während sein Argument (gleichmässig in σ) gegen $\frac{\pi}{2} \varrho$ konvergiert. Daher gilt für die Funktion $\frac{f(s)}{s^\varrho}$ für grosse ν auf der ganzen Strecke $\sigma_1 \leq \sigma \leq \sigma_2, t = t_\nu$, dass ihr Betrag grösser ist als etwa

$$\frac{c t_\nu^{1-\sigma}}{2 t_\nu^\varrho} = \frac{c}{2} t_\nu^{1-\sigma-\varrho} \geq \frac{c}{2} t_\nu^{1-\sigma_2-\varrho},$$

welche letztere Grösse mit ν gegen ∞ strebt wegen $\sigma_2 < 1 - \varrho$, während ihr Argument in einem festen Winkelraum kleiner als π gelegen ist, weil dies für das Argument des Zählers gilt, während das Argument des Nenners gegen eine feste Grösse konvergiert. Dies involviert aber offenbar, dass das Integral

$\int_{\sigma_1 + it_\nu}^{\sigma_2 + it_\nu} \frac{f(s)}{s^\varrho} ds$ dem Betrage nach mit ν über alle Grenzen wächst,

womit, wie oben ausgeführt, die PERRONSche Frage erledigt ist.

Beweis des Satzes durch die Konstruktion eines Beispiels.

Wir wählen eine beliebige, aber feste Folge von positiven Zahlen $t_1 < t_2 < \dots$ mit $t_\nu \rightarrow \infty$, derart dass $\frac{t_{\nu+1}}{t_\nu} \rightarrow \infty$. Der Bequemlichkeit halber werden wir die t_ν ganz und sogar durch 12 teilbar annehmen und so, dass schon von Anfang an $\frac{t_{\nu+1}}{4} > \frac{t_\nu}{3}$ ist. Für diese t_ν (und $0 < \sigma < 1$) soll $f(\sigma + it_\nu)$ gross ausfallen. Für ein festes t_ν betrachten wir die obige Differenz

$$\frac{1}{m^{\sigma + it_\nu}} - \frac{1}{(m+1)^{\sigma + it_\nu}}$$

und zwar für Werte von m , die unterhalb t_ν gelegen, aber von der Grössenordnung von t_ν sind, genau gesprochen, für welche

$$\frac{t_\nu}{4} \leq m \leq \frac{t_\nu}{3}$$

ist. Hierbei ist das Argument des ersten Gliedes $-t_\nu \log m$, das des zweiten Gliedes $-t_\nu \log(m+1)$, also (abgesehen vom Vor-

zeichen) die Änderung des Arguments beim Übergang von $\frac{1}{m^{\sigma+it_\nu}}$ zu $\frac{1}{(m+1)^{\sigma+it_\nu}}$ gleich

$$t_\nu (\log(m+1) - \log m) = t_\nu \log\left(1 + \frac{1}{m}\right).$$

Für grosse m verhält sich diese letztere Grösse wie $\frac{t_\nu}{m}$; genauer gesprochen (wegen $\frac{5}{6} \cdot \frac{1}{m} < \log\left(1 + \frac{1}{m}\right) < \frac{1}{m}$ für $m \geq 3$) liegt sie für alle in Betracht kommenden Werte von m zwischen $\frac{5}{6} \cdot \frac{t_\nu}{m}$ und $\frac{t_\nu}{m}$, also zwischen $\frac{5}{6} \cdot 3$ und 4 und damit a fortiori zwischen $\pi - \frac{\pi}{3}$ und $\pi + \frac{\pi}{3}$, d. h. in einem festen Winkelraum um π links von der imaginären Achse. Die beiden Glieder $\frac{1}{m^{\sigma+it_\nu}}$ und $\frac{1}{(m+1)^{\sigma+it_\nu}}$ sind somit »im wesentlichen« entgegengesetzt gerichtet. Für den absoluten Betrag der Differenz $\frac{1}{m^{\sigma+it_\nu}} - \frac{1}{(m+1)^{\sigma+it_\nu}}$ gilt somit offenbar

$$\left| \frac{1}{m^{\sigma+it_\nu}} - \frac{1}{(m+1)^{\sigma+it_\nu}} \right| > \left| \frac{1}{m^{\sigma+it_\nu}} \right| = \frac{1}{m^\sigma},$$

während das Argument der Differenz in einem festen Winkelraum von $-\frac{\pi}{3}$ bis $+\frac{\pi}{3}$ um das Argument des ersten Gliedes $\frac{1}{m^{\sigma+it_\nu}}$ liegt.

Nach diesen Vorbereitungen gehen wir nunmehr zur direkten Angabe der gewünschten Funktion $f(s) = \sum \frac{a_n}{n^s}$ über. Statt den Koeffizienten a_n selbst anzugeben, ziehen wir vor, die Koeffizientensumme $S_m = \sum_1^m a_n$ vorzuschreiben. Wir wählen bei jedem $\nu = 1, 2, \dots$

$$S_m = m^{it_\nu} \quad \text{für } \frac{t_\nu}{4} \leq m \leq \frac{t_\nu}{3} \quad (\nu = 1, 2, \dots)$$

$$S_m = 0 \quad \text{für alle } m \text{ ausserhalb dieser Intervalle.}$$

Da $|S_m| \leq 1$ für alle m , ohne dass S_m für $m \rightarrow \infty$ einem Grenzwert zustrebt, so dass die Reihe $\sum a_n = \sum \frac{a_n}{n^0}$ zwischen end-

lichen Grenzen oscilliert, muss der Punkt 0 auf der Konvergenzgeraden liegen, d. h. die Konvergenzabszisse gleich 0 sein. Auf der Strecke $0 < \sigma < 1$, $t = t_\nu$ erhalten wir dann

$$f(\sigma + it_\nu) = \sum_{m=4}^{\frac{t_\nu}{3}} m^{it_\nu} \left(\frac{1}{m^{\sigma + it_\nu}} - \frac{1}{(m+1)^{\sigma + it_\nu}} \right) + R_\nu(s),$$

wobei

$$R_\nu(s) = \sum_{m=1}^{\frac{t_\nu-1}{3}} S_m \left(\frac{1}{m^{\sigma + it_\nu}} - \frac{1}{(m+1)^{\sigma + it_\nu}} \right) + \sum_{m=\frac{t_\nu+1}{4}}^{\infty} S_m \left(\frac{1}{m^{\sigma + it_\nu}} - \frac{1}{(m+1)^{\sigma + it_\nu}} \right).$$

In der Summe erstreckt von $\frac{t_\nu}{4}$ bis $\frac{t_\nu}{3}$ ist nach dem obigen und der getroffenen Wahl von $S_m = m^{it_\nu}$ (d. h. S_m hat den absoluten Betrag 1 und sein Argument ist gleich dem Argument von $\frac{1}{m^{\sigma + it_\nu}}$ mit entgegengesetztem Vorzeichen) jedes Glied $\frac{1}{m^{\sigma + it_\nu}} - \frac{1}{(m+1)^{\sigma + it_\nu}}$ dem Betrage nach grösser als $\frac{1}{m^\sigma}$ und fällt in einen Winkelraum $\left(-\frac{\pi}{3}, +\frac{\pi}{3}\right)$ um die positive Halbachse. Somit gilt für diese Summe, dass ihr Realteil und damit a fortiori ihr absoluter Betrag grösser ist als

$$\cos \frac{\pi}{3} \cdot \sum_{m=4}^{\frac{t_\nu}{3}} \frac{1}{m^\sigma} > \frac{1}{2} \int_{\frac{t_\nu}{4}}^{\frac{t_\nu}{3}} \frac{du}{u^\sigma} = c(\sigma) t_\nu^{1-\sigma},$$

wo die positive Konstante $c(\sigma)$ nicht von t_ν abhängt und in jedem festen Intervall $0 < \sigma_1 \leq \sigma \leq \sigma_2 < 1$ grösser als eine feste Konstante $c = c(\sigma_1, \sigma_2)$ bleibt. Ferner ist das Argument dieser Summe gleichfalls in dem betrachteten festen Winkelraum $\left(-\frac{\pi}{3}, +\frac{\pi}{3}\right)$ gelegen. Wir haben nunmehr das Restglied $R_\nu(s)$ abzuschätzen. Dies ist aber mit wenigen Worten getan. Nach den Abschätzungen (A) und (B) haben wir

$$|R_\nu(s)| < \frac{t_\nu - 1}{\sum_{m=1}^3} \frac{2}{m^\sigma} + \sum_{m=\frac{t_\nu+1}{4}}^{\infty} \frac{2t_\nu}{m^{\sigma+1}},$$

also für $\nu \rightarrow \infty$

$$\begin{aligned} R_\nu(s) &= O\left(2 \int_1^{\frac{t_\nu-1}{3}} \frac{du}{u^\sigma}\right) + O\left(2t_\nu \int_{\frac{t_\nu+1}{4}}^{\infty} \frac{du}{u^{1+\sigma}}\right) \\ &= O(t_{\nu-1}^{1-\sigma}) + O(t_\nu \cdot t_{\nu+1}^{-\sigma}) = o(t_\nu^{1-\sigma}), \end{aligned}$$

und zwar gilt diese Abschätzung gleichmässig in jedem festen Intervall $0 < \sigma_1 \leq \sigma \leq \sigma_2 < 1$. Schliesslich ergibt sich also für $\nu \rightarrow \infty$ in jedem festen Intervall $0 < \sigma_1 \leq \sigma \leq \sigma_2 < 1$ die Ungleichung

$$|f(\sigma + it_\nu)| > c t_\nu^{1-\sigma} + o(t_\nu^{1-\sigma}).$$

Hieraus ist aber ersichtlich, dass $f(\sigma + it_\nu)$ für hinreichend grosse ν auf der Strecke $\sigma_1 \leq \sigma \leq \sigma_2$ absolut genommen grösser als $\frac{c}{2} t_\nu^{1-\sigma}$ ist, und ferner ist klar, dass ihr Argument in einem festen Winkelraum, etwa $\left(-\frac{2}{5}\pi, +\frac{2}{5}\pi\right)$, von einer Öffnung kleiner als π liegt. Hiermit ist der Satz bewiesen.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 3

ON THE REACTION BETWEEN CARBON MONOXIDE AND ALCOHOL CATALYZED BY ALCOHOLATE

BY

J. A. CHRISTIANSEN AND J. C. GJALDBÆK



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1942

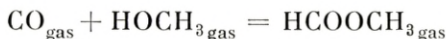
Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

Introduction.

In 1914, A. STÄHLER (1) showed that CO combines with dry R₂RONa (R = alcy) to form NaCOOR or C(OR)(ONa), but not RCOONa in detectable amounts. Moreover, he found that at a pressure of a few hundred atmospheres and at a temperature of 100—200°, CO also combines with alcoholates dissolved in the corresponding alcohol, forming alcyformates, the alcoholate acting as a catalyst.

The experiments were merely preparative, and no exact investigation of the kinetics of the reaction was intended. The author states, however, that the reaction was faster with methanol than with ethanol. Our experiments prove that, at pressures below 1 atm. and correspondingly lower temperatures, the reaction still proceeds at a measurable rate, but about 3 times faster in the case of ethanol than in the case of methanol.

In 1926, it was shown by one of the present authors (2) that the equilibrium



can be attained using solid NaOCH₃ as a catalyst, and that the equilibrium constant K_c is approximately given by the expression

$$\log K_c = 1880/T + 4.82.$$

The present paper deals with the kinetics of the reaction which so far has not been investigated, except by STÄHLER.

Materials.

CO was prepared according to L. MOSER (3). Anhydrous formic acid was added drop by drop to concentrated sulphuric acid in a small flask; the CO evolved was led through soda

lime to free it from CO_2 and was collected in a glass bulb gasometer of the BODENSTEIN type over concentrated sulphuric acid. By analyses in a micro-gas-burette (4), the gas was found to contain about 98% CO —the balance being air and, strangely enough, traces (about 0.5%) of CO_2 . These small impurities were considered harmless and no further purification was attempted.

Anhydrous alcohols were prepared according to HAKON LUND and J. BJERRUM (5), omitting, however, the distillation from tribromobenzoic acid. The preparations were tested as follows.

Methanol:

a) 10 ml. did not decolorize 1 drop of 0.1 n KMnO_4 in the course of 15 min. (no aldehyde).

b) 10 ml. consumed less than 1 drop of 0.1 n NaOH (no formate) and less than 1 drop of 0.1 n HCl (no base).

c) 5 ml. + 10 ml. of 2 n NaOH + 25 ml. of 0.1 n I_2 remained clear after standing for 5 minutes and gave no odour of iodoform (no acetone).

Ethanol:

10 ml. did not decolorize 5 drops of 0.1 n KMnO_4 in the course of 30 min. (no aldehyde).

No acid or basic impurities as under methanol.

Preparation of sodium alcoholate. Freshly cut sodium was freed from kerosene by means of filter paper and dissolved in anhydrous alcohol. It was stored in portions of 100 to 200 ml. in cork-stoppered flasks and was protected against CO_2 by means of soda lime. After dilution with water the solutions were titrated with 0.1 n HCl and methyl red as indicator. The concentrations at higher temperatures were calculated by means of the following table which is based on YOUNG'S (6) measurements for pure alcohols.

Table of Specific Gravities.

	35°	45°	54°
CH_3OH	0.7781	0.7693	0.7615
$\text{C}_2\text{H}_5\text{OH}$	0.7762	0.7678	0.7599

Dioxane was purified according to KURT HESS and HERMANN FRAHM (7). The commercial product was freed from water and

ethylene-acetal by prolonged boiling with thin slices of sodium and subsequent distillation from the sodium. Shaking with mercury showed the absence of peroxides, and the test for aldehydes with fuchsin-sulphuric acid was also negative.

Apparatus.

A shaking apparatus according to BRØNSTED was used (Fig. 1). R is a glass-stoppered flask holding 75.5 ml. connected with a mercury manometer through the capillary g. The dead space

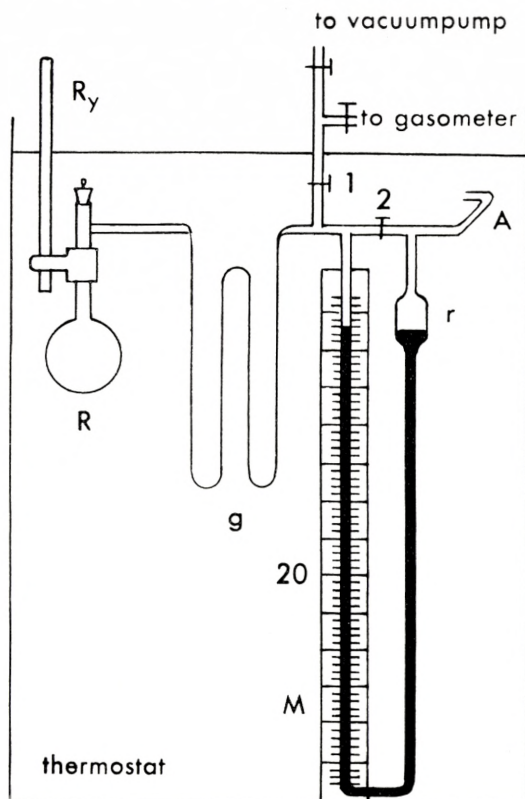


Fig. 1.

from the neck of the flask to the surface of the mercury in the left branch of the manometer is about 7.8 ml. It is essential to make this space as small as possible, since diffusion occurs at

the start of the experiment, giving rise to pressure variations which have nothing to do with the reaction itself. At A the tube bends backwards, connecting the mercury container r with a compensation flask behind R through a capillary similar to g. M is a milk glass scale graduated in millimeters; the manometer is read through a lens. R_y is a shaking device carrying R and the compensation flask, so that both are shaken simultaneously. Through stopcock 1, the apparatus may be evacuated and filled with CO. Stopcock 2 serves to open between the two flasks and the two manometer branches (for evacuation) and to close it for filling R with CO. Obviously, this stopcock remains closed during the measurements. 1 and 2 are lubricated with a rather stiff grease capable of standing the highest temperatures used. The whole apparatus, including stopcocks 1 and 2, is kept in a thermostat, electrically heated and regulated to $\pm 0.01^\circ$ in the usual manner by means of a toluene mercury regulator with relay. To avoid sparking, the current from the regulator is led to the grid of an ordinary triode radio valve whose anode current then actuates the relay. The temperature is read on a thermometer graduated in $1/10$ th degrees, the readings of which are corrected by comparison with a standardized thermometer. Only the corrected temperatures are given in the tables. For time measurements, a stopwatch is used in slow experiments, and a wall clock sounding the minutes in fast ones.

Procedure.

By means of a pipette, 22.00 ml. of catalyst solution (e. g., 1 molar sodium methylate in methanol) were transferred into the reaction flask (R), and 22 ml. of pure alcohol into the compensation flask. Stopcocks 1 and 2 were opened, and the apparatus was pumped until the liquids began to boil; subsequently, stopcock 1 was closed. If the pumping were continued, the concentration of the catalyst solution would be altered by evaporation.

The flasks were now shaken for about 10 minutes until the liquids had attained the temperature of the thermostat. Stopcock 2 was closed, and a portion of CO was let in through

stopcock 1, causing a depression of the manometer to, say, mark 20.

The shaking of the flasks was now resumed, the readings of time and pressure being started simultaneously; the manometer received a few taps before each reading. As the experiment progressed, the CO in the reaction flask was absorbed and the pressure decreased, as indicated by a rise of the mercury in the left branch of the manometer to about "mark zero".

Since the alcohol vapour pressure was eliminated by means of the compensation flask, the manometer recorded only the CO-pressure in the reaction flask.

Since the reaction proved to be of the first order, and it therefore seemed appropriate to apply GUGGENHEIM's method (8) to the calculations, an initial set of readings was performed (column 2 of table 1) at definite times, which was followed by a second set of readings (column 3 of table 1) at the same time intervals, beginning at a given time (Z) after the start of the experiment. The time Z which is given in the table was always so chosen as to be of the order of magnitude of 2 to 5 times the half-time of the reaction.

After each experiment, 5.00 ml. of the cooled reaction mixture were titrated with 0.1 n HCl.

Experimental data.

At each of the temperatures of about 35°, 45°, and 54°, experiments were made with approximately 1, 1/2, and 1/4 molar sodium methylate in methanol as well as with sodium ethylate in ethanol. At 54°, additional experiments were performed with approximately 1/2 molar concentrations of catalysts in a mixture of 50 % dioxane and 50 % alcohol.

All pressure readings were corrected according to the correction table given below, since it was found that blank experiments with pure alcohol in both flasks and CO excess pressure in flask R gave a reproducible small increase in pressure. Inasmuch as this increase did not follow a course of reaction of either the first or the second order, a blank experiment was carried out at each temperature, and the corrections obtained were subtracted from all pressures measured later.

In principle, objections can be raised to this method of correction, but the results obtained will nevertheless be substantially better than the uncorrected ones.

Originally, the corrections were much greater. However, after a re-dimensioning of the apparatus they became smaller. Unfortunately we have not succeeded in eliminating them altogether. The errors are probably caused by diffusion phenomena in the apparatus.

Correction Table.

Methanol					Ethanol					
P _{init.}	35.05°	45.05°	45.05°	54.16°	P _{init.}	35.05°	45.05°	54.16°	54.16°	54.16°
Min.	20 cm.	10 cm.	20 cm.	20 cm.	Min.	20 cm.	20 cm.	10 cm.	20 cm.	30 cm.
0	0	0	0	0	..	0	0	0	0	0
5	0.29	0.30	0.65	1.60	2	0.02	0.03	0.01	0.20	0.28
10	0.43	0.68	0.81	1.90	5	0.04	0.07	0.02	0.49	0.54
15	0.50	0.73	0.87	2.02	10	0.06	0.11	0.04	0.54	0.74
20	0.53	0.81	0.93	2.12	15	0.10	0.12	0.05	0.59	0.80
25	0.56	0.86	0.97	2.20	25	0.12	0.12	0.05	0.67	0.85
30	0.59	0.90	1.00	2.26	35	0.14	0.13	0.07	0.74	0.88
40	0.62	1.02	1.05	2.37	45	0.14	0.14	0.10	0.77	0.90
50	0.64	1.05	1.08	2.48	60	0.15	0.16	0.12	0.80	0.94
60	0.66	1.07	1.10	2.55	90	0.16	0.18	0.16	0.84	0.96
75	0.68	1.09	1.13	2.64	120	—	0.24	0.16	0.85	0.96
90	0.69	1.10	1.15	2.71	180	—	0.32	0.20	0.87	1.03
120	0.72	1.12	1.22	2.80	240	—	0.33	0.23	0.92	..
150	0.74	1.15	1.24	3.85	300	—	0.34	..	0.96	..
240	0.79	1.16	1.31	2.96	∞	..	0.51
300	0.80	1.16	1.32	3.00
∞	0.90	1.18	1.46	3.20

Calculation of the experiments.

If the reaction is assumed to be of the first order, and $k = -\frac{1}{t} \ln \frac{p-p_{\infty}}{p_0-p_{\infty}}$ is calculated for experiments 3, 4, and 14, 15, and 17 where the ratios of the initial pressures are 2/1 and 3/2/1, respectively, we find almost the same value for k when converting to the same concentration of catalyst. Thus, the reaction is in all probability of the first order, which is further

proved by the fact that k is almost constant throughout each experiment. A slight "drift" of the constant could, however, be demonstrated in almost all experiments, k first increasing, then reaching a maximum, and finally decreasing again. The maximum value was 1—4 % higher than the mean value.

It was assumed that all the time the shaking of the flasks was sufficiently effective to secure equilibrium between the CO in the gas phase and in the liquid phase.

The ester formed during the experiment has a low boiling point (HCOOCH_3 : B. p. 31.8° ; HCOOC_2H_5 : B. p. 54.1° at 1 atm.) so that considerable amounts of ester must pass over into the vapour phase. The pressure p read is thus the sum of the CO-pressure (p_{CO}) and the ester pressure (p_{E})

$$p = p_{\text{CO}} + p_{\text{E}} \quad (1)$$

If only the velocity constants have to be calculated, the ester pressure causes no error in the calculation, as will be seen from the following considerations.

Since the ester concentration is small, its pressure must be proportional to its concentration in the solution, which again is proportional to the amount of CO absorbed. Consequently, at the time t we find

$$p_{\text{E},t} = a \cdot (p_0 - p_{\text{CO},t}), \quad (2)$$

where a is a proportionality factor, p_0 the initial pressure of CO, and $p_{\text{CO},t}$ the CO-pressure at the time t . Substituting (2) into (1) we find

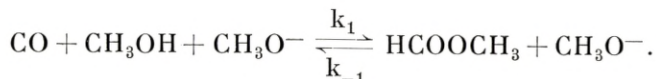
$$p = p_{\text{CO},t}(1 - a) + p_0 a. \quad (3)$$

In addition, we have

$$p_\infty = p_{\text{CO},\infty}(1 - a) + p_0 a \quad (4)$$

$$p - p_\infty = (p_{\text{CO},t} - p_{\text{CO},\infty})(1 - a) \quad (5)$$

As the reaction in solution, like the corresponding reaction in the gas phase, is assumed to be reversible, it follows



The alcohol concentration is almost constant because the alcohol is present in a large excess; furthermore, the concentration of catalyst is constant, and therefore it is not surprising that the reaction follows the unimolecular equation to which applies

$$\log (N_1 - N_{1,\infty}) = \log N \frac{k_1}{k_1 + k_{-1}} - (k_1 + k_{-1}) t \cdot \log e, \quad (6)$$

where N_1 and N_2 are the number of CO- and ester molecules, respectively, $N_{1,\infty}$ the number of CO-molecules in the equilibrium state, and $N_1 + N_2 = N$.

Since, according to (5), $N_1 - N_{1,\infty}$ is proportional to $p - p_\infty$, we find

$$\log (p - p_\infty) = -\text{constant}_1 \cdot t + \text{constant}_2,$$

i. e., the logarithm of the pressure difference decreases linearly with time; the ester pressure causes only a parallel displacement of the straight lines, their slopes give the velocity constant.

The calculation of the experiments was carried out according to GUGGENHEIM (8) with the modification that the proposed geometric method for the determination of k was replaced by a mathematical method proposed by J. A. CHRISTIANSEN (9). An example of this calculation is given in table 1 (experiment 11).

According to GUGGENHEIM's method, $\log (p' - p)$ is plotted as ordinates and t as abscissae. A straight line is then drawn through the resulting points, and k is obtained from the slope of this line (cf. fig. 2).

However, the slope of the line may also be obtained by dividing differences in the logarithm with corresponding differences in time. This is most accurately done by extrapolating the value of $\log (p' - p)$ at $t = 0$ by plotting $\log (p' - p)$ against t on a large scale; when drawing a straight line through the points obtained, the desired value is found as the point of intersection of the line with the log-axis (1.3500). Applying this and the following values of $\log (p' - p)$, the differences in column 6 were calculated successively. The corresponding time differences are in column 1. A series of k -values throughout the experiment was then obtained by calculating $\frac{\Delta \log}{\Delta t}$; these values are recorded in column 7.

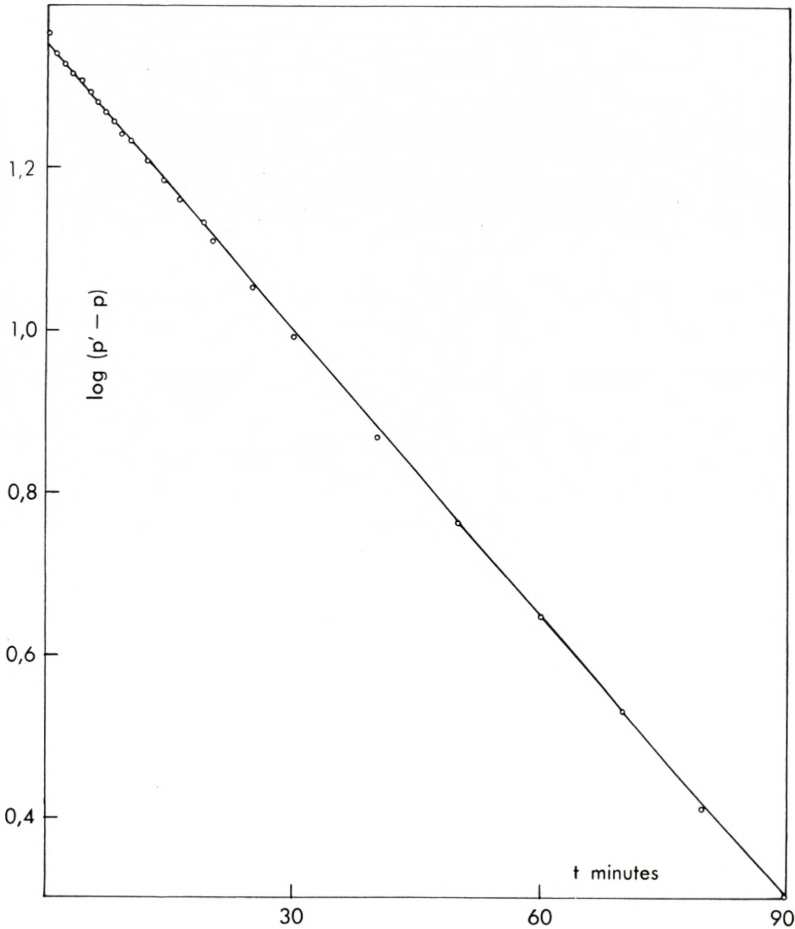


Fig. 2. The equation for the curve is $\Delta \log (p' - p) = 11.74 \cdot 10^{-3} \cdot \Delta t$.

According to J. A. CHRISTIANSEN (9), however, the most correct mean value of k is obtained by using the formula

$$k_{\text{mean}} = \frac{\sum G_i \cdot K_i}{\sum G_i},$$

where G_i denotes the statistical weight of the constant k_i . G_i (column 9) for the different stages (x) of a first order reaction is obtained from a table computed according to (9).

Practically, G which need not be particularly accurate, is read from a curve drawn on the basis of the values given in the table.

Table 1.

Experiment 11. CO + 1.501 mol CH₃ONa in CH₃OH at 54.16°.

1	2	3	4	5	6	7	8	9
Min.	p	p'	p'—p	log (p'—p)	△ log	k · 10 ³	x	G _i
0	17.90	41.19	23.29	1.3672 ¹	0		0	0
1	19.31	—23	21.92	1.3408	0.0092	9.2	0.06	0.20
2	20.03	—31	21.28	1.3280	0.0220	11.00	0.08	0.28
3	20.68	—40	20.72	1.3164	0.0336	11.20	0.11	0.38
4	21.22	—42	20.20	1.3053	0.0447	11.18	0.13	0.44
5	21.81	—47	19.66	1.2936	0.0564	11.28	0.15	0.50
6	22.41	—51	19.10	1.2810	0.0690	11.50	0.18	0.57
7	22.97	—55	18.58	1.2691	0.0809	11.56	0.20	0.61
8	23.53	—61	18.08	1.2572	0.0928	11.60	0.22	0.65
9	24.10	—65	17.55	1.2443	0.1057	11.74	0.24	0.68
10	24.64	—69	17.05	1.2317	0.1183	11.83	0.26	0.72
12	25.62	—78	16.16	1.2085	0.1415	11.79	0.30	0.75
14	26.59	—87	15.28	1.1841	0.1659	11.85	0.34	0.78
16	27.50	—94	14.44	1.1596	0.1904	11.90	0.38	0.80
18	28.36	—98	13.62	1.1331	0.2169	12.05	0.41	0.79
20	29.17	42.06	12.89	1.1103	0.2397	11.99	0.44	0.78
25	30.93	—20	11.27	1.0519	0.2981	11.92	0.52	0.72
30	32.52	—30	9.78	0.9904	0.3596	11.99	0.58	0.68
40	35.15	—53	7.38	0.8681	0.4819	12.05	0.68	0.51
50	36.96	—75	5.79	0.7627	0.5873	11.75	0.75	0.38
60	38.44	—88	4.44	0.6474	0.7026	11.71	0.81	0.25
70	39.59	—98	3.39	0.5302	0.8198	11.71	0.86	0.18
80	40.53	43.10	2.57	0.4099	0.9401	11.75	0.89	0.11
90	41.19	—20	2.01	0.3032	1.0468	11.63	0.92	0.07

$$\Sigma G = 11.63$$

$$Z = 1.5 \text{ hours.}$$

$$\Sigma k \cdot G = 136.5078$$

$$k = \frac{136.5078}{11.63} = 11.74 \cdot 10^{-3} \text{ min.}^{-1}$$

¹ log (p'—p) extrapolated to t = 0 gives 1.3500.

Finally, we calculate $\Sigma G = 11.63$ and $\Sigma Gk = 136.5078$, whereupon k_{mean} is found to be $11.74 \cdot 10^{-3}$ (¹⁰logarithms, min.⁻¹).

All k's are calculated according to the above described method.

The accuracy of the measured values of k is impaired for

different reasons. Owing to the high vapour pressure of the alcohols (cf. the table below), a slight difference in the temperatures of the reaction flask and the compensation flask causes an error of the pressure measured. At the most (CH_3OH at 55°), a temperature difference of 0.02° between the two flasks will cause a pressure difference of 0.4 mm. Moreover, the correction mentioned on p. 7 may slightly falsify the results. The following alcohol vapour pressures given in cm Hg are interpolated on the basis of RICHARDSON'S results (10).

Temp.	CH_3OH	$\text{C}_2\text{H}_5\text{OH}$
35°	19.8	11.6
45°	38.4	17.6
55°	47.0	28.0

Experimental results and discussion.

A survey of the experimental results is given in tables 2 and 3. The k -values are calculated on the basis of natural logarithms and seconds.

Table 2.
Experiment with $\text{CO} + \text{CH}_3\text{OH}$; catalyst CH_3ONa .

1	2	3	4	5	6	7	8
Expe- riment	Abs. temp.	c_{cat}	$P_{\text{abs.}}$ cm. Hg	$k \cdot 10^5$ sec. ⁻¹	$\frac{k \cdot 10^5}{c_{\text{cat}}}$	$k^\circ 10^5$ sec. ⁻¹	$k'' 10^5$ sec. ⁻¹
6	308.15	1.004	27.70	7.25	7.25	4.15	11.3
8	—	0.4181	18.95	2.16	5.16		
7	—	0.2219	18.17	1.05	4.73		
3	318.15	0.9960	25.63	17.61	} 18.3	11.2	30.2
4	—	—	11.09	18.76			
20 a	—	0.5421	21.32	8.18	} 13.9		
1	—	0.4403	25.10	5.69			
5	—	0.2201	24.00	2.82			
11	327.26	1.051	25.30	45.06	42.9	31.0	82.8
9	—	0.4123	17.01	14.44	35.0		
12	—	0.2207	19.47	7.33	33.2		

In the following experiment the reaction mixture contained 50 % dioxane.

26	327.26	0.4811	19.32	51.45			(286)
----	--------	--------	-------	-------	--	--	-------

Table 3.
Experiment with $\text{CO} + \text{C}_2\text{H}_5\text{OH}$; catalyst $\text{C}_2\text{H}_5\text{ONa}$.

1	2	3	4	5	6	7	8
Expe- riment	Abs. temp.	c_{cat} .	$P_{\text{abs.}}$ cm. Hg	$k \cdot 10^5$ sec. ⁻¹	$\frac{k \cdot 10^5}{c_{\text{cat}}}$	$k \cdot 10^5$ sec. ⁻¹	$k'' \cdot 10^5$ sec. ⁻¹
22	308.15	0.9976	19.87	16.33	16.4		
23	—	0.4972	21.97	7.14	14.3	12.8	34.8
24	—	0.2540	19.34	3.48	13.7		
19	318.15	0.9712	21.00	39.76	41.0		
20	—	0.4835	20.66	18.93	39.2	34.3	94.0
21	—	0.2369	20.89	8.31	35.0		
13	327.26	0.9616	20.46	88.62	97.1		
14	—	0.4837	34.53	41.76	} 86.5	79.8	212
15	—	—	20.66	41.37			
16	—	—	30.97	43.41			
17	—	—	11.97	41.03			
18	—	0.2434	19.56	20.20			
25	327.26	0.5187	21.30	118.8			(611)

In the following experiment the reaction mixture contained 50 % dioxane.

After each experiment, 5.00 ml. of the reaction mixture were diluted with 50 ml. of CO_2 -free water and titrated with 0.1 n HCl; the ester concentrations found are recorded in table 4. After dilution, the following reactions take place:



The decrease in the quantity of base during each experiment must therefore be equivalent to the amount of ester formed, which again is equivalent to the amount of CO absorbed. The results which are summarized in table 4 show but a poor accuracy, partly because of the small amount of ester formed as compared to the catalyst concentration, and partly because of the volatility of the ester.

The above experiments prove that the reaction is of the first order with respect to CO. Moreover, at high dilution of CH_3O^- , the constant k is almost proportional to the concentration of CH_3O^- , increasing somewhat more rapidly at higher concentrations.

Presumably, we have here to do with some kind of salt

Table 4.

Experiment	CO _{abs.} mmol	Ester formed mmol	$\frac{\text{mmol ester}}{\text{mmol CO}_{\text{abs.}}}$	
6	0.8659	0.754	0.87	CO + CH ₃ O ⁻
8	0.5924	0.577	0.98	
7	0.5680	0.505	0.89	
3	0.7761	0.781	1.01	
4	0.3470	0.324	0.93	
20 a	0.6456	0.686	1.06	
1	0.7600	0.638	0.84	
5	0.7267	0.797	1.10	
11	0.7451	0.817	1.10	
9	0.5033	0.510	1.02	
12	0.5734	0.546	0.95	
22	0.6211	0.680	1.09	
23	0.6868	0.714	1.04	
24	0.6045	0.633	1.05	
19	0.6359	0.678	1.07	
20	0.6256	0.683	1.09	
21	0.6325	0.678	1.07	
13	0.6025	0.523	0.87	
14	1.0169	1.029	1.01	
15	0.6084	0.573	0.94	
16	0.9121	0.878	0.96	
17	0.3407	0.360	1.05	
18	0.5670	0.575	1.00	

effect similar to what is known from the cane sugar inversion, where LEININGER and KILPATRICK (11) found the relation

$$\log \frac{k}{C_{\text{HCl}}} = \alpha + \beta \cdot C_{\text{HCl}}$$

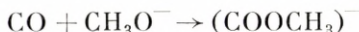
a relation which, incidentally, was found as early as in 1850 by WILHELMY (12).

An analogous relation seems to apply in our case; for data, cf. tables 2 and 3, column 6. By graphical extrapolation to $c_{\text{CH}_3\text{O}^-} = 0$, $k^\circ = \frac{k}{c_{\text{cat}}}$ was obtained, as recorded in tables 2 and 3, column 7.

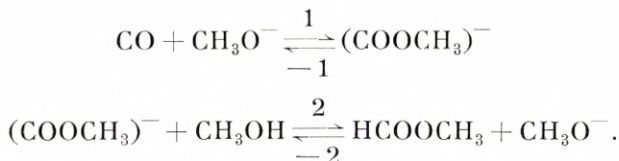
From the foregoing, it may be assumed that the velocity expression is

$$v = k \cdot c_{\text{CO}} \cdot c_{\text{CH}_3\text{O}^-}$$

and it follows that the velocity-determining reaction is



so that the complete reaction sequence becomes



However, only the existence of reaction 1 has been demonstrated directly or has been made probable by the experiments described. Since the stoichiometry of the reaction is known, the above stated reaction-sequence becomes the simplest one consistent with the experimental results. As early as in 1868, A. GEUTHER (13) showed that the decomposition reaction $\text{HCOOCH}_3 \rightarrow \text{CO} + \text{CH}_3\text{OH}$ is also catalyzed by CH_3O^- .

The velocity constant k defined previously by the equation $\frac{dp_{\text{CO}}}{dt} = -k \cdot p_{\text{CO}}$ depends, however, on the dimensions of the apparatus, since the reaction takes place in the liquid phase only. The velocity constant of the reaction in the liquid phase may be defined by the equation

$$\frac{dx}{dt} = k' \cdot c_{\text{CO}} \cdot c_{\text{cat.}}$$

where x is the amount of formate formed per litre of liquid; c_{CO} and $c_{\text{cat.}}$ denote the concentration of CO and catalyst, respectively, in the liquid phase. Disregarding the vapour pressure of the formate, we find, however

$$-\frac{dx}{dt} = \frac{1}{RT} \cdot \frac{dp}{dt} \cdot \frac{V_g}{V_v},$$

where V_g and V_v are the volumes of the gas phase and the liquid phase, respectively, and

$$c_{\text{CO}} = \frac{p}{RT} \cdot \alpha,$$

where α is the distribution coefficient of CO between the liquid and the gas phase. Consequently, we have

$$-\frac{dp}{dt} \cdot \frac{V_g}{V_v} = k' \cdot p \cdot \alpha \cdot c_{\text{cat.}}$$

from which

$$k' = k \cdot \frac{V_g}{V_v} \cdot \frac{1}{\alpha \cdot c_{\text{cat.}}} \rightarrow k^{\circ} \cdot \frac{V_g}{V_v} \cdot \frac{1}{\alpha}$$

i. e.
$$\frac{k}{c_{\text{cat.}}} \rightarrow k^{\circ} \text{ when } c_{\text{cat.}} \rightarrow 0.$$

The literature available contains no data on the solubility of CO in alcohols at temperatures as high as those of the present experiments; furthermore, an extrapolation would be too inaccurate owing to lack of data; it was therefore decided instead of k' to calculate k'' defined by

$$k'' = k^{\circ} \frac{V_g}{V_v}.$$

k'' is recorded in tables 2 and 3, column 8.

In order to get an impression of the "energy of activation" (4.571 A) and the "frequency exponent" (H), A and H were calculated from the equation

$$\log k'' = -\frac{A}{T} + H$$

using the method of least squares. The figures obtained were for the reaction in methanol A = 4555, H = 10.82, and for the reaction in ethanol A = 4125, H = 9.93.

$k' = \frac{k''}{\alpha}$ should actually be used in the calculation of A and H. Applying the solubility data of G. JUST (14) who determined the solubility of CO in methanol and ethanol at 20° and 25°, a and b were calculated for each reaction according to the formula $\log \alpha = -\frac{a}{T} + b$. In the case of methanol, the results were a = 501 and b = +0.97, which, subtracted from the previously calculated A- and H-values leads to A' = 4054 and H' = 9.85.

In the case of ethanol, the results were $a = 78.64$ and $b = -0.46$, from which we find $A' = 4046$ and $H' = 10.39$. However, the calculation of a and b is rather inaccurate, since the temperatures 20° and 25° lie so close together.

The apparent disagreement between A. STÄHLER's results and ours (cf. the introduction) may perhaps be explained by means

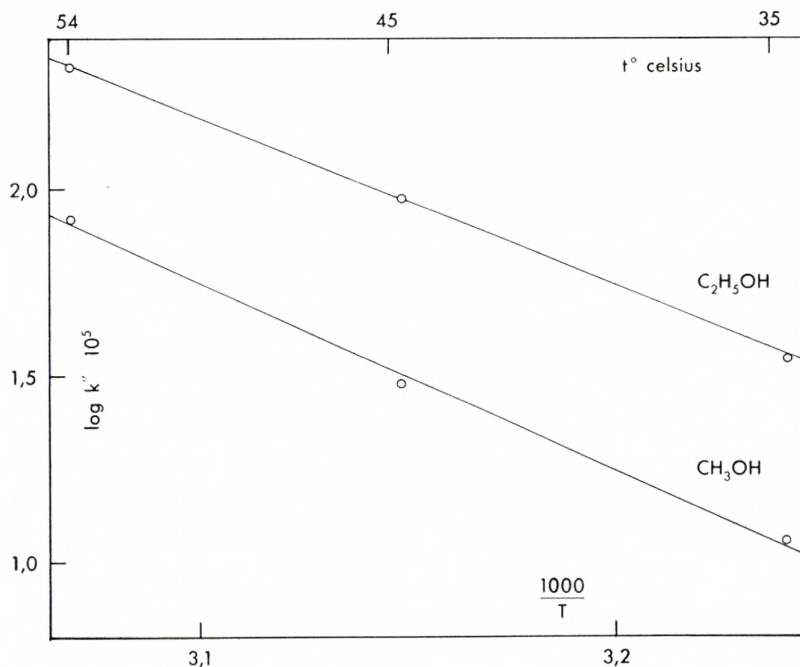


Fig. 3. The equation for the curve CH_3OH is $\log k'' = -\frac{4555}{T} + 10.82$,
the equation for the curve C_2H_5OH is $\log k'' = -\frac{4125}{T} + 9.93$.

of the calculated A- and H-values, as the curves for $\log k''$ intersect at about 200° . Consequently, above 200° the reaction in ethanol proceeds slower than in methanol, and it is faster below 200° , which might be in agreement with the observations by STÄHLER and by the present authors. However, this explanation seems somewhat uncertain, since the A-values used are only based upon three experiments within a small temperature interval (19°).

A calculation of the number of collisions S from the formula

$$S = \sigma_{1,2}^2 \cdot n_1 \cdot n_2 \cdot \sqrt{8\pi RT \cdot \frac{M_1 + M_2}{M_1 M_2}},$$

where $\sigma_{1,2}$ denotes the sum of the radii of the molecules, n the number of molecules per ml., and the M 's the molecular weights, leads to the following values at unit concentrations (1 mol/litre): For methanol $H_{\text{calc.}} = \log S_1 = 11.22$, and for ethanol $H_{\text{calc.}} = \log S_2 = 11.14$, where S is expressed in mol/litre per sec.

As radii of CH_3O^- and $\text{C}_2\text{H}_5\text{O}^-$ are used the radii of CH_3OH and $\text{C}_2\text{H}_5\text{OH}$, respectively. The following values were applied: $\sigma_{\text{CO}} = 1.6 \cdot 10^{-8}$ cm., $\sigma_{\text{CH}_3\text{OH}} = 2 \cdot 10^{-8}$ cm., $\sigma_{\text{C}_2\text{H}_5\text{OH}} = 1.8 \cdot 10^{-8}$ cm. $T = 318$, and $R = 8.32 \cdot 10^7 \frac{\text{erg.}}{\text{degr.}}$.

Hence, the experimentally found values for 10^H are not markedly different from the values calculated for the number of collisions (at unit concentrations) in the gas phase; this was found for many reactions in solution—cf., for example, LANDOLT BÖRNSTEIN'S tables (15)—and has been thoroughly discussed in the reaction-kinetic literature (TRAUTZ, LEWIS, and later authors).

Experiments 25 and 26 were performed in order to obtain information regarding the influence of the alcohol concentration on the velocity of the reaction. The initial solution in experiment 26 was 0.5 molar CH_3ONa in a mixture of equal parts of CH_3OH and dioxane; in experiment 25, 0.5 molar $\text{C}_2\text{H}_5\text{ONa}$ in a mixture of equal parts of $\text{C}_2\text{H}_5\text{OH}$ and dioxane.

The results are given in tables 2 and 3. In the methanol experiment, the velocity constant was about 3.6 times that of the corresponding experiment without dioxane, and in the ethanol experiment it was about 2.9 times as great. However, the greater velocity thus observed can hardly be due to the smaller alcohol concentration, but must be ascribed to the fact that the reaction now takes place in another medium.

In order to find whether CO reacts with a sodium hydroxide solution at low temperatures and pressures, an experiment was made analogous to those described above, using an approximately 1 molar NaOH solution at 54.16° . k of the reaction was found to be $0.96 \cdot 10^{-5} \text{ sec.}^{-1}$, which, giving due consideration to the solubility of CO in H_2O and the factor $\frac{V_g}{V_v}$, may be converted

into $k' = 19.8 \cdot 10^{-4}$. The corresponding constant of the methanol experiment at 1 molar c_{cat} is $k' = 42.0 \cdot 10^{-4}$. The velocity in water is thus about 2.1 times as slow as in methanol.

Our thanks are due to N. HOFFMANN-BANG, civil engineer, for valuable participation in the experiments.

Summary.

1) The reaction between CO and CH_3OH , catalyzed by CH_3ONa , is investigated manometrically. HCOOCH_3 is formed. The reaction is found to be of the first order with respect to CO.

Its rate is shown to be determined by the reaction



the complex formed reacting to give HCOOCH_3 and CH_3O^- which thus acts as a catalyst.

2) The velocity constant increases somewhat faster than proportional to the concentration of the catalyst, and is approximately described by the following equation

$$\log \frac{k}{c_{\text{cat}}} = \alpha + \beta \cdot c_{\text{cat}}. \quad (1)$$

3) Experiments with CO and $\text{C}_2\text{H}_5\text{ONa}$ dissolved in $\text{C}_2\text{H}_5\text{OH}$ proceed analogously, and here too, the relation between velocity constant and concentration is approximately described by equation (1). The velocity constant is about 3 times as great as for the reaction in CH_3OH .

4) Both series of experiments are carried out at approximately 35° , 45° , and 54° . For the reactions in CH_3OH and $\text{C}_2\text{H}_5\text{OH}$, the energies of activation are both about 18500 small calories, respectively, and the frequency factors about $10^{9.9}$ and $10^{10.4}$, respectively.

5) The reaction proceeds more rapidly if the solution contains dioxane.

6) The velocity constant for a reaction between CO and a solution of sodium hydroxide in water is of the same order of magnitude as the constants for the above reactions. However, the observed velocity is about 45 times smaller, mainly because CO is less soluble in water than in alcohol.

Chemical department A.

Royal Technical College of Denmark.

Copenhagen, July 1942.

Bibliography.

- (1) ARTHUR STÄHLER: Über die Addition von Kohlenoxyd an Alkohole. Ber. **47** (1914) 580.
- (2) J. A. CHRISTIANSEN: The Equilibrium between Methyl Formate and Methyl Alcohol, and Some Related Equilibria. J. Chem. Soc. (1926) 413.
- (3) L. MOSER: Die Reindarstellung von Gasen. 1920.
- (4) J. A. CHRISTIANSEN und JOHN R. HUFFMAN: Über eine modifizierte Mikrogasbürette. Z. anal. Chem. **80** (1930) 435.
- (5) HAKON LUND und JANNIK BJERRUM: Eine einfache Methode zur Darstellung wasserfreier Alkohole. Ber. **64** (1931) 210.
- (6) S. YOUNG: The Vapour-Pressures, Specific Volumes, Heats of Vaporation, and Critical Constants of Thirty Pure Substances. Scient. Proc. of Royal Dublin Soc. **12** (1910) 374.
- (7) KURT HESS und HERMANN FRAHM: Über Darstellung und Eigenschaften von reinem Dioxan. Ber. **71** (1938) 2627.
- (8) E. A. GUGGENHEIM: On the Determination of the Velocity Constant of a Unimolecular Reaction. Phil. Mag. [2] **7** (1926) 538.
- (9) J. A. CHRISTIANSEN: Über die Berechnung von Geschwindigkeitskonstanten aus Versuchsdaten. Z. physik. Chem. (A) **189** (1941) 126.
- (10) A. RICHARDSON: Determination of Vapour-Pressures of Alcohols and Organic Acids. J. Chem. Soc. **49** (1886) 761.
- (11) P. M. LEININGER and M. KILPATRICK: The Inversion of Sucrose. J. Am. Chem. Soc. **60** (1938) 2891.
- (12) L. WILHELMY: Über das Gesetz, nach welchem die Einwirkung der Säuren auf den Rohrzucker stattfindet. Poggendorffs Annalen **81** (1850) 499.
- (13) A. GEUTHER: Über die Einwirkung des Aethernatrons auf die Aether einiger Kohlenstoffsäuren. Chem. Zentralblatt 1868, 630.
- (14) G. JUST: Löslichkeit von Gasen in organischen Lösungsmitteln. Z. physik. Chem. **37** (1901) 361.
- (15) LANDOLT-BÖRNSTEIN: Physikalisch-chemische Tabellen. 2. Ergänzungsband p. 1622, and 3. Ergänzungsband p. 2883.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 4

ON THE VELOCITY OF UNIMOLECULAR REACTIONS

BY

J. A. CHRISTIANSEN



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1942

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

1. It is well known that the explanation of reactions which follow the unimolecular law (reactions of the first order) involves theoretical difficulties concerning the necessary rate of transfer of the activation energy. These difficulties seem to have been overcome mainly by O. K. RICE and H. C. RAMSPERGER¹, R. H. FOWLER and E. K. RIDEAL², and C. N. HINSHELWOOD³. Further contributions to the elucidation of these problems were made by G. N. LEWIS and D. F. SMITH⁴, and by the present author⁵. However, it should not be forgotten that already several years earlier M. POLANYI⁶ applied a quite similar treatment of the problem. In the present paper, we shall not be concerned with these difficulties, as it will be assumed that the reacting molecules are dissolved in a sufficiently dense solvent, so that the necessary transfer of energy to the reacting molecules occurs by collisions with the solvent molecules. This way of avoiding the theoretical difficulties by means of realisable experimental conditions was proposed many years ago by F. A. LINDEMANN⁷. The necessary assumptions may be expressed as follows. All molecules of the same kind, the reacting ones included, have the same activity, i. e., the number of molecules in a given state of energy multiplied by the reciprocal Boltzmann-factor is independent of this state.

2. The problem investigated here is mainly that of the absolute values of the velocity constants but moreover, other sides of the velocity-problem will be treated.

¹ J. Am. Chem. Soc. **49** (1927) 1617.

² Proc. Roy. Soc. **113 A** (1926) 570.

³ Proc. Roy. Soc. **113 A** (1926) 530.

⁴ J. Am. Chem. Soc. **47** (1925) 1514.

⁵ Proc. Cambridge Phil. Soc. **23** (1926) 438.

⁶ Zs. f. Phys. **1** (1920) 341.

⁷ Trans. Far. Soc. **17** (1922) 599.

As is well known, the constants may be expressed as a product of two factors, one of which is almost independent of the temperature and may conveniently be called the frequency factor (f), and the other which has the exponential form $e^{\frac{-Q}{RT}}$. Both factors can be determined by experiments at different temperatures, the factor f , however, only with very moderate accuracy. This has always been a great difficulty for the theories, since, at least in many cases, any theory which gives only the right order of magnitude of f can be used to describe the experiments.

The experiments show that f is generally of the order of magnitude $10^{+13} - 10^{+14}$ reciprocal seconds, a value which roughly corresponds to the frequencies of molecular oscillations. On the other hand, this value is not very different from $\frac{RT}{h}$, where R is the gas constant per molecule. With $T = 300^\circ$, $RN = 8.316 \cdot 10^7$ erg/ $^\circ\text{C}$, $N = 6.06 \cdot 10^{23}$, $h = 6.55 \cdot 10^{-27}$ we get $\frac{RT}{h} = 6.28 \cdot 10^{12}$. Incidentally, it is a rather old observation that f is of this order of magnitude (cf. for instance K. F. HERZFELD¹, M. TRAUTZ², and S. DUSHMANN³).

One further result of the experiments may be mentioned here. Investigations of groups of analogous reactions seem to indicate that a correlation between f and Q exists, viz. f is almost proportional to $e^{\frac{+dQ}{r}}$ ⁴, where r is a positive constant of the order of magnitude RT , and dQ is the deviation of Q from some common value. However, it is doubtful in some cases whether this correlation is real, since an inaccuracy dQ in the determination of Q from a given series of experiments necessarily introduces an erroneous factor $e^{\frac{+dQ}{RT_1}}$ into f . T_1 is the mean temperature of the respective series of experiments. In view of the experimental evidence it seems difficult, however, completely to deny the existence of a correlation between Q and f . Finally, Q depends in some cases on T ⁵. Obviously, such a dependency

¹ Ann. d. Phys. **59** (1919) 613.

² Zs. f. Physik **2** (1920) 113.

³ J. Am. Chem. Soc. **43** (1921) 397.

⁴ J. K. SYRKIN: Z. anorg. allg. Chem. **199** (1931) 28. G.-M. SCHWAB, Z. physik. Chem. **B 5** (1929) 406. Comp. the discussion by M. G. EVANS and M. POLANYI. Trans. Far. Soc. **32** (1936) 1334.

⁵ Cf. f. inst. V. K. LA MER. J. chem. Phys. **1** (1933) 289. P. M. LEININGER and

is not at all surprising, since, in analogy with the equilibrium case, a heat capacity of the reacting molecules different from that of the normal molecules might very well be assumed. Many years ago, F. E. C. SCHEFFER¹ showed that the influence of this heat capacity effect, when it is only a few times R , on the velocity constants in dependence on temperature could not be detected, not even in experiments accurate to about 1 % and extending over the experimentally accessible temperature interval of 40 to 50 degrees. It is therefore astonishing that the existence of the effect has actually been proved beyond doubt in several cases; the corresponding "difference in heat capacity" was calculated from the experiments to as much as 40 to 50 times R ; regarding the relatively simple molecules investigated, these values are incredibly high.

In view of all these difficulties, a theory which is suited to meet the needs of the experimenter would be very useful. Numerous authors have treated this subject² and also the present author has made an attempt to give some contributions to this end³. These papers must to some extent be considered tentative and they are unsatisfactory in some respects, so that it was deemed necessary to review the whole question.

It is the trend of the present paper to look for solutions of the velocity problem on the basis of such knowledge which may be considered reasonably accurate. This should comprise:

- 1) The value of Q known from experiments with a reasonable degree of accuracy.
- 2) The force function for the atom displaced inside the molecule during the reaction in the field arising from the rest of the molecule. However, we shall only use the force function valid in the vicinity to the stable equilibrium points, and we shall endeavour to avoid a detailed formulation of the force function in the proximity of the unstable equilibrium point, cf. B in Fig. 1. The numerical value of the potential energy at this point is obviously given by Q .

M. KILPATRICK, J. Am. Chem. Soc. (1938) 2891. E. A. MOELWYN HUGHES, Proc. Roy. Soc. London **A 164** (1938) 295.

¹ Koninkl. Akad. Wetensch. Amsterdam **19** (1917) 636.

² H. EYRING, M. POLANYI, E. WIGNER: For references, cf. Trans. Far. Soc. **34** (1938), J. chem. Phys. **7** (1939).

³ Z. physik. Chem. **B 33** (1936) 145, **37** (1937) 374, **40** (1938) 183.

- 3) The knowledge of the force function in the vicinity to the unstable equilibrium point will be replaced by an assumption concerning the definition of the surface in the space of configuration¹ (e. g. in ordinary space) which separates the two regions belonging to the two possible types of the molecule.

To make the essential features appear more clearly, the actual calculation will be restricted to unidimensional reactions, the mathematical difficulties thus being reduced to a minimum. In this case, the "surface" mentioned above is reduced to a point on a line which then represents the "space".

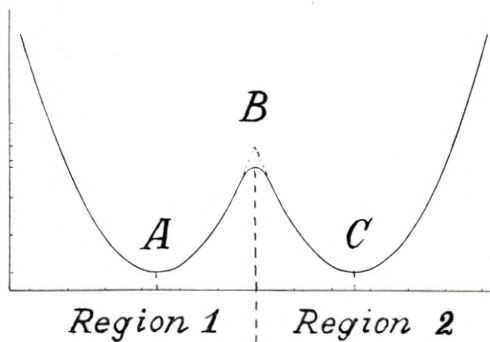


Fig. 1.

3. In recent years, it has become the general view that the "normal" value of f is $\frac{RT}{h}$; this view has been shared by the present author. However, this assumption would be in contradiction to the classical theory, as then the equilibrium condition between two isomeric molecules would be $K = e^{-\frac{\epsilon_1 - \epsilon_2}{RT}}$, where $\epsilon_1 - \epsilon_2$ is the energy difference between two isomeric molecules. According to classical statistical mechanics this does not hold, at least not in the simple case we are going to discuss.

We consider a model of a pair of isomeric molecules consisting of a unidimensional field of force in which a particle of a mass m is moving. The field of force is supposed to have the form of two parabolas as indicated in Figure 1. The two parabolas may not have the same parameter, i. e. we assume that

¹ Cf. J. chem. Physics 7 (1939) 23, and p. 12 of this paper.

the force acting on the particle is $-\alpha_1 x_1$ in the region 1 and $-\alpha_2 x_2$ in the region 2, where α_1 and α_2 are the force constants and x_1 and x_2 the distances from the respective equilibrium points A and C . In order to describe the state of the particle in the classical way we need two coordinates, hence in the present case simply one coordinate in ordinary space and one which specifies the momentum of the particle. If these coordinates are denoted as q and p , respectively, we get for the energy in the region 1

$$\varepsilon = \frac{1}{2} \alpha_1 q^2 + \frac{1}{2m} p^2 + \text{const.} \tag{1}$$

According to GIBB's statistical mechanics, the phase density of the probability P in our two-dimensional space is a function of the energy alone and has the form $A \cdot e^{\frac{-\varepsilon}{RT}}$. Of course, we may put the potential energy at the boundary between the two regions equal to zero and, consequently, the phase density of particles in rest at this boundary equal to A , which constant thus is common to both regions.

The density in phase becomes accordingly

$$P = A \cdot e^{\frac{\varepsilon_1}{RT}} \cdot e^{\frac{-\frac{1}{2}(\alpha_1 q^2 + \frac{1}{m} p^2)}{RT}}, \tag{2}$$

where ε_1 is the height from the vertex of parabola 1 to the maximum point, as we have $\varepsilon = -\varepsilon_1 + \frac{1}{2} \alpha q^2 + \frac{p^2}{2m}$ (1 a).

The total number of particles in region 1 is thus

$$N_1 = A \cdot e^{\frac{\varepsilon_1}{RT}} \iint e^{\frac{-\frac{1}{2}(\alpha_1 q^2 + \frac{1}{m} p^2)}{RT}} dq dp. \tag{3}$$

The integration with respect to p has to be performed from $p = -\infty$ to $p = +\infty$, and thus we get

$$N_1 = A \cdot e^{\frac{\varepsilon_1}{RT}} \sqrt{2mRT\pi} \int e^{-\frac{1}{2} \alpha_1 q^2} dq. \tag{4}$$

Here, the integration must be performed from $q = -\infty$ to a positive value of q given by the equation $\varepsilon_1 = \frac{1}{2} \alpha_1 q^2$. If, however, ε_1 is

some times (f. inst. 10 times) greater than RT , it is admissible to choose the upper limit positively infinite; in this way we obtain

$$N_1 = A \cdot e^{\frac{\varepsilon_1}{RT}} \sqrt{2mRT\pi} \sqrt{\frac{2}{\alpha_1} RT\pi}. \quad (5)$$

However, $\frac{m}{\alpha_1} = \frac{1}{4\pi^2\nu_1^2}$ (6), where ν_1 is the frequency of the oscillator in the region 1 (which is independent of the energy) which leads to

$$N_1 = A \cdot e^{\frac{\varepsilon_1}{RT}} \frac{RT}{\nu_1}. \quad (5_1)$$

In exactly the same way we get N_2 .

$$N_2 = A \cdot e^{\frac{\varepsilon_2}{RT}} \frac{RT}{\nu_2} \quad (5_2)$$

(where A is the same constant as above while ν_2 and ε_2 may be different from ν_1 and ε_1), and the equilibrium constant becomes

$$\frac{N_1}{N_2} = \frac{\nu_2}{\nu_1} e^{\frac{\varepsilon_1 - \varepsilon_2}{RT}}. \quad (6)$$

If k_1 is the velocity constant of the reaction $1 \rightarrow 2$, and k_{-1} is the constant of the reverse reaction $2 \rightarrow 1$, we have, as usual,

$$\frac{k_1}{k_{-1}} = \frac{N_2}{N_1} \quad \text{or} \quad \frac{k_1}{k_{-1}} = \frac{\nu_1 e^{\frac{-\varepsilon_1}{RT}}}{\nu_2 e^{\frac{-\varepsilon_2}{RT}}}. \quad (7)$$

Hence, the classical expression for the velocity of a reaction in the harmonic case must contain the factor ν , the frequency of the oscillation in question.

4. It is now the question whether these considerations may also be extended to the case where an equilibrium between the two forms is lacking. This means that we must try to evaluate the flow of particles at the boundary which separates the two forms of molecules from each other. We consider again the case of quasi-elastic oscillations around one of two centers.

Statistical mechanics then offers an expression for the flow across a boundary q_1 , given by $\frac{1}{2} \alpha_1 q_1^2 = \epsilon_1$. In the equilibrium case, the flow will be

$$s = \int P \dot{q} dp, \text{ where } \dot{q} \equiv \frac{dq}{dt}$$

and

$$P = A \cdot e^{\frac{\epsilon_1}{RT}} \cdot e^{-\frac{1}{2} \alpha_1 q_1^2} \cdot e^{-\frac{1}{2m} \frac{p^2}{RT}}.$$

According to the canonical equation, $\dot{q} = \frac{\partial \epsilon}{\partial p} = \frac{\partial \epsilon_p}{\partial p}$, and we thus get the total flow for positive values of p

$$s_+ = A \int_0^\infty e^{-\frac{1}{2m} \frac{p^2}{RT}} d\epsilon_p = ART \quad \text{or} \quad s_+ = N_1 e^{\frac{-\epsilon_1}{RT}} \cdot \nu_1. \quad (8)$$

The flow in the opposite direction will be of exactly the same value but of a negative sign, and the resultant flow thus becomes zero. If it was permissible to suppose that the reaction in the direction from left to right occurs with the velocity s_+ and that in the opposite direction with the same velocity, the above expression would be the solution of the problem. Unfortunately, however, this is not the case, as we have to assume that the motion of the particle is disturbed by collisions with the solvent molecules. In this case we must treat the problem as a diffusion problem, and our expression for the flow must contain a diffusion coefficient which is difficult to evaluate.

5. We shall, however, attack the problem just from this point of view. Recently, the author showed that a natural extension of the Arrhenius picture concerning the mechanism of a chemical reaction would lead to a process in which particles diffuse inside the molecule (or complex of molecules) in the intramolecular field of force¹. When applied to actual problems, this picture of the mechanism is not very different, if at all, from that known as the transition state method developed by H. EYRING, M. POLANYI, E. WIGNER and others². However, I consider it an advantage to state explicitly that the problem will be treated as a diffusion problem.

¹ Cf. ref. 3) p. 5, and H. A. KRAMERS, *Physica* **7** (1940) 284.

² Cf. ref. 2) p. 5.

6. It can easily be shown that EINSTEIN's expression for the intensity of flow s in a unidimensional field of force may be written

$$s = -D \varphi^{-1} \frac{d}{dx} c \varphi, \quad (9)$$

where D is the diffusion coefficient which we assume to be constant in time and space. c is the linear density of the points representing the position of the particle on the x -axis and φ is defined by $\ln \varphi = V$, where V is the potential energy $U(x)$ divided by RT . This leads to

$$s = -D \frac{dc}{dx} + c \frac{D}{RT} K_x$$

which is the well known EINSTEIN expression for the diffusion in a force-field.

From the equation expressing the conservation of the number of particles

$$-s' = \dot{c} \quad (10)$$

and (9) we get by substitution of $c \varphi = y$

$$D(y'' - V' y') = \dot{y}, \quad (10 \text{ a})$$

as φ is independent of time.

7. In reaction kinetics we are only interested in reactions which are so slow that y with a very high approximation can be considered stationary in time. This can easily be seen in the special case $V = \frac{x^2}{4a} = \frac{\alpha x^2}{2RT}$, i. e. the harmonic oscillator. In this case, our equation becomes

$$2ay'' - xy' = \frac{2a\dot{y}}{D}, \quad (11)$$

and may be solved as follows: We look for a solution which obeys the differential equation $\dot{y} = -ky$, indicating that the linear point density decreases with time everywhere in the same proportion. k is the ordinary unimolecular velocity constant. Thus, the partial differential equation (11) becomes an ordinary differential equation

$$2ay'' - xy' = -y \frac{2ak}{D}, \quad (12)$$

where the pure number $\frac{2ak}{D}$ is exceedingly small as compared to 1. The estimation is as follows: Per definitionem $2a = \frac{RT}{\alpha}$; at room temperatures and the usual values of α this is of the order of magnitude 10^{-18} cm². We know almost nothing about D , but as a reasonable assumption we might choose the values found for ordinary diffusion coefficients in liquids; in the case of water, this value happens to be of the same order of magnitude as $\frac{h}{4\pi m}$, where h is PLANCK'S constant and m the mass of the molecule. Hence, we consider D to be of the order of magnitude 10^{-4} cm²/sec. For reactions with measurable velocities, k must be less than 1 sec⁻¹, and the order of magnitude of $\frac{2ak}{D}$ thus becomes 10^{-14} or less.

The exact solution of the equation

$$2ay'' - xy' + \nu y = 0,$$

where ν is a fraction, was given by N. NIELSEN¹; for practical purposes, however, this solution seems too laborious. It is easier and sufficiently accurate to use an approximation which we obtain in the following way:

In accordance with the foregoing remarks, the solutions of (11) must differ but slightly from the solutions of

$$2ay'' - xy' = 0 \tag{13}$$

which are $y = A = \text{const.}$ (in space) and $y = \int_0^x \varphi dx$, where $\ln \varphi = \frac{x^2}{4a}$. However, when no reaction occurs, the distribution must evidently be given by $y = A$ which expresses the "barometric formula" $c = A\varphi^{-1}$. The solution of (11) must therefore be $y = A + \eta$ which, by multiplication with φ^{-1} and integration, gives the total number of particles $N = A \int_{-\infty}^l \varphi^{-1} dx$, since the integral $\int_{-\infty}^l \eta \varphi^{-1} dx$ must be small as compared to N .

Correspondingly, we may assume $\dot{y} = B + \dot{\eta}$ and $\dot{N} = B \int_{-\infty}^l \varphi^{-1} dx$, since the integral $\int_{-\infty}^l \dot{\eta} \varphi^{-1} dx$ must be small as compared to \dot{N} . Disregarding $\dot{\eta}$ as compared to B we therefore get from (11)

¹ D. Kgl. Danske Vidensk. Selsk. Math.-fys. Medd. I, 6 p. 61 (1918).

$$2a\eta'' - x\eta' = B\frac{2a}{D} \quad (14)$$

from which by "variation of the constants"

$$\eta = -\frac{B}{D} \left[\int^x J\varphi^{-1} dx - J \int^x \varphi^{-1} dx \right]. \quad (15)$$

Here, $y_1 = 1$, $y_2 = J = \int_0^x \varphi dx$ and $y_1 y_2' - y_1' y_2 = \varphi$, respectively, are the two particular solutions of (13) and their differential-determinant.

In order to employ these expressions for the determination of $\frac{\dot{N}}{N} = -k$, we introduce a definition of the point $x = l$, which separates the regions (on the x -axis) corresponding to the two forms of the molecule (the transition point): In one region c (or y) decreases everywhere with time, and in the other one c increases everywhere with time. Consequently, at $x = l$, c must remain constant; on account of the time-dependent constants A and B , this is only possible if c , and thus y , are zero at that point. At $x = l$ we get therefore

$$0 = y = A - \frac{B}{D} \left[\int_{-\infty}^l J\varphi^{-1} dx - J \int_{-\infty}^l \varphi^{-1} dx \right], \quad (16)$$

where the lower limit of integration is now determined by the condition that the flow s defined by $-D\varphi^{-1}y'$ must be zero in the negative infinite; at $x = l$ it becomes $-B \int_{-\infty}^l \varphi^{-1} dx$.

Since J is an odd function of x and $J\varphi^{-1} \rightarrow \frac{2a}{x}$ for great values of x^1 , the second term on the right side must cancel as compared to the third one, and we get with good approximation

$$k = -\frac{B}{A} = -\frac{\dot{N}}{N} = \frac{D}{J\sqrt{4\pi a}}, \quad (17)$$

as $\int_{-\infty}^l \varphi^{-1} dx \rightarrow \sqrt{4\pi a}$ when $\ln \varphi = \frac{l^2}{4a}$ is only moderately high.

¹ T. J. STIELTJES: Acta mathematica 9 (1887), 157.

As said above J approximates $\frac{2a}{l} \cdot \frac{l^2}{e^{4a}}$ for high values of $\frac{l^2}{4a} = \frac{Q}{RT}$. It seems, however, that this solution cannot be the correct one, because if it were applied to two reciprocal reactions it would lead to an expression for the equilibrium constant $K = \frac{k_1}{k_{-1}}$

$$K = \sqrt{\frac{a_2}{a_1} \frac{J_2(l_2)}{J_1(l_1)}} = \frac{\nu_1}{\nu_2} \frac{J_2(l_2)}{J_1(l_1)}$$

which is not identical with the classical value found on p. 6 eq. (7).

8. Therefore, we must consider the problem of two opposing reactions more closely.

If we have two types of molecules (1,2) transforming into each other, the resulting velocity becomes, as well known,

$$s = -\frac{dN_1}{dt} = k_1 N_1 - k_{-1} N_2 \quad (18)$$

from which we obtain by integration

$$\left. \begin{aligned} N_1 &= k_{-1} \frac{N}{k} + C \cdot e^{-kt} & (k \equiv k_1 + k_{-1}) \\ N_2 &= k_1 \frac{N}{k} - C \cdot e^{-kt}. \end{aligned} \right\} \quad (19)$$

This means obviously that we have to look for solutions of (11) of the types

$$\left. \begin{aligned} y_1 &= A + B_1 + \eta_1 \\ y_2 &= A + B_2 + \eta_2 \end{aligned} \right\}, \quad (20)$$

where A is constant in time and space, while B_1, B_2 which are constant in space and η_1, η_2 contain the common factor e^{-kt} . It is now the question, how to define the total "concentration" (number of particles per cm) or the density of probability c and the corresponding y , so that both c and s are continuous at the point of separation. The simplest procedure seems to be the following: We assume everywhere

$$c \varphi = y = A + \varphi [(B_1 + \eta_1) \varphi_1^{-1} + (B_2 + \eta_2) \varphi_2^{-1}], \quad (21)$$

where φ is the true value of φ everywhere, while φ_1 and φ_2 are defined as the functions valid in the two different regions corresponding to the two types of the molecule. In the harmonic case, $\ln \varphi_1 = \frac{x_1^2}{4a_1}$, $\ln \varphi_2 = \frac{x_2^2}{4a_2}$, where x_1 and x_2 are the distances measured from the points A and C respectively on the potential curve (cf. Fig. 1).

It is supposed that φ coincides with φ_1 or with φ_2 in the main parts of the regions corresponding to the two types of the molecule, so that the expressions given below for the number of molecules of the two types remain essentially correct. Around the transition point itself we assume a difference, since here $\varphi' = 0$ while φ'_1 and φ'_2 are different from 0. In order to preserve continuity, we assume $\varphi_1 = \varphi_2$ at the same point; without further loss of generality we may place both equal to 1 at the same point. Thus, the number N_1 of molecules of type 1 becomes with good approximation

$$N_1 = \int_{-\infty}^{l_1} c \, dx = (A + B_1) \int_{-\infty}^{l_1} \varphi_1^{-1} dx \equiv (A + B_1) P_1 \quad (22 \text{ a})$$

as we may assume $\int_{-\infty}^{l_1} \eta_1 \varphi_1^{-1} dx$ and $B_2 \int_{-\infty}^{l_1} \varphi_2^{-1} dx$ to be vanishingly small. Analogously,

$$N_2 = (A + B_2) \int_{-l_2}^{+\infty} \varphi_2^{-1} dx \equiv (A + B_2) P_2. \quad (22 \text{ b})$$

To calculate \dot{N} , we consider the flow s at the point of transition where it is equal to $-\dot{N}$. Generally, we have $s = -D \varphi^{-1} y'$; y' at the mentioned point becomes

$$y' = \varphi \frac{d}{dx} [(B_1 + \eta_1) \varphi_1^{-1} + (B_2 + \eta_2) \varphi_2^{-1}]. \quad (23)$$

Furthermore, we shall introduce two conditions analogous to that used in the case of a onesided reaction viz.

$$B_1 + \eta_1 = 0; \quad B_2 + \eta_2 = 0, \quad (\text{at } x_1 = l_1, x_1 = -l_2); \quad (24)$$

hence, we get

$$-\dot{N}_1 = s = -D [\eta'_1 \varphi_1^{-1} + \eta'_2 \varphi_2^{-1}]. \quad (25)$$

Just as in the former case, we now assume approximately

$$\dot{y}_1 = C_1; \quad \dot{y}_2 = C_2, \tag{26}$$

where C_1 and C_2 both contain the factor e^{-kt} but are constant in space.

We thus get

$$\left. \begin{aligned} 2 a_1 \eta_1'' - x_1 \eta_1' &= \frac{2 a_1}{D} C_1 \\ 2 a_2 \eta_2'' - x_2 \eta_2' &= \frac{2 a_2}{D} C_2 \end{aligned} \right\} \tag{27}$$

from which

$$\left. \begin{aligned} \eta_1 &= -\frac{C_1}{D} \left[\int_{-\infty}^{x_1} \varphi_1^{-1} J_1 dx - J_1 \int_{-\infty}^{x_1} \varphi_1^{-1} dx \right]; \quad \eta_1' = \frac{C_1}{D} \varphi_1 \int_{-\infty}^{x_1} \varphi_1^{-1} dx \\ \eta_2 &= -\frac{C_2}{D} \left[\int_{+\infty}^{x_2} \varphi_2^{-1} J_2 dx - J_2 \int_{+\infty}^{x_2} \varphi_2^{-1} dx \right]; \quad \eta_2' = \frac{C_2}{D} \varphi_2 \int_{+\infty}^{x_2} \varphi_2^{-1} dx \end{aligned} \right\} \tag{28}$$

$-\dot{N}_1 = s$ at the transition point $x_1 = l_1$; $x_2 = -l_2$ thus becomes

$$\begin{aligned} s &= -C_1 \int_{-\infty}^{l_1} \varphi_1^{-1} dx - C_2 \int_{+\infty}^{-l_2} \varphi_2^{-1} dx = -C_1 \int_{-\infty}^{l_1} \varphi_1^{-1} dx + C_2 \int_{-l_2}^{+\infty} \varphi_2^{-1} dx \\ &= -C_1 P_1 + C_2 P_2. \end{aligned}$$

The condition that $\dot{y} = 0$ is fulfilled on account of the two conditions (24) introduced above. From these we get

$$\left. \begin{aligned} DB_1 + C_1 J_1(l_1) P_1 &= 0 \\ DB_2 + C_2 J_2(-l_2)(-P_2) = DB_2 + C_2 J_2(l_2) P_2 &= 0. \end{aligned} \right\} \tag{29}$$

Thus

$$s = D \left(\frac{B_1}{J_1} - \frac{B_2}{J_2} \right). \tag{30}$$

From the equations (22)

$$\begin{aligned} N_1 &= (A + B_1) P_1 \\ N_2 &= (A + B_2) P_2 \end{aligned}$$

we get by addition

$$N_1 + N_2 = N = A(P_1 + P_2) \tag{31}$$

as the total number of molecules must be independent of time, and the B 's are time dependent. Therefore,

$$B_1 = \frac{1}{P_1} \frac{N_1 P_2 - N_2 P_1}{P_1 + P_2}$$

$$B_2 = -\frac{1}{P_2} \frac{N_1 P_2 - N_2 P_1}{P_1 + P_2}.$$

Substitution in (30) gives

$$s = -\dot{N}_1 = D \frac{N_1 P_2 - N_2 P_1}{P_1 + P_2} \left(\frac{1}{P_1 J_1} + \frac{1}{P_2 J_2} \right). \quad (32)$$

By comparison with

$$-\dot{N}_1 = k_1 N_1 - k_{-1} N_2 \quad (33)$$

we get

$$k_1 = D \left(\frac{1}{P_1 J_1} + \frac{1}{P_2 J_2} \right) \frac{P_2}{P_1 + P_2}$$

$$k_{-1} = D \left(\frac{1}{P_1 J_1} + \frac{1}{P_2 J_2} \right) \frac{P_1}{P_1 + P_2}$$

and

$$k = k_1 + k_{-1} = D \left(\frac{1}{P_1 J_1} + \frac{1}{P_2 J_2} \right), \quad (34)$$

while

$$\frac{k_1}{k_{-1}} = \frac{P_2}{P_1} = K. \quad (35)$$

According to their definitions, the products PJ in the expression for k are independent of the choice of the zero point for the potential energy, and they are equal to $\sqrt{4\pi a} \int_0^l e^{\frac{x^2}{4a}} dx = \sqrt{4\pi a} \cdot j$. In experiments on a reaction of the first order we actually measure $k = k_1 + k_{-1}$, and the expression (34) is thus generally valid. The value of the equilibrium constant is identical with the classical expression, as may be seen from the following.

As we have made $\varphi_1 = \varphi_2 = 1$ at the point of transition

$$\ln \varphi_1 = \frac{x_1^2}{4a_1} - V_1 \quad \text{where} \quad V_1 = \frac{Q_1}{RT} = \frac{l_1^2}{4a_1}$$

and

$$\ln \varphi_2 = \frac{x_2^2}{4a_2} - V_2 \quad \text{where} \quad V_2 = \frac{Q_2}{RT} = \frac{l_2^2}{4a_2}.$$

Thus $K = \frac{P_2}{P_1} = \sqrt{\frac{a_2}{a_1}} \cdot e^{V_2 - V_1} = \frac{\nu_1}{\nu_2} \cdot e^{-(V_1 - V_2)}$ which is identical with the classical expression.

9. In the numerous cases in which the reaction is practically irreversible, i. e. $K \gg 1$ or $V_2 > V_1$, j_2 will be large as compared to j_1 , and then we get the expression (17)

$$k = \frac{D}{\sqrt{4\pi a}} \cdot \frac{1}{j}.$$

In the special case of an experiment on racemisation $j_1 = j_2$ and $P_1 = P_2$ so that

$$k = 2k_1 = \frac{D}{\sqrt{\pi a}} \cdot \frac{1}{j}, \quad K = 1.$$

Using the approximation

$$j = \frac{2a}{l} \cdot e^{+\frac{l^2}{4a}}$$

we get in the former case

$$k = \frac{D}{2a} \cdot \frac{l}{\sqrt{4\pi a}} \cdot e^{-\frac{l^2}{4a}}.$$

By means of $\frac{\alpha}{m} = 4\pi^2\nu^2$ and $2a = \frac{RT}{\alpha}$ we may introduce ν and get

$$k = \frac{4\pi mD}{RT} \cdot \nu^2 \sqrt{\pi} \sqrt{\frac{Q}{RT}} \cdot e^{-\frac{Q}{RT}}.$$

If it were literally true that $4\pi mD = h$, which at least is not impossible as D signifies an intramolecular diffusion constant, the frequency factor f would be $\nu \sqrt{\pi} \cdot \frac{h\nu}{RT} \cdot \sqrt{\frac{Q}{RT}}$, hence of the same order of magnitude as ν , i. e. about 10^{14} and this is the order of magnitude of f found empirically in numerous cases. It is also seen that this factor depends on T and on Q , however, their powers are far too small to account for the dependence found empirically.

10. In this connection it should be mentioned that in the case of strong interaction between the displaced particle and its surroundings, the "potential energy" must be replaced by "the free energy" given as a function of the position of the particle. Thus, it becomes possible to understand that thermodynamic relations can successfully be applied to problems in kinetics,

as shown repeatedly by M. POLANYI and his school, apart from several earlier authors.

So we may say that considerations based on classical statistical mechanics in connection with the assumption of EINSTEIN's equation for diffusion in a force field, which may also be called "classical", lead to an expression which contains the unknown quantity D , but otherwise seems to be reasonably consistent with the experimental facts. It should be emphasized that H. A. KRAMERS¹ discussed the application of the diffusion theory to the problem of chemical reactions more generally. KRAMERS came to the following result: if it is assumed that a "damping" of the oscillations arising from the interaction between the solute and the solvent molecules is very strong, which means that the acceleration forces may be neglected, EINSTEIN's diffusion equation may be applied. Actually this assumption was our starting point.

For the evaluation of the diffusion constant, one might of course resort to a procedure similar to that used in the kinetic theory of gases. We want a mechanism which is sufficiently effective to maintain the normal thermal distribution. Such a mechanism would demand the assumption of very numerous collisions as present in liquids, and the problem can hardly be solved with any reasonable degree of accuracy. Furthermore, one would be inclined in this case to expect a rather narrow parallelism between the viscosity of the solvent and the velocity of unimolecular reactions; however, such a parallelism does not seem to exist. It may be mentioned in this connection that in a few cases the same reaction has been studied as gas-reaction and in different solvents, but no appreciable difference was found between the frequency factors obtained with different solvents.

My best thanks are due to Dr. CHR. MØLLER at the Institute for Theoretical Physics for several illuminative discussions on the above subject.

¹ Cf. ref. 1) p. 8.

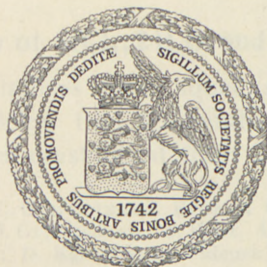
University and Royal Institute of Technology, Copenhagen, July 1942.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 5

THE STRUCTURE OF $\text{SrCl}_2, 2\text{H}_2\text{O}$

BY

AKSEL TOVBORG JENSEN



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1942

INTRODUCTION

The salt hydrates in which both cation and anion are noble gas ions would form the best basis for a simple discussion of the chemical forces operating in solid hydrates. So far, only a few of such hydrates have been completely investigated by the methods of X-ray crystallography. The structures of the following compounds seem to be well established: $\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$ ¹ with which $\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$ is isomorphous, $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ and the isomorphous bromide², $\text{SrCl}_2 \cdot 6\text{H}_2\text{O}$, and a long series of isomorphous Ca, Sr and Ba halide hexahydrates³, the tetragonal mixed crystal of carnallite $\text{KMgCl}_3 \cdot 6\text{H}_2\text{O}$, and brom-carnallite $\text{KMgBr} \cdot 6\text{H}_2\text{O}$ ⁴. In these hexahydrates the bi- or trivalent cations are completely surrounded by water molecules.

The atomic arrangement in a hydrate in which the number of water molecules per cation is so small that direct contact between cation and anion can be expected is of some interest. It shows whether the structure is in agreement with BERNAL and FOWLER's⁵ ideas concerning the coordination of water molecules, a theory which has previously been confirmed by the present author in the case of hexahydrates.

The Crystallography of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ and Related Compounds.

$\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ crystallizes from aqueous solutions at temperatures above 66.5° . Below that temperature, $\text{SrCl}_2 \cdot 6\text{H}_2\text{O}$ is the stable hydrate⁶. When crystallized from a hot aqueous solution

¹ K. R. ANDRESS and C. CARPENTER: *Z. Krist. A* **87** (1934) 446.

² K. R. ANDRESS and J. GUNDERMANN: *Z. Krist. A* **87** (1934) 345.

³ A. TOVBORG JENSEN: *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* XVII, 9 (1940).

⁴ K. R. ANDRESS and O. SAFFE: *Z. Krist. A* **101** (1939) 451.

⁵ D. BERNAL and R. H. FOWLER: *J. chem. Phys.* **1** (1933), 515.

⁶ TILDEN: *J. Chem. Soc.* **45** (1884), 269.

it appears according to the author's experience as thin fragile plates. According to WULFF and HEIGL¹, the crystals show double refraction, they are "biaxial and non-orthogonal". $\text{SrCl}_2, 2 \text{H}_2\text{O}$ may also be prepared by precipitating a saturated solution of $\text{SrCl}_2, 6 \text{H}_2\text{O}$ with concentrated HCl, or by dehydrating $\text{SrCl}_2, 6 \text{H}_2\text{O}$ at a suitable water vapour tension. Powder specimens were prepared in the two last mentioned ways at several temperatures. They gave identical powder diagrams (i. e. identical apart from minor differences in intensity due to the fact that $\text{SrCl}_2, 2 \text{H}_2\text{O}$ precipitated by HCl forms thin flaky crystals, while prepared by dehydration it shows no developed faces). The correct composition of all preparations was ascertained by titration with AgNO_3 . The powder diagrams could be indexed with the lattice constants found on single crystals prepared in the first mentioned way. Hence, all three methods of preparation led to identical hydrates.

$\text{SrCl}_2, 2 \text{H}_2\text{O}$ is neither isomorphous with $\text{CaCl}_2, 2 \text{H}_2\text{O}$ nor with $\text{BaBr}_2, 2 \text{H}_2\text{O}$, since these compounds showed totally different powder diagrams. Nor is it isomorphous with $\text{BaCl}_2, 2 \text{H}_2\text{O}$, as may be seen by comparing its cell dimensions with those of $\text{BaCl}_2, 2 \text{H}_2\text{O}$ (cf. NARÁY-SZABÓ and SASVÁRI²). Moreover, the mixed crystals of $\text{BaCl}_2, 2 \text{H}_2\text{O}$ and $\text{BaBr}_2, 2 \text{H}_2\text{O}$ are of a fifth type, as follows from the crystallographic data for $\text{BaBr}_2, 2 \text{H}_2\text{O}$ and $\text{Ba}(\text{Cl}, \text{Br})_2, 2 \text{H}_2\text{O}$ ³. Powder diagrams of these compounds were also taken and were found to lack any resemblance to each other.

Some attempts to prepare mixed crystals of $\text{SrCl}_2, 2 \text{H}_2\text{O}$ and $\text{BaCl}_2, 2 \text{H}_2\text{O}$ (LEHMANN'S „Kristalle ähnlich den wasserärmeren von Chlorbarium, doch weit schöner ausgebildet, somit wahrscheinlich isomorphe Mischungen”⁴), which might have been useful for the determination of the structure of $\text{BaCl}_2, 2 \text{H}_2\text{O}$ by Fourier methods proved quite unsuccessful. Crystallization by evaporation or cooling of solutions containing 5 mols of SrCl_2 to 1 mol of BaCl_2 always gave ordinary crystals of $\text{BaCl}_2, 2 \text{H}_2\text{O}$ which, in semi-quantitative analyses, were found to contain far less than 5 per cent of $\text{SrCl}_2, 2 \text{H}_2\text{O}$. It may be mentioned on

¹ Z. Krist. 77 (1931), 99.

² Z. Krist. 97 (1937), 235—37.

³ Gmelins Handb. 30, Ba, 225 and 235.

⁴ Z. Krist. 8 (1884), 451.

this occasion that LEHMANN described his experiments but briefly, and they seem, moreover, to have been made in a somewhat casual way. His observations are not completed by quantitative optical or chemical data; but they only inform us of what the microscope reveals.

From the information collected above it is obvious that the problem is of a perplexing diversity as compared, for instance, with the hexahydrates of the same alkaline earth halides.

Determination of Crystal System, Unit Cell, and Space Group of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$.

The single crystals of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ prepared by evaporation from a solution at 80° were thin plates and presented no other crystallographic elements. They were far from perfect, since they had to be broken off, removed from the mother liquor, and wiped with filter paper as rapidly as possible, because while cooling the adhering mother liquor solidifies to a mass of hexahydrate crystals.

Of a plate of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ 80 KV Laue photographs were taken with the primary beam perpendicular to the developed face. The diagram, plate I, Fig. 1, shows that a plane of symmetry is present in the crystal. Another diagram was taken with the primary beam perpendicular to the plane of symmetry found. The diagram, plate I, Fig. 2, shows a twofold axis, only. Hence, the crystal belongs to the monoclinic system with the b-(ortho)axis parallel to the developed face. A small fragment of the crystal was mounted on the goniometer head and was orientated with the b-axis as axis of rotation by aid of a number of Laue-photographs. From oscillation photographs, the length of the b-axis was found to be 6.44 \AA . The $h0l$ plane of the reciprocal lattice was constructed¹ by means of a Weissenberg diagram.

From the reciprocal net the crystal lattice dimensions were found to be $\beta = 72^\circ$, $b = 3.33 \text{ \AA}$, $a = 5.87 \text{ \AA}$. According to WULFF and HEIGL (loc. cit.) $d = 2.6715$. This gives $Z = 0.996$.

¹ The Bernal reciprocal lattice. Proc. Roy. Soc. A **113** (1926) 117.

The diagrams were taken with a Buerger-Weissenberg goniometer. Z. Krist. A **94** (1936) 87—99.

² The angles between the reciprocal axes were taken to be $> 90^\circ$. Hence, all indexes and fractional coordinates refer to a cell with $\beta = 72^\circ$.

A h1l Weissenberg photograph revealed that the a and c axes were twice the axes first obtained, and the number of mols in the true unit cell must be 4. The h1l and h2l nets were constructed from Weissenberg diagrams, and the following reflections were found to be present:

$$\left. \begin{array}{l} \text{in } h0l \quad h = 2n, l = 2n \\ h1l \quad h+1 = 2n+1 \\ h2l \quad h+1 = 2n \end{array} \right\} h+k+l = 2n.$$

With the axes chosen, the lattice is body-centered. The condition $h+k+l = 2n$ was also fulfilled in the $hk0$ and the $hk1$ Weissenberg diagrams, and in the powder diagrams. Cu-radiation was used for the Weissenberg diagrams, Cr-radiation for the powder diagrams. The space group has glide planes with the translations a , c , and n . Conventionally, a is included in its symbol, but not c or n , which with the axes mentioned above leads to the following orientated space groups for $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$:

$$\begin{array}{l} I2/a (C_{6h}^2) \text{ which is centrosymmetrical and} \\ Ia (C_s^4) \text{ which is not centrosymmetrical.} \end{array}$$

The complete structure shows that $Ia (C_s^4)$ is the correct space group.

Bragg-Fourier projection on 010.

The visually estimated intensities from 48 reflections on the $h0l$ diagram were employed in the usual way¹ to give a projection of the structure on 010 (Figs. 3 and 4). The unit area contains only one mol of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ and, since the Sr atom with 38 of the total of 92 electrons is per definitionem at the origin of the projection to be calculated, all reflections stronger than "very weak" could safely be assumed to be positive. The unit of projection (of which

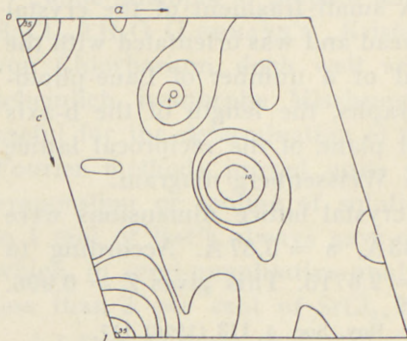


Fig. 3. Bragg-Fourier projection on 010 in $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$.

¹ Cf. $\text{SrCl}_2 \cdot 6\text{H}_2\text{O}$. D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd. XVII. 9. 1940.

Fig. 3 is the left half) contains—apart from the heavy atom at the origin, which was assumed in order to determine the signs of the F 's—two medium and two light atoms, in accordance with the formula $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$. The knowledge of this formula has, however, nowhere been applied to the calculations. The peaks are practically circular; moreover, the number of peaks is correct and their relative heights are reasonable. This shows that the atomic arrangement when projected on 010 is centrosymmetrical or at least very close to it with the origin as center of symmetry. However, it will be shown later that the centrosymmetry does not extend to the spatial arrangement of the atoms. The x - and z -coordinates referring to the unit of projection are

$$\begin{array}{ll} \text{Cl} & x = 0.287 \quad z = 0.510 \text{ etc.} \\ \text{O} & x = 0.226 \quad z = 0.240 \text{ etc.} \end{array}$$

and referring to the larger unit cell with Sr at $\frac{1}{4} \frac{1}{4} \frac{1}{4}$ as has been chosen later

$$\begin{array}{ll} \text{Cl} & x = 0.106 \quad z = 0.005 \\ \text{O} & x = 0.137 \quad z = 0.130. \end{array}$$

Table 1 shows the very satisfactory agreement between calculated and observed intensities.

As should be expected, the lowest orders of reflection from the only developed face, (100), are registered with an abnormally high intensity.

If the structure had been based on the space group $I2/a$ (C_{2h}^6) it would have been determined by 3 Cl parameters, 3 O parameters, and one or no Sr parameter¹.

Of these six or seven parameters, four were accurately known from the Fourier synthesis. Numerous attempts were now made

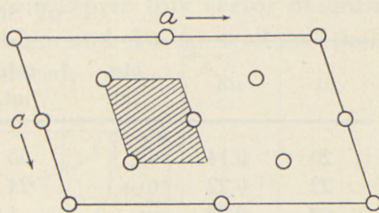


Fig. 4. Symmetry elements in 010 projection of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$. The density was calculated in the shaded area, only.

¹ Cf. Intern. Tabellen z. Best. v. Kristallstrukturen p. 102, where the orientation, however, is $C2/c$.

Table 1.

Calculated and observed intensities in the h0l zone
of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$.

hl	sin	obs. int.	calc. int.	hl	sin	obs. int.	calc. int.
20	0.14	s	65	46	0.70	w-m, m	12
22	0.22	m-s	24	06	0.72	0	0.6
02	0.24	w	1.0	6 $\bar{4}$	0.72	w-m	7.4
40	0.28	w, w-m	0.6	104	0.73	0	0
42	0.30	m-s, s	62	86	0.76	0	0.2
2 $\bar{2}$	0.31	m-s	55	2 $\bar{6}$	0.78	v	2.0
60	0.41	s	18	10 $\bar{2}$	0.79	m, m-s	9.4
62	0.41	w-m, m	10	122	0.79	w-m	2.9
4 $\bar{2}$	0.42	m-s	38	120	0.82	w	1.1
24	0.45	w-m	4.9	124	0.82	w-m, m	2.1
44	0.47	w-m	3.4	106	0.83	w-m	8.9
04	0.48	m-s	16	8 $\bar{4}$	0.83	w, w-m	4.8
6 $\bar{2}$	0.53	0	0.4	4 $\bar{6}$	0.85	w-m	8.0
64	0.53	m-s	17	126	0.92	w-m	10.3
82	0.53	w-m, m	3.2	68	0.93	w-m	6.4
2 $\bar{4}$	0.55	w-m	3.9	144	0.93	m-s	26
80	0.55	m-s	19	10 $\bar{4}$	0.93	0	0.5
4 $\bar{4}$	0.60	0	1.9	142	0.94	0	1.1
84	0.62	w-m	7.3	12 $\bar{2}$	0.96	w-m, m	13
8 $\bar{2}$	0.66	0	0.2	140	0.96	w-m	13
102	0.66	m-s	7.1	08	0.96	w-m, m	21
66	0.67	0	0.4	28	0.96	0	2.5
100	0.68	v, 0	1.8	48	0.96	0	1.5
26	0.70	w-m	7.0	88	0.97	w-m	10

on the basis of these parameters and of space considerations to build up a structure of the crystal. However, all attempts were unsuccessful. It was not possible to find an arrangement which led to reasonable distances between the atoms. It was deemed necessary to find the still unknown parameters independently, thereby settling the question whether the higher or the lower space group is the true one.

Patterson-Fourier projection on 001.

The material was a hk0 Weissenberg diagram. The tabulated intensities showed the expected symmetry properties of reflections

from Cmm. A set of approximate F^2 's was obtained from visually estimated spot intensities, and a Patterson synthesis was performed (Figs. 5 and 6). In order to interpret this vector diagram in terms of atomic coordinates, $I2/a$ and Ia as well as their projections on 001 must be considered.

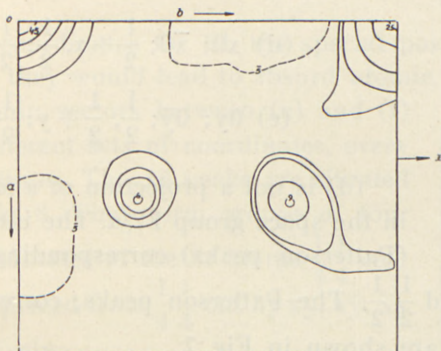


Fig. 5. Patterson-Fourier projection on 001 in $SrCl_2, 2 H_2O$.

$I2/a$ point positions: $\left(000; \frac{1}{2} \frac{1}{2} \frac{1}{2}\right) +$

4: (a) $000; 00 \frac{1}{2}$ (b) $0 \frac{1}{2} 0; 0 \frac{1}{2} \frac{1}{2}$

(c) $\frac{1}{4} \frac{1}{4} \frac{3}{4}; \frac{3}{4} \frac{1}{4} \frac{3}{4}$ (d) $\frac{1}{4} \frac{1}{4} \frac{1}{4}; \frac{3}{4} \frac{1}{4} \frac{1}{4}$

(e) $0y \frac{1}{4}; 0\bar{y} \frac{3}{4}$

8: (f) $xyz; \bar{x}\bar{y}\bar{z}; \bar{x}, y, \frac{1}{2} - z; x, \bar{y}, \frac{1}{2} + z$.

Its general point projects on 001 in

$$\left(00; \frac{1}{2} \frac{1}{2}\right) + xy; \bar{x}\bar{y}; \bar{x}y; x\bar{y}$$

which is the general point of the plane group Cmm (PATTERSON: C_{2v}^{IV})¹

¹ BRAGG and LIPSON. Z. Krist. (A) 95 (1936) 238. PATTERSON ib. 90 (1935) 546.

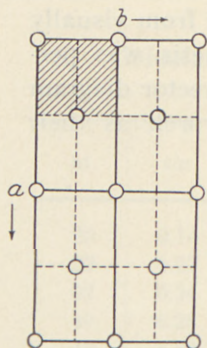


Fig. 6. Symmetry elements of 001 Patterson vector map.

Cmm contains the following special points

$$2: (a) 00; \frac{1}{2} \frac{1}{2} \quad (b) \frac{1}{2} 0; 0 \frac{1}{2}$$

$$4: (c) \frac{1}{4} \frac{1}{4}; \frac{3}{4} \frac{1}{4}; \frac{1}{4} \frac{3}{4}; \frac{3}{4} \frac{3}{4}$$

$$(d) x0; \bar{x}0; \frac{1}{2} + x, \frac{1}{2}; \frac{1}{2} - x, \frac{1}{2}$$

$$(e) 0y; 0\bar{y}; \frac{1}{2}, \frac{1}{2} + y; \frac{1}{2}, \frac{1}{2} - y.$$

(d) is not a projection of a four-fold position in the space group $I2/a$. The interatomic vectors (Patterson peaks) corresponding to (a) and (b) are at 00 and $\frac{1}{2} \frac{1}{2}$. The Patterson peaks, corresponding to (c), (d), and (e), are shown in Fig. 7.

Apart from the very large peaks at 00 and $\frac{1}{2} \frac{1}{2}$ which must always be present, the Patterson projection Fig. 5 shows a large peak at $x = 0, y = \frac{1}{2}$, about half the size of the 00 peak. This peak corresponds to the smaller peak in Fig. 5 (c)¹ and must be

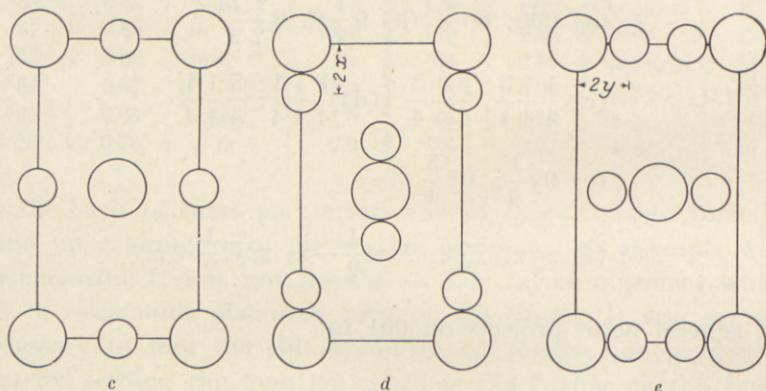


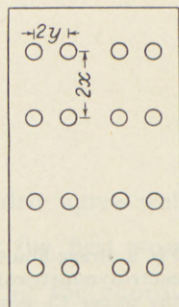
Fig. 7. Interatomic vectors (Patterson peaks) corresponding to the special positions (c), (d) and (e) in the plane group Cmm.

¹ It should be mentioned that Fig. 5 may also be interpreted as analogous to Fig. 7 (d) or (e) with the special value of $\frac{1}{4}$ of the variable parameter; however, this represents no new solution, but only a change of the origin of the coordinate system, even when the corresponding special positions in space ($I2/a$) are considered.

the Sr — Sr peak. Still assuming the space group to be I2/a, we have now to look for the next largest peaks which are due to interaction between 4 Sr at (c) and 8 Cl in the general position (f)

$$xy; \bar{x}y; x\bar{y}; \bar{x}\bar{y}; \frac{1}{2}+x, \frac{1}{2}+y; \frac{1}{2}-x, \frac{1}{2}+y; \frac{1}{2}+x, \frac{1}{2}-y; \frac{1}{2}-x, \frac{1}{2}-y$$

As possible coordinates for Cl, the special positions may be discarded, since they would lead to absurd atomic arrangements. The 64 interatomic vectors between (c) and (f) consist of 16 different sets of coordinates, every set occurring 4 times. The 16 peaks are situated as shown in Fig. 8. They form groups of four around the coordinates of the Sr atoms $\frac{1}{4} \frac{1}{4}$ etc., and they are displaced from $\frac{1}{4} \frac{1}{4}$ etc. by $xy, x\bar{y}, \bar{x}y$ and $\bar{x}\bar{y}$, respectively.



The projection, Fig. 8, actually shows 16 peaks, much smaller than the Sr—Sr peaks, arranged in sets of four around $\frac{1}{4} \frac{1}{4}$ etc. These are the Sr—Cl peaks, situated at a distance of $x = 0.109$ and $y = 0.100$ from $\frac{1}{4} \frac{1}{4}$ etc. $x = 0.109$ $y = 0.100$ would be the Cl-coordinates in the unit cell, and $x = 0.100$ agrees well with $x = 0.106$ found independently by Fourier projection on 010.

Fig. 8. Interatomic vectors between 4 atoms at $\frac{1}{4} \frac{1}{4}$ etc. and 8 at xy in Cmm.

The solution $xy = 0.11; 0.10$ is not the only one possible. $xy = 0.11; 0.40$ is another solution, but it corresponds to $0.11; 0.10$ and a shift of the origin. So far, no objection to a structure built on the higher space group has been encountered. A set of calculated intensity contributions from Sr and 2 Cl with the coordinates found did not agree very well with the observed values. The non-included oxygen contributions were made responsible for this failure, and the calculated Sr + 2 Cl contributions were employed to determine the sign of the $F(hko)$'s; a Bragg-Fourier synthesis was performed which is reproduced in Fig. 7, although it represents a false track. It is interesting to note that the erroneous assumption of centrosymmetry appears in the projection. The peaks occur at other values of xy than expected,

they are badly shaped, and the background is very far from smooth.

What was expected when it proved impossible to find the y parameters by space considerations now becomes apparent, viz. the structure is definitely not based on $I2/a$, and therefore the acentric group Ia has to be discussed.

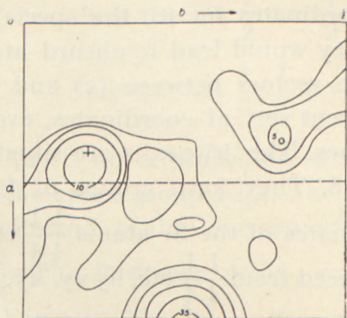


Fig. 9. Bragg-Fourier synthesis on 001 based upon the incorrect assumption of a centrosymmetrical space group. The large peak is a Sr atom. + indicates where the upper Cl peak should appear according to the 001 Patterson projection. The dotted line indicates where the 0 peak should be expected according to the 010 Bragg projection.

The group Ia has no special points and its general point is

$$xyz; x, \bar{y}, \frac{1}{2} + z; \frac{1}{2} + x, \frac{1}{2} + y, \frac{1}{2} + z; \frac{1}{2} + x, \frac{1}{2} - y, z$$

which projects onto 001 in

$$xy; x\bar{y}; \frac{1}{2} + x, \frac{1}{2} + y; \frac{1}{2} + x, \frac{1}{2} - y$$

the latter being the general point of the plane group Cm in the orientation $C1m^1$.

¹ BRAGG and LIPSON. *Z. Krist. A* **95** (1936) 327, PATTERSON. *Z. Krist. A* **90** (1935) 545. PATTERSON'S table contains an error, viz. $\bar{x}, \bar{y}; \frac{1}{2} - x, \frac{1}{2} - y$, should read $\bar{x}, y; \frac{1}{2} - x, \frac{1}{2} + y$. Another error was found on p. 546. C_{2v}^{II} (c) where $\bar{x}, y + \frac{1}{2}$ should read $\frac{1}{2} - x, \frac{1}{2} + y$. In view of the great applicability of the Patterson tables and because they have already been incorporated uncorrected into the handbook of Halla-Mark, the author ventures to draw the reader's attention to these errors.

A Sr atom in general position in C1m will give rise to the following interatomic peaks in a Patterson projection

xy	Weight
00	$4 z^2$
$\frac{1}{2} \frac{1}{2}$	$4 z^2$
$0, 2y$	$2 z^2$
$0, 2\bar{y}$	$2 z^2$
$\frac{1}{2}, \frac{1}{2} + 2y$	$2 z^2$
$\frac{1}{2}, \frac{1}{2} - 2y$	$2 z^2$

} Fig 7.

The positions of the large peaks in Fig. 5 also agree well with Ia. y_{Sr} is found to be $\frac{1}{4}$. In Ia, x and z for the first atom may be chosen at will, thereby fixing the origin of the coordinate system; taking $xz = \frac{1}{4} \frac{1}{4}$, we obtain the same set of coordinates for Sr as previously found in I 2/a. The special position (a) in the plane group C1m may be disregarded. It would give Sr—Sr peaks at 00 and $\frac{1}{2} \frac{1}{2}$ only, and the peak at $0 \frac{1}{2}$ is too large to be due to interaction between Cl and Sr or Cl and Cl.

A Sr atom at $\frac{1}{4} \frac{1}{4}$ etc. and a Cl atom at the general point xy etc. give rise to 16 Patterson peaks of weight $2Z_{Sr}Z_{Cl}$ arranged in sets of four at distances of $x_{Cl}y_{Cl}$, $\bar{x}_{Cl}y_{Cl}$, $x_{Cl}\bar{y}_{Cl}$, $\bar{x}_{Cl}\bar{y}_{Cl}$ from $\frac{1}{4} \frac{1}{4}$ etc. The 16 peaks of secondary size found in the Patterson synthesis were—when I 2/a was discussed—interpreted as interatomic peaks from 4 Sr at $\frac{1}{4} \frac{1}{4}$ etc., and 8 Cl at xy etc. (Cmm). In Ia, they must be considered to be superpositions of 2 sets of 16, each set originating from the interaction of 4 Sr at $\frac{1}{4} \frac{1}{4}$ etc., and Cl at xy etc. in Cm. Any one of the 16 peaks in Fig. 3 is at distances of 0.14, 0.15 from $\frac{1}{4} \frac{1}{4}$ etc. This leads to the following 4 equations for x_{Cl}

$$\left. \begin{array}{l} \frac{1}{4} - x \quad \frac{1}{4} + x \\ \frac{3}{4} - x \quad \frac{3}{4} + x \end{array} \right\} = 0.14$$

but only $x = 0.11$ and 0.39 are independent solutions. In y , all four solutions are independent. $y = 0.90; 0.10; 0.40; 0.60$. From

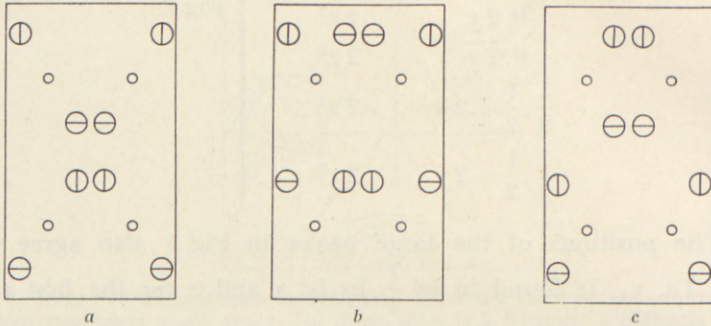
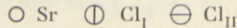


Fig. 10. The three projections of Sr and Cl on 001 according to a Patterson-Fourier synthesis.



(a) is the Cmm solution previously discarded, in (b) both axes are divided by 2, only (c) is possible.

these equations, the coordinates of Cl_I and Cl_{II} are determined. If mirror images etc. are eliminated, the number of independent solutions (atomic arrangements) is reduced to 3 shown in Fig. 10.

(a) is the Cmm solution previously found to be incorrect. In (b) the cell is no longer primitive, both axes are halved so this solution cannot be correct; it would mean that all reflections with h and k odd were due to oxygen atoms only, which is not the case, since a number of them are fairly strong.

(c) is the only remaining possibility. The Cl contributions were calculated according to (c).

$$\begin{array}{l} \text{Cl}_I \quad xy = 0.11 \quad 0.10 \\ \text{Cl}_{II} \quad xy = 0.39 \quad 0.10 \end{array}$$

$$\begin{array}{l} \text{Structure factor for Clm: } A = 4 \cos 2\pi hx \cos 2\pi ky \\ B = 4 \sin 2\pi hx \cos 2\pi ky \end{array}$$

and they were found to be a very definite improvement of the Cmm solution. All more striking discrepancies were removed.

The new arrangement agrees with the Cl parameter $x = 0.106$ from the 010 projection just as satisfactorily as did the discarded one.

The Patterson diagram Fig. 5 can also yield some information as to the O coordinates. According to the projection on 010 $x_0 = 0.14$. Sr—O peaks in the Patterson diagram should be

Table 2.
Calculated and observed intensities in the
hko zone of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$.

hk	sin	obs. int.	calc. int.	hk	sin	obs. int.	calc. int.
20	0.14	vs	65	93	0.73	v	0
11	0.14	v, w	3.4	06	0.73	m-s	10.4
31	0.24	m-s	6.2	26	0.74	w-m, m	2.7
02	0.26	m-s	11	84	0.75	w, w-m	0.5
40	0.28	w-m	0.9	111	0.77	w-m, m	1.1
22	0.29	s	29	75	0.78	w-m	1.2
13	0.38	0	0	46	0.78	0	0.1
42	0.38	s	27	120	0.83	w-m	1.0
51	0.38	v	0.2	66	0.84	m	4.1
33	0.42	w	0.3	113	0.84	v	0.1
60	0.41	vs	17	17	0.85	0	0
62	0.49	m-s	5.6	104	0.85	m-s, s	5.1
04	0.49	w	0	122	0.86	m-s	4.3
71	0.50	m, m-s	2.5	95	0.87	v	0.3
24	0.51	m-s	7.8	37	0.88	w	0
53	0.53	0	0	86	0.91	s	9.9
80	0.55	s, vs	19.3	131	0.91	w-m, m	0.7
44	0.56	s	18.8	57	0.92	0	0
82	0.60	m	2.5	08	0.96	m-s	10.5
15	0.62	w-m	0.1	140	0.96	m-s	14.4
73	0.62	v	0.2	133	0.96	0	0
91	0.63	w-m	0.5	124	0.96	m-s	11.3
64	0.64	w-m, m	2.8	77	0.97	0	0
35	0.64	v	0.9	115	0.97	w-m	3.1
100	0.70	w, w-m	1.6	142	0.98	m-s	7.8
55	0.70	v	0	28	0.98	m-s	9.1
102	0.73	m-s	3.7				

Table 3.

Atomic parameters xyz for $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$.Space group Ia (C_2^4), all atoms in general position:

$$xyz; \quad x\bar{y}\left(\frac{1}{2}+z\right); \quad \left(\frac{1}{2}+x\right)\left(\frac{1}{2}+y\right)\left(\frac{1}{2}+z\right); \quad \left(\frac{1}{2}+x\right)\left(\frac{1}{2}-y\right)z.$$

Sr	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	etc.
Cl _I	0.11	0.10	0.00	etc.
Cl _{II}	0.39	0.10	0.50	etc.
O _I	0.14	0.40	0.63	etc.
O _{II}	0.36	0.40	0.87	etc.

looked for at $x = 0.11$. The line $x = 0.11$ is seen to cut the upper parts of the two peaks of which the larger part is due to Sr and Cl. In fact, the slopes of the upper areas fall off more slowly than those of the lower areas, thus indicating other influences than Sr and Cl. The peak on the left extends from 0.13 to 0.18, the peak on the right from 0.32 to 0.39 at $x = 0.11$. Y_0 in the crystal hence must lie somewhere between 0.13 and 0.07; 0.10 was taken as best value.

A discussion similar to that in the case of Cl shows that there are only two possible sets of coordinates for the O atoms, provided they cannot be placed almost on top of the Cl atoms; the length of the c-axis 6.36 Å allows no such arrangement. The coordinates are

$$O_I \quad xy = 0.14 \quad 0.60 \quad \text{etc.}$$

$$O_{II} \quad xy = 0.36 \quad 0.60 \quad \text{etc.}$$

The assignment of proper z-coordinates to O_I , O_{II} , Cl_I and Cl_{II} is the next problem. In the full projection on 010 we have O-atoms at $z = 0.13 \quad 0.27 \quad 0.63 \quad 0.77$ and Cl atoms at $x = 0.00$ and 0.50; but only the arrangement whose full set of coordinates are given in table 3 was possible without almost superimposing atoms in several places.

Calculated and observed intensities are given in table 2.

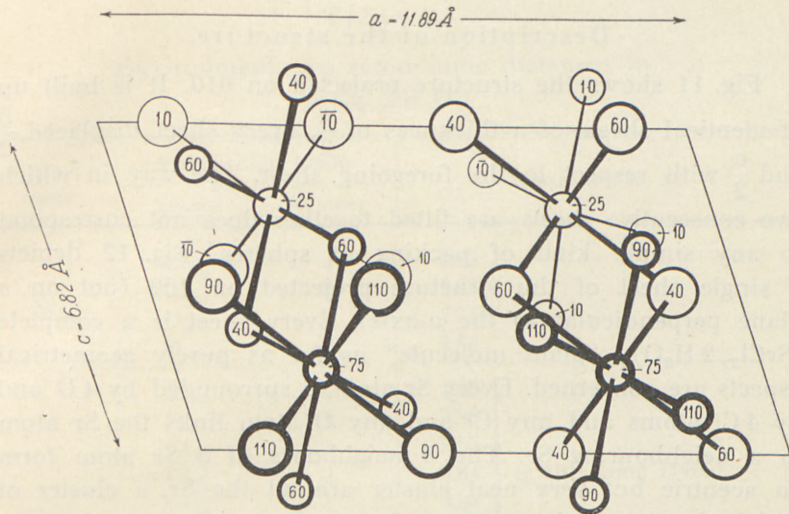


Fig. 11 $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ projected on 010, only bonds within the layers parallel to 100 are drawn. Where two atoms are superimposed, the lower one is slightly displaced in the figure.

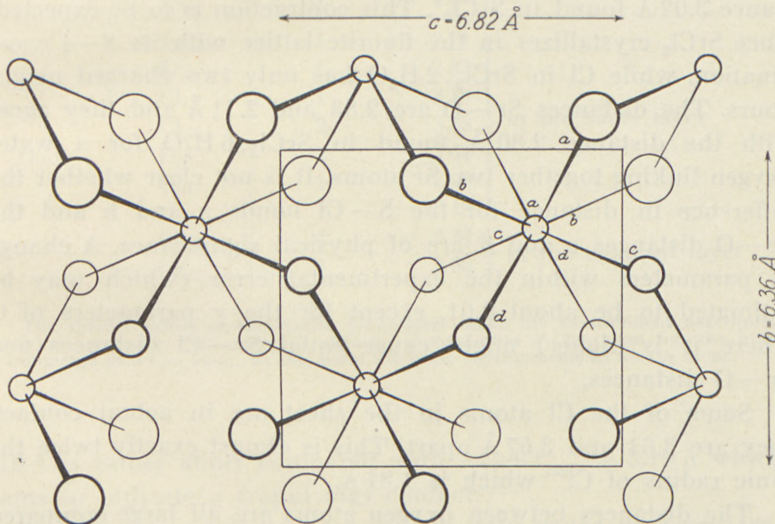


Fig. 12 A layer of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ projected on 100. Sr atoms lie in the plane of the drawing. Cl atoms 1.59 Å above and below. O atoms 1.23 Å above and below.

Description of the structure.

Fig. 11 shows the structure projected on 010. It is built up of identical sheets of a thickness of $\frac{a}{2}$, every sheet displaced $\frac{b}{2}$ and $\frac{c}{2}$ with respect to the foregoing sheet. The way in which two consecutive sheets are fitted together does not correspond to any simple kind of packing of spheres. Fig. 12 depicts a single sheet of the structure projected on 100 (not on a plane perpendicular to the a-axis). Every sheet is a complete $(\text{SrCl}_2, 2\text{H}_2\text{O})_n$ "giant molecule" as far as purely geometrical aspects are concerned. Every Sr atom is surrounded by 4 O and by 4 Cl atoms and any Cl and any O atom links the Sr atom to a neighbouring Sr. The 8 neighbours of a Sr atom form an acentric but very neat cluster around the Sr, a cluster of which the surrounding atoms of any adjacent Sr form the mirror image.

Table 4 gives some of the interatomic distances. The distances from Sr to Cl (2.88 and 2.99 Å) are somewhat below the distance 3.02 Å found in SrCl_2 ¹. This contraction is to be expected, since SrCl_2 crystallizes in the fluorite lattice with its 8—4 coordination, while Cl in $\text{SrCl}_2, 2\text{H}_2\text{O}$ has only two charged neighbours. The distances Sr—O are 2.83 and 2.71 Å and they agree with the distance 2.80 Å found in $\text{SrCl}_2, 6\text{H}_2\text{O}$ for a water oxygen linking together two Sr atoms. It is not clear whether the difference in distance for the S—Cl bonds a and b and the Sr—O distances c and d are of physical significance. A change of parameters within the experimental error (which may be estimated to be about 0.01, except for the y parameters of O where it is higher) might cause equal Sr—Cl distances and Sr—O distances.

Some of the Cl atoms in the sheet are in actual contact, they are 3.64 and 3.57 Å apart. This is almost exactly twice the ionic radius of Cl^- which is 1.81 Å.

The distances between oxygen atoms are all large compared with 2.76 Å which, according to W. L. BRAGG, is twice the radius of a water molecule; however, one distance from O to Cl of

¹ H. OTT: Z. Krist. **63** (1926) 222. Strukturberichte **1** 187.

Table 4.
Environment and interatomic distances in
SrCl₂, 2 H₂O.

Central Atom	Surrounding Atoms		Separation in Å
	Nature	Number	
Sr	Cl _(a)	2	2.99
	Cl _(b)	2	2.88
	O _(c)	2	2.71
	O _(d)	2	2.83
Cl _(I) ¹	Sr _(a)	1	2.99
	Sr _(b)	1	2.88
	Cl _(II)	1	3.57
	Cl _(I)	2	3.64
	Cl _(II)	1	3.98
	O _(I)	1	3.10
	O _(II)	2	3.35
	Cl _(II)	2	4.11
	O _(II)	1	3.34
	O _(II)	1	3.50
O _(I) ¹	Sr _(c)	1	2.71
	Sr _(d)	1	2.83
	Cl _(I)	1	3.10
	Cl _(II)	2	3.35
	O _(I)	2	3.64
	O _(II)	1	2.97
	O _(II)	1	3.48
	Cl _(II)	1	3.34
	Cl _(II)	1	3.50

¹ The special values of the atomic parameters make the environment description somewhat cumbersome. Particularly, it should be noted that a Cl_(II) atom has the same environments as a Cl_(I) atom, an O_(II) atom the same environments as an O_(I) atom.

3.10 Å is rather short compared with $r_{Cl^-} + r_{H_2O} = 3.19 \text{ Å}$ which seems to indicate a (repulsing) contact.

The interatomic distances from one layer to the next are large compared with the sums of the radii which are $2 r_{Cl^-} = 3.62 \text{ Å}$, $r_{Cl^-} + r_{H_2O} = 3.19 \text{ Å}$, but they are not improbable, since the lattice is decidedly a layer lattice. The layers are parallel to the only face developed; unfortunately, it has, however, not been

Table 5.

Calculated and observed intensities in the powder diagram of
 $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$. $\text{CrK}\alpha$ radiation.

hkl	\sin^2 calc.	\sin^2 obs.	obs. int.	calc. int.		
211	0.0838	0.0846	m-s	83		
202	0.1244	0.1258	m	{ 36.4		
310	0.1250				0	{ 26.4
002	0.1272					
$\bar{2}11$	0.1276	0.1284	{ 204			
020	0.1281			0	{ 22	
112	0.1495					0
121	0.1597	w-m, m	28			
400	0.1654			0	12	
220	0.1695					m-s
$\bar{1}21$	0.1814	m	71			
411	0.1859			vs	241	
312	0.1884					m
$\bar{1}12$	0.1930	w-m	29			
402	0.2069			m-s, s	71	
$\bar{2}02$	0.2143					m-s, s
321	0.2205	w-m, m	29			
222	0.2548			s	46	
022	0.2572					s
$\bar{1}11$	0.2736	m-s	31			
$\bar{3}21$	0.2863			..	(v)	
510	0.2905					0
420	0.2936	0	0.4			
213	0.2981			s	60	
130	0.2986					s-vs
512	0.3100	0	0.2			
$\bar{3}12$	0.3200			0	0.6	
031	0.3206					0
013	0.3226	0.3215	{ 47			
422	0.3350			m-s, s broad	{ 38	
231	0.3400					w-m
$\bar{2}22$	0.3430	s	59			
413	0.3554			..	(w-m)	
521	0.3657					w
602	0.3698	w	12			
611	0.3708			v	6.5	
600	0.3722					0.3726
330	0.3813	s, broad	{ 42			
$\bar{4}02$	0.3822			0	16	
$\bar{2}31$	0.3838					0
132	0.4050	0	37			
323	0.4130			0	6.6	
						0
		0	4.9			

possible to ascertain how perfect the cleavage is, parallel to the layers. The crystals were too thin and too fragile as to permit an investigation of this kind.

A more detailed discussion of the bonding and dynamics of the lattice will be given in a later communication.

Powder photographs.

In the derivation of the structure, no reflections of the general type hkl have been employed. For a final check, the observed intensities of powder photographs were compared with a set of intensities calculated from the set of parameters found. The powder photographs were taken with the focussing cameras of PHRAGMÉN-HÄGG with Cr-radiation. The reflections from low to medium orders, only, have been indexed. The difficult indexing of the few and feeble lines found at higher angles was considered unnecessary. The work of refining 15 parameters by trial and error would have been very great, however, the agreement between calculated and observed intensities is good enough to show the fundamental correctness of the structure.

The structure factor of C_s^4 in the orientation $I_a (= I_c = I_n)$ is

$$\left. \begin{aligned} h+k+l &= 2n+1 \\ h+k+l &= 2n \\ l &= 2n \end{aligned} \right\} \begin{aligned} A &= B = 0 \\ A &= 4 \cos 2\pi (hx+lz) \cos 2\pi ky \\ B &= 4 \sin 2\pi (hx+lz) \cos 2\pi ky \end{aligned}$$

$$\left. \begin{aligned} h+k+l &= 2n \\ l &= 2n+1 \end{aligned} \right\} \begin{aligned} A &= -4 \sin 2\pi (hx+lz) \sin 2\pi ky \\ B &= 4 \cos 2\pi (hx+lz) \sin 2\pi ky \end{aligned}$$

The accurate lattice constants found were: $a = 11.76 \text{ \AA}$.
 $b = 6.38 \text{ \AA}$. $c = 6.66 \text{ \AA}$. $\beta = 72.9^\circ$

Summary.

$\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ prepared in different ways always crystallizes in the same lattice.

Its structure has been determined from Weissenberg photographs by Bragg- and Patterson-Fourier syntheses. The structure is determined by 15 parameters and is a layer lattice. Every layer is a two-dimensional giant molecule of $\text{SrCl}_2 \cdot 2\text{H}_2\text{O}$ in which every Sr is surrounded by 4 Cl and 4 O atoms. Every

Cl and every O link two Sr atoms. The interatomic distances show that the forces within the layers are strong, from layer to layer, however, relatively weak.

I am indebted to Professor GUNNAR HÄGG and the University of Uppsala for the hospitality granted me during the spring term of 1940 when the early part of the present work was done, and moreover to Professor HÄGG for many helpful discussions. My stay at Uppsala was made possible by a grant from K. A. LARSEN'S Legat, for which I express my sincere thanks.

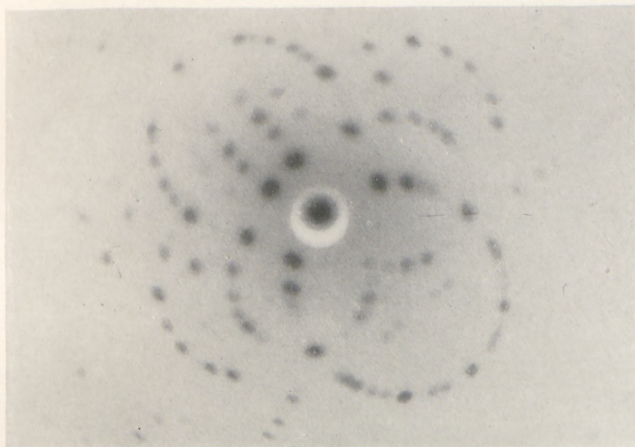


Fig. 1. Laue photograph of $\text{SrCl}_2 \cdot 2 \text{H}_2\text{O}$. Beam perpendicular to developed face.

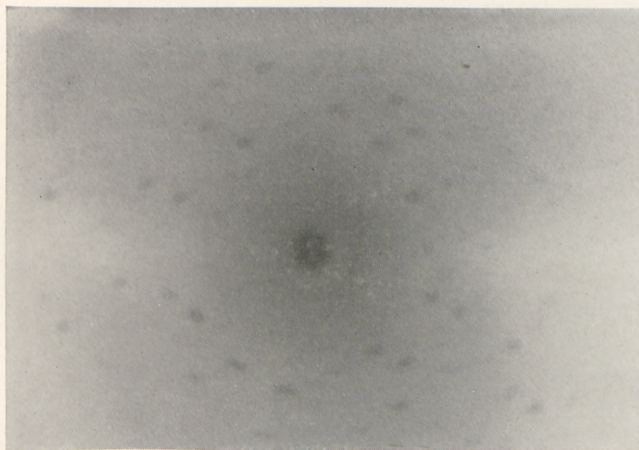


Fig. 2. Laue photograph of $\text{SrCl}_2 \cdot 2 \text{H}_2\text{O}$. Beam perpendicular to symmetry plane.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 6

THE BEGINNING OF THE
PHOTON-INITIATED ROSSI-CURVE FOR
Pb, Fe and Al

BY

B. TRUMPY



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

1. Introduction.

The beginning of the Rossi-Curve has in recent years been the object of extensive examination by a long series of investigators. Thus HU CHIEN SHAN, KISILBASCH, KETILADGE^{1) 2)}, WATASE³⁾ and CLAY and JONKER⁴⁾ have by means of Geiger-Müller counters examined the shower-production in thin sheets of several different materials. The prevailing view among these investigators seems to be that the usual experimental arrangement with three or more counter tubes in coincidence gives a Rossi-Curve, the course of which—at least for small thicknesses—is essentially determined by the soft component of the radiation complex.

The thickness of the shower-producing material was accordingly indicated in shower-units— l —, and it was then found—with essential agreement among the authors—that the Rossi-Curves for the different materials were completely identical within the limits $l = 0$ to $l = 1$. CLAY and JONKER⁴⁾ consider this result a confirmation of the cascade-theory of electron-showers.

To this, however, must be remarked that a theoretical calculation—based upon the cascade-theory—of the number of showers produced by the soft radiation component, leads to results actually contradicting the experiments mentioned above. Calculations to this effect are made by ARLEY and ERIKSEN^{5) 6)}. They have calculated the theoretical Rossi-Curve for different shower-producing materials, both for a primary radiation consisting of electrons only and of photons only. By combining these results the Rossi-Curves for the soft component are determined as a whole, and the results may to a large extent be directly applied to experiments with counter tubes. The authors find that the Rossi-Curves for small thicknesses *do not coincide* for different materials—the curve within the limits $0 < l < 1$ lying the higher, the greater the atom number of the material. In a paper

previously published on the shower production of mesons we⁷⁾ have pointed out, as already pointed out by ARLEY^{5) 6)}, that this disagreement between theory and experiment is probably due to the fact that the experiments also include the hard component. By the usual determination of the Rossi-Curve—by means of counter tubes—those showers which are produced by the hard component will also be measured. Our experiments with a Wilson-Chamber⁷⁾ have shown that the equilibrium between the mesons and their secondaries is reached in very thin metal sheets, and the influence of the meson showers on the course of the beginning of the usual Rossi-Curve is, therefore, not to be ignored. Our experiments showed, further, that the curves giving the number of meson secondaries as a function of the sheet thickness for different materials very nearly coincide, if the thickness is stated in g/cm^2 . This fact shows that the Rossi-Curves of the hard component do not coincide if the thickness is stated in shower units—l—, lying in this case the lower, the greater the atom number of the material.

As the Rossi-Curves plotted on an l-scale for the whole radiation according to the above experiments coincide for small thicknesses of different materials we may, further, draw the conclusion that this can not be the case for *the soft radiation alone*, the curves lying in this case necessarily higher, the greater the atom number of the material.

This fact is qualitatively in accordance with the theory.

In order to find a base for a quantitative verification of the cascade theory on this point we have, by means of the method of anti-coincidences, made an experimental determination of the photon-initiated Rossi-Curve for thin sheets of different materials. In the following we shall give a summary of the preliminary results of this investigation made with Pb, Fe and Al, and compare them with the theoretical calculations of ARLEY and ERIKSEN.

2. Experimental Arrangement.

Fig. 1 gives a vertical cut through the experimental arrangement perpendicular to the axis of the counter tubes. The combination K_1 , K_2 and K_3 consists of counter tubes with a diameter

of 2 cm and an effective length of 15 cm. The three tubes K_3 are in parallel. K_1 , K_2 and K_3 are placed in an ordinary coincidence coupling, and this part of the apparatus has, according to performed measurements, a resolving power of $3 \cdot 10^{-6}$ min.

Above the counter tubes K are placed in parallel a total of 11 counter tubes, A , each with a length of 35 cm. These counters are coupled in anti-coincidence, i. e.: a coincidence in the K tubes will only be registered provided that no simultaneous discharge is taking place in one or more of the tubes A . In the space between the tubes A and K sheets F of different materials and varying thickness may be placed.

What is measured with this arrangement is thus, obviously, the showers—containing at least one electron—which are produced in the metal sheet F by a non-ionising radiation. In a previously published paper⁸⁾ we have shown that in the cosmic radiation complex there do not exist neutral mesons which produce ionising radiation in traceable quantities in metal sheets of the thickness applied here. The effect measured by the experimental arrangement in fig. 1 will, therefore, exclusively be due to photons, and what is measured will accordingly be the photon-initiated Rossi-Curve, corresponding to a number of electrons per shower of $N \geq 1$. As the zero-effect by measurements with thin metal sheets may have a considerable influence on the result, the attempt has been made to reduce the former as much as possible by completely covering the sheet F with tubes in anti-coincidence. The remaining zero-effect, 4.8—which is relatively insignificant—is then chiefly due to a production of showers by photons in the lower walls of the tubes A . This wall is made of brass, 0.4 mm thick.

The resolving power of the anti-coincidence coupling was

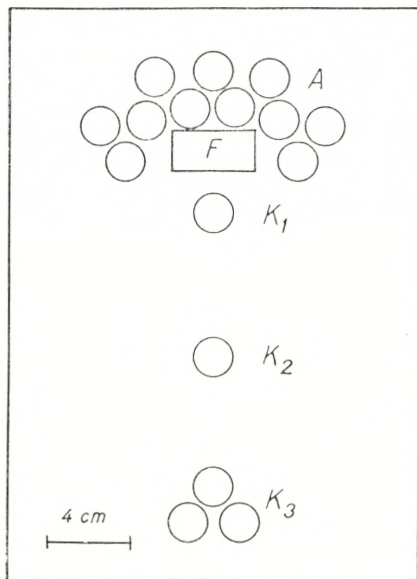


Fig. 1. Arrangement of counter tubes.

determined by placing the tubes A besides the tubes K at a distance of about 1 m. from the latter. Between A and K there was placed a lead filter with a thickness of 10 cm. The number of coincidences N_1 in the K tubes was measured while the A-tubes were excluded and next the number of anti-coincidences $-N_2-$ was counted.

$N = N_1 - N_2$ then gives the number of accidental coincidences between the combinations A and K, and the resolving power may be calculated in the usual way. It was with 11 A-tubes about 10^{-5} min.

3. Results.

Measurements were made with Pb, Fe and Al. Maximum of thickness used for the sheet F was in the first experiments the same for all materials, restricted by the geometry of the experimental arrangement, and it is therefore—expressed in l-units—decreasing from Pb to Al.

In the tables 1, 2 and 3 we give the results of the measurements, the zero-effect having been subtracted.

Table 1.
Pb.

Thickness in l-units	Hours	Anti-coincidences per Hour
0.28	142.5	8.5 + 0.3
0.56	164.0	14.5 + 0.3
0.84	167.0	18.5 + 0.4
1.12	165.0	20.5 + 0.5
1.84	294.0	23.3 + 0.6
2.93	145.0	21.0 + 0.4
3.69	74.0	19.0 + 0.6
4.80	120.0	16.6 + 0.4
6.39	22.0	10.9 + 0.9

The last measurement in Table 3 is made with about 73 mm Al. In this case the upper counter tubes—A—had to be raised in order to make room for the thick aluminium sheet. This circumstance means that the geometry of the experimental arrangement was somewhat changed, and it is possible that

we have, consequently, to allow for a greater margin of error in this measurement than for the other measurements, the correction

Table 2.
Fe.

Thickness in l-units	Hours	Anti-coincidences per Hour
0.25	69.0	8.2 ± 0.5
0.50	77.0	11.1 ± 0.5
0.75	72.0	14.5 ± 0.5
1.00	167.8	15.7 ± 0.4
1.50	183.3	17.5 ± 0.4
2.43	167.5	15.7 ± 0.4

for the zero-effect being among other things greater. In fig. 2 the results in the tables 1—3 are shown in curves, the fully drawn

Table 3.
Al.

Thickness in l-units	Hours	Anti-coincidences per Hour
0.09	124.0	3.3 ± 0.3
0.27	120.0	7.9 ± 0.3
0.45	116.5	10.0 ± 0.4
1.10	258.8	13.8 ± 0.2^1

¹ Carried out with the tubes A raised.

curves being marked Pb, Fe and Al. As to the broken curves with the same notations see next paragraph.

4. Discussion.

Our experiments with the anti-coincidence method give according to fig. 2 the following chief results for $N \geq 1$, N being the number of electrons per shower.

1. The photon-initiated Rossi-Curves for different materials *do not coincide* for small sheet-thicknesses when the latter are plotted in shower units, the curves lying within the limits

$0 < l < 1$ the higher, the greater the atom number of the shower-producing material.

2. The number of anti-coincidences at the maximum of the curve increases with increasing atom number of the shower-producing material and is for Pb 1.7, for Fe 1.3 times greater than for Al.

3. The maximum of the photon-initiated Rossi-Curve advances with increasing atom number of the shower producing material towards greater values of l . Thus it lies for Pb, Fe and Al by the values $l = 2.0$, 1.7 and about 1.1 respectively.

Point 1 stands in direct conflict with the conclusion drawn by HU, KISILBASCH, KETILADGE, WATASE and CLAY based on their experiments of the course of the Rossi-Curve by means of the usual coincidence arrangement. This confirms our previously⁷⁾ advanced hypothesis that the hard component in the cosmic radiation has a perceptible influence on the course of the Rossi-Curve, even for very thin sheet of the absorbing material.

It is of great interest to compare our results with the cascade theory for the soft radiation which through ARLEY has developed so far as to allow for a direct comparison. In their last work ARLEY and ERIKSEN have made a theoretical calculation of the course of the photon-initiated Rossi-Curve for different materials. These calculations must, of course, chiefly be founded on our knowledge of the energy spectrum of the photon radiation at the surface of the earth and as experiments cannot as yet give exact information on this subject, the authors have tried to make a calculation of the spectral distribution based on certain simplified assumptions. They suppose e. g. that all electrons and photons found at sea-level are produced by multiplication in the atmosphere of the primary radiation reaching the earth from outside. They further suppose that the primary radiation chiefly consists of high energy electrons, with a spectrum:

$$f(E_0) dE_0 = \text{const.} \frac{dE_0}{E_0^{1+\gamma}}$$

where the numerical value of γ lies between 1 and 2.

These assumptions are not in full accordance with the results of the investigations made in recent years on the transformation

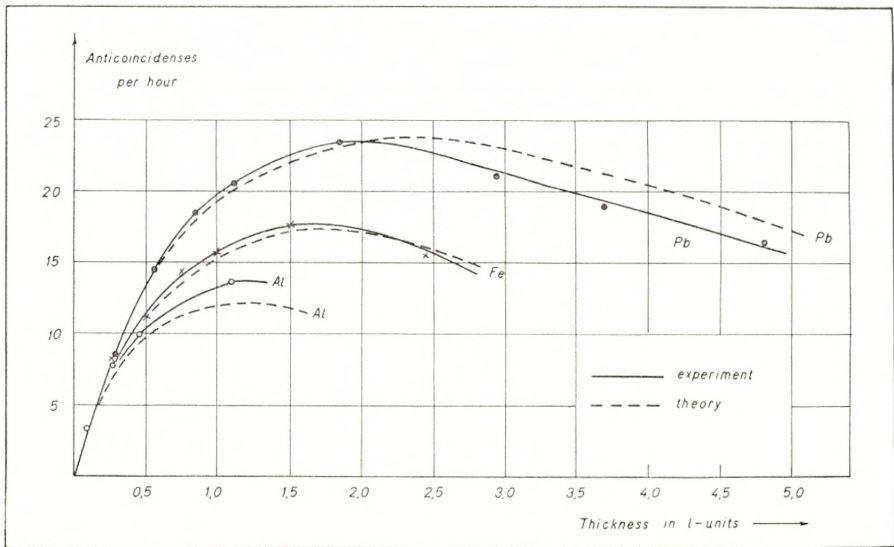


Fig. 2. Photon-initiated Rossi-Curves for Pb, Fe and Al.

of the cosmic radiation complex from the top of the atmosphere to the surface of the earth. It seems especially open to criticism that the photons produced in the atmosphere by multiplication of the electron secondaries of the mesons are not taken into consideration.

We may, however, from numerical calculations confirm the statement of ARLEY and ERIKSEN that the ultimate result does not depend significantly upon the form of the spectrum, and the calculated Rossi-Curves may, therefore, be expected to be correct in the main features, even if too much importance must not be attached to a quantitative comparison with the experimental results.

ARLEY and ERIKSEN have calculated the photon-initiated Rossi-Curves for Pb, Fe and Al. In these calculations are, however, included both "fast" and "slow" electrons, whereas our counter tube experiments will exclude electrons not being able to penetrate through the tube walls. The limit lies here presumably about 10^7 e. V., and the "slow" electrons ought accordingly not to be included in the calculations.

We have made some numerical calculations in order to fix the influence on the result when the slow electrons are excluded. The alteration thereby made in the course of the curves is of

essential importance only for sheet thicknesses greater than $l = 1.0$, and within the limits which are experimentally examined in this paper it will have a perceptible influence only for the Pb curve. This curve for "fast" electrons is calculated by ARLEY and ERIKSEN (l. c.). The broken curves in fig. 2 represent the theoretical values for Pb, Fe and Al in the case of $N \geq 1$. The theoretical curve for Fe is placed in such a manner that the maximum of the experimental and of the theoretical curve lie on the same level.

As will be seen, the theory renders the chief results in point 1—3 quite correctly. The curves do not coincide within the limits $0 < l < 1$, lying higher, the higher the atom number of the material; the maximum shower number increases with the atom number, and the maximum at the same time advances towards greater values of l . One dare say that also the quantitative accordance between theory and experiment is so good as may be expected, when the approximations used in the calculations are taken into account.

Finally we have to express our sincere thanks to Mr. T. BJORDAL who has rendered valuable aid during the experiments. We shall further express our thanks to The Scientific Research Fund of the Norwegian State (Statens Vitenskapelig Forsknings-fond) which has financed the work.

*Geofysisk Institutt
Bergen July 1942.*

References.

- 1) HU CHIEN SHAN: Proc. Roy. Soc. A. **158**, 581, 1937.
 - 2) HU CHIEN SHAN, KISILBASCH and KETILADGE: Proc. Roy. Soc. A. **161**, 95, 1937.
 - 3) WATASE: Proc. Phys. Math. Soc. Japan. **19**, 749, 1937.
 - 4) J. CLAY and K. H. JONKER: Physica **VII** 921, 1940.
 - 5) N. ARLEY: Proc. Roy. Soc. A. **168**, 519, 1938.
 - 6) N. ARLEY and B. ERIKSEN: D. Kgl. Danske Vidensk. Selskab, Mat.-fys. Medd. **XVII** 11, 1940.
 - 7) B. TRUMPY: Det Kgl. Norske Vid. Sel. Forh. **XIV** 63, 1941.
 - 8) B. TRUMPY and T. BJORDAL: Bergens Museums Årbok 1942.
-

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 7

THE ACID DISSOCIATION OF
THE HYDRATED CUPRIC ION
THE FORMATION OF DIMERIC CUPRIC
COMPOUNDS

BY

KAI JULIUS PEDERSEN



KØBENHAVN

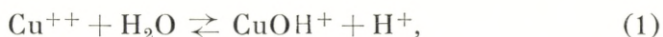
I KOMMISSION HOS EJNAR MUNKSGAARD

1943

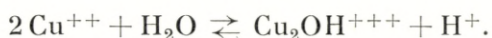
Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

Our knowledge of the hydrogen ion concentration of solutions of cupric salts, and the equilibria which determine it, is very incomplete. The hydrogen ion concentration of a cupric sulphate solution was first measured by CARRARA and VESPIGNANI⁽¹⁾ using the method of inversion of sucrose. KULLGREN⁽²⁾ applied the same method to solutions of cupric chloride and nitrate. The quinhydrone electrode was used by O'SULLIVAN⁽³⁾ and by FRIEDMAN and STOKES⁽⁴⁾ in cupric sulphate solutions, by QUINTIN⁽⁵⁾ in cupric nitrate and sulphate solutions, and by ČUPR and ŠIRŮČEK⁽⁶⁾ in different cupric salt solutions. Finally, the glass electrode was employed for measuring the pH of cupric sulphate solutions by HUGHES⁽⁷⁾ and by CRANSTON and BROWN⁽⁸⁾. A selection of the best measurements is given in Fig. 1.

It is usually assumed that the hydrogen ion concentration is determined by the equilibrium



but the measurements are not sufficiently accurate to test this assumption. If anything, they rather disfavour it. Thus, QUINTIN as well as ČUPR and ŠIRŮČEK state that, when pH is plotted against the logarithm of the concentration of the copper salt, the points fall close to a straight line with the slope one (cf. Fig. 1), or, in other words, the hydrogen ion concentration is roughly proportional to the cupric ion concentration. If this holds, equilibrium 1 alone does not determine the hydrogen ion concentration. The relation found by these investigators rather indicates the equilibrium



JANDER and MÖHR⁽⁹⁾ recently studied the diffusion of solutions of cupric perchlorate and cupric chloride in the presence of neutral salt or strong acid. For all their solutions, they found the diffusion coefficient to be the same within a few per cent, while pH varied from 5 to 0. However, their pH values were often several units lower than should be expected from measure-

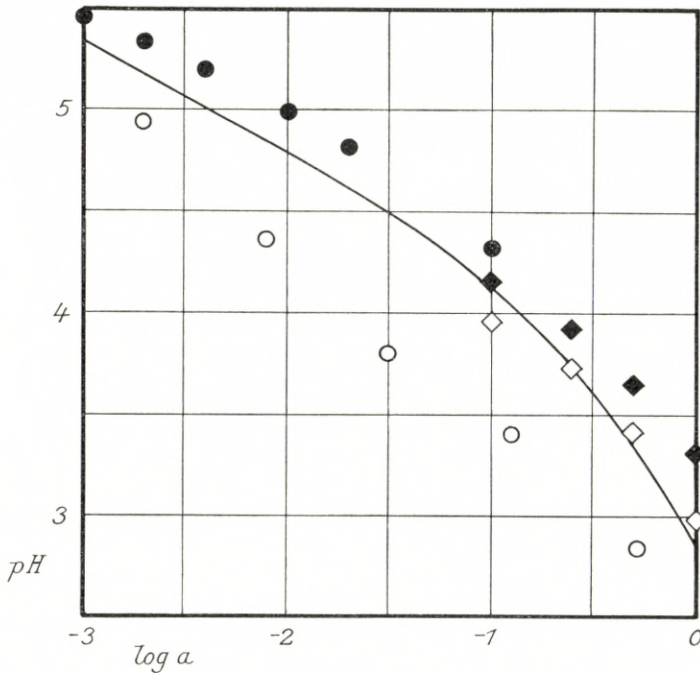


Fig. 1. Measurements of the pH of solutions of cupric nitrate and sulphate (a molar) by the investigators: QUINTIN (quinhydrone electrode, cupric nitrate, open circles), CRANSTON and BROWN (glass electrode at 15°C ., cupric sulphate, solid circles), and ČUPR and ŠIRŮČEK (quinhydrone electrode at 20°C ., cupric nitrate, open squares, and cupric sulphate, solid squares). The curve represents the measurements of the present author.

ments by other investigators, including the author of the present paper. It appears from their results that the degree of hydrolysis varies from 0 to a very considerable value ($> 20\%$), while the diffusion coefficient is practically constant. Since a dimeric ion would have a smaller diffusion coefficient than the usual cupric ion, they conclude that only monomeric hydrolysis products are present. However, it seems safe to maintain that the degree of hydrolysis of their solutions is always less than 1% .

Therefore, their measurements cannot give any information concerning the nature of the hydrolysis products.

The object of the present paper is to throw light on the problem of acid-base equilibria in aqueous solutions of cupric ions. For this purpose, the hydrogen ion concentrations of solutions of cupric nitrate, varying in concentration from 1 to 0.001 molar, were measured by means of the glass electrode at 18.0° C. Cupric nitrate was chosen rather than sulphate, because the presence of sulphate ions would make the problem more complicated owing to their power of forming complexes with cupric ions and of combining with hydrogen ions. Furthermore, the greater electric charge of the sulphate ion would cause a greater salt effect. Measurements were carried out both on pure aqueous solutions of cupric nitrate and on solutions containing, in addition, various amounts of nitric acid or sodium hydroxide. It is well known that only very little sodium hydroxide can be added before a precipitate is formed. According to J. BJERRUM⁽¹⁰⁾, the green crystalline substance which precipitates in weakly alkaline solutions of cupric nitrate has the composition $\text{Cu}_2(\text{OH})_3\text{NO}_3$. It was often possible to carry out measurements on solutions containing so much sodium hydroxide that they were supersaturated with respect to the substance forming the precipitate. For such solutions, the e. m. f. remained constant for a period of time, but when they were kept standing sufficiently long, a gradual increase in the hydrogen ion concentration was measured. At the same time, the solutions turned turbid. The greatest concentration of sodium hydroxide which could be maintained without disturbing the measurements was: in 1 molar $\text{Cu}(\text{NO}_3)_2$ only 0.2 % of the concentration of cupric nitrate, in 0.1 molar $\text{Cu}(\text{NO}_3)_2$ it was 1.2 %, and in 0.005 molar $\text{Cu}(\text{NO}_3)_2$ it was more than 10 %. These examples show that the solutions measured were very weakly buffered. There is no sense in carrying out measurements on such systems if the amount of acid or basic impurity present is not known with great accuracy. We shall see later how these impurities may be estimated.

Cupric nitrate hexahydrate was used for the experiments. It was prepared by recrystallization of the best trihydrate available (Schering-Kahlbaum's *Cuprum nitr. puriss.*). 1000 g. of the

trihydrate were dissolved in 400 ml. of warm water. The solution was cooled down to 0°, seeded with the hexahydrate, and was left for crystallization at 0°. The crystals were separated from the mother liquor and were dried in a desiccator over anhydrous calcium chloride. When the crystals began to effloresce the drying agent was removed. The preparation contained an excess of base, or a deficiency of acid, which increased gradually during storage. At the beginning of the measurements it amounted to 3×10^{-5} , at the end to 18×10^{-5} equivalents per mole of cupric nitrate. This increase is probably due to the fact that the glass basin containing the preparation was inadvertently placed in the desiccator on a perforated iron plate. The nitric acid lost may have distilled from the slightly moist crystals to the iron plate.

Standard solutions of about 1 or 2 molar cupric nitrate were made from the preparation at different times. The solutions, which were kept in a silica flask, were analyzed for copper according to the modified iodometric titration method of HAGEN⁽¹¹⁾, which gave excellent results (accuracy about 0.1 %). Metallic copper (electrolytically deposited, *pro analysi*) was used as a standard substance.

The barium nitrate used in some of the measurements was a Merck *pro analysi* preparation which was recrystallized and dried at 110° C. It contained no appreciable basic or acid impurities.

The water used for the solutions was redistilled in a still supplying water of a specific conductance of less than 3×10^{-7} ohmcm⁻¹. It was stored in bottles covered inside with a layer of paraffin wax.

The cells measured had the composition



where X is the solution under investigation. The glass electrode was of the plane type described by MACINNES and DOLE⁽¹²⁾. It was tested before use, and was found to follow practically completely the law holding for the hydrogen electrode through-

out the pH interval in which it was going to be used. As container for the solution under investigation served the vessel shown in Fig. 2. Its total volume was 100 ml., but only 70 ml. of solution were used for the measurement. For the solutions containing more than 0.1 molar cupric nitrate, a similar, but smaller vessel requiring only 10 ml. was applied. The opening D was closed with a rubber stopper carrying the glass electrode which dipped into the solution. The side tube E was completely

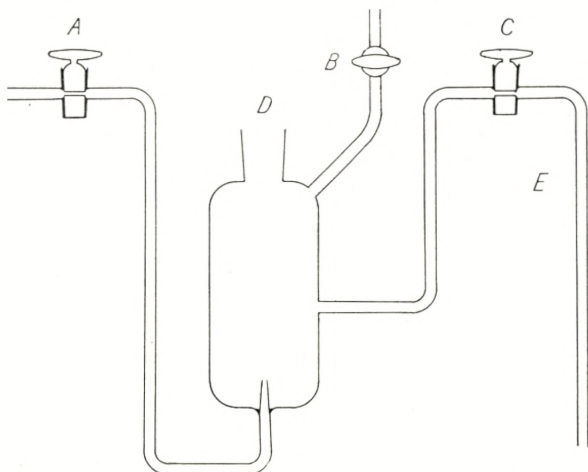


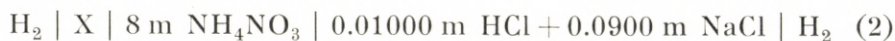
Fig. 2. Electrode vessel.

filled with the solution. Before the measurement, possible traces of carbon dioxide were removed by a vigorous current of pure, carbon dioxide free air which was passed through the solution for 10 minutes. The air entered through the stopcock A and escaped through B.

The ready cell was placed into a liquid paraffin thermostat at 18.0°C . The lower end of E dipped into a test tube containing 8 molar ammonium nitrate which provided connection with the quinhydrone electrode. The levels of the solutions were so arranged that the bridge solution rose a few centimetres into E when the stopcock C was opened. The e. m. f. was measured by means of a valve potentiometer (made by the firm "Radiometer", Copenhagen). The potential had become constant after a few minutes, and generally it did not change for hours. Only when the solution was supersaturated, a gradual decrease was

observed, as already mentioned, after a longer or a shorter period of constant potential.

At least once every day, a solution of the composition 0.01000 m HCl + 0.0900 m NaCl was measured instead of X. The small potential measured here is the asymmetry potential of the cell. When it is subtracted from the e. m. f. measured for X, the potential E of a cell of the composition



is obtained.

To a series of measurements at constant cupric nitrate concentration but varying nitric acid or sodium hydroxide concentration, we may apply the equation

$$E = E_a - \frac{RT}{F} \ln(\text{H}^+),$$

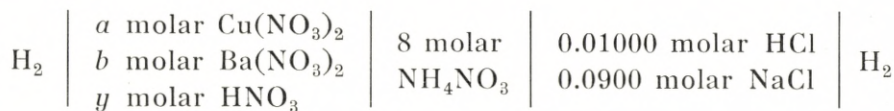
where (H^+) is the hydrogen ion *concentration*, and E_a includes the constant potential of the right half cell, the liquid-liquid potential, and the salt effect of the left half cell. When a sufficiently small amount of acid or base is added, E_a is constant. For the temperature 18° C. the equation may be written

$$-\log(\text{H}^+) = A + 17.32 E, \quad (3)$$

where $A \equiv -17.32 E_a$. The value of A corresponding to an infinitely small concentration of acid or base added will be denoted by A_0 .

A was determined in the following way. E was measured for solutions of cupric nitrate containing so much nitric acid that the hydrogen ion concentration produced by the cupric ions was either negligible or could be estimated from preliminary values of the equilibrium constants. In no case, this correction exceeded 0.5 % of the concentration of nitric acid added. Also the small acid deficiency of the cupric nitrate, estimated from later measurements, was taken into account. (H^+) and E were introduced into equation 3, and A was computed. A_0 , the value of A corresponding to the concentration 0 of nitric acid, was found for each series by extrapolation. The results are given in Table 1. The table includes two series of measurements in

Table 1.

Electromotive force E volts at 18.0° C. of the cell

measured by means of the glass electrode. $y = -x - a\delta$, where $-x$ is the concentration of nitric acid added, and δ is the deficiency of nitric acid in the preparation of cupric nitrate expressed in equivalents per mole. A has been computed from equation 3. A_0 is the value of A corresponding to $y = 0$.

a	b	$\delta 10^5$	y	E	A	A_0
1.000	0	18	0.04975	-0.0578	2.304	2.308
			0.02468	-0.0403	2.305	
			0.01002	-0.0180	2.309	
0.500	0	16	0.04098	-0.0424	2.122	2.123
			0.02008	-0.0247	2.125	
			0.00801	-0.0016	2.123	
0.200	0	8	0.01027	-0.0023	2.028	2.033
			0.004104	0.0205	2.031	
			0.002059	0.0377	2.032	
0.100	0	3	0.01000	-0.0005	2.009	2.018
			0.005015	0.0166	2.012	
			0.002525	0.0335	2.017	
0	0.100		0.01000	-0.0005	2.009	2.018
			0.005017	0.0166	2.012	
			0.002528	0.0335	2.017	
			0.001025	0.0561	2.017	
0.0500	0	8	0.005134	0.0160	2.012	2.021
			0.002055	0.0387	2.017	
			0.001033	0.0558	2.019	
0.0200	0	3	0.005138	0.0154	2.022	2.040
			0.002059	0.0378	2.032	
			0.001033	0.0558	2.033	
0.0100	0	3	0.00504	0.0153	2.033	2.058
			0.002021	0.0373	2.048	
			0.001018	0.0542	2.054	
0.0020	0.0080	3	0.002021	0.0374	2.046	2.058
			0.001018	0.0544	2.050	
			0.000413	0.0767	2.056	

Table 1 (continued).

a	b	$d \cdot 10^5$	y	E	A	A_0
0.0050	0	18	0.005005	0.0154	2.034	2.072
			0.002006	0.0372	2.053	
			0.001011	0.0540	2.060	
			0.000410	0.0762	2.067	
0.0020	0	18	0.005040	0.0152	2.034	2.086
			0.002020	0.0364	2.064	
			0.001018	0.0531	2.072	
			0.000413	0.0753	2.080	
0.0010	0	18	0.005005	0.0145	2.050	2.097
			0.002006	0.0363	2.069	
			0.001011	0.0528	2.081	
			0.000410	0.0749	2.090	

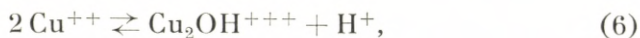
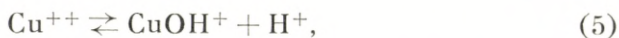
solutions containing barium nitrate. It is seen that, both in 0.1 and 0.01 molar cupric nitrate, the cupric ions may be replaced by barium ions without altering the value of A_0 .

In the following, the hydrogen ion concentration of solutions of cupric nitrate containing small concentrations of sodium hydroxide or nitric acid will be computed from the measurements by means of the formula

$$-\log(\text{H}^+) = A_0 + 17.32 E, \quad (4)$$

where A_0 has the values given in Table 1. Before analyzing the results given in Table 2, we must consider the theory to be applied.

We may explain the hydrogen ion concentrations of the solutions under the assumption that the following three equilibria take place



and



In these schemes the water of hydration has been omitted. The mass action constants are

$$k_1 = \frac{(\text{CuOH}^+) (\text{H}^+)}{(\text{Cu}^{++})}, \quad (8)$$

$$k_2 = \frac{(\text{Cu}_2\text{OH}^{+++}) (\text{H}^+)}{(\text{Cu}^{++})^2}, \quad (9)$$

and

$$k_3 = \frac{(\text{Cu}_2\text{O}^{++}) (\text{H}^+)^2}{(\text{Cu}^{++})^2}. \quad (10)$$

We consider a series of measurements in a molar cupric nitrate. The cupric nitrate has an acid deficiency of δ equivalents per mole. x molar sodium hydroxide, or $-x$ molar nitric acid, has been added. It follows from the theory that

$$\begin{aligned} (\text{H}^+) + x + a\delta &= (\text{CuOH}^+) + (\text{Cu}_2\text{OH}^{+++}) + 2(\text{Cu}_2\text{O}^{++}) \\ &= k_1 \frac{(\text{Cu}^{++})}{(\text{H}^+)} + k_2 \frac{(\text{Cu}^{++})^2}{(\text{H}^+)} + 2k_3 \frac{(\text{Cu}^{++})^2}{(\text{H}^+)^2}, \end{aligned}$$

which may also be written as follows

$$\frac{[(\text{H}^+) + x] (\text{H}^+)}{(\text{Cu}^{++})} + \frac{a\delta (\text{H}^+)}{(\text{Cu}^{++})} = k_1 + k_2 (\text{Cu}^{++}) + 2k_3 \frac{(\text{Cu}^{++})}{(\text{H}^+)}. \quad (11)$$

If the decrease in cupric ion concentration owing to the reactions 5 to 7 is denoted by \mathcal{A} , we get

$$(\text{Cu}^{++}) = a - \mathcal{A} \quad (12)$$

and

$$\begin{aligned} \mathcal{A} &= (\text{CuOH}^+) + 2(\text{Cu}_2\text{OH}^{+++}) + 2(\text{Cu}_2\text{O}^{++}) \\ &= (\text{H}^+) + x + a\delta + (\text{Cu}_2\text{OH}^{+++}) \\ &= (\text{H}^+) + x + a\delta + k_2 \frac{(\text{Cu}^{++})^2}{(\text{H}^+)}. \end{aligned}$$

An approximation leads to

$$\mathcal{A} = (\text{H}^+) + x + k_2 \frac{a^2}{(\text{H}^+)}. \quad (13)$$

From the equations 11 and 12 we obtain

$$k = (k_1 + ak_2) + 2k_3 \frac{(\text{Cu}^{++})}{(\text{H}^+)}, \quad (14)$$

where

$$k \equiv \frac{[(\text{H}^+) + x](\text{H}^+)}{(\text{Cu}^{++})} + \frac{\alpha\delta(\text{H}^+)}{(\text{Cu}^{++})} + k_2\mathcal{A}. \quad (15)$$

In order to test the theory, we first compute \mathcal{A} from equation 13 and a preliminary value of k_2 , and then (Cu^{++}) from equation 12. k of equation 15 is a sum of three terms, the first and most important of which may now be calculated. The third term is a very small correction and may be estimated from a preliminary value of k_2 . The relative importance of the second term, which contains the unknown δ , decreases when x increases within a series of measurements at constant a . If we as a first approximation put $\delta = 0$, preliminary values of k may be calculated and plotted against $\frac{(\text{Cu}^{++})}{(\text{H}^+)}$ (the solid circles in Fig. 3). According to equation 14, the true value of k leads to a straight line when a is constant; actually, the points computed under the assumption that $\delta = 0$ fall on a curve which approaches a straight line with increasing x . Generally, the approach is the more rapid the smaller a . By the method of trial and error, we find the value of δ for which a rectilinear plot is obtained (open circles in Fig. 3). From the straight line, the numerical values of the two constants $(k_1 + ak_2)$ and $2k_3$ in equation 14 may be determined. The series chosen as an example in Fig. 3 is one in which the deviation from the straight line of the points computed for $\delta = 0$ is fairly great. Only for two of the 15 series, namely when $a = 1$ and $a = 0.5$, the deviations were greater.

The results of the measurements and computations are presented in Tables 2 and 3. The numbers of the series, given in the first column of Table 2, indicate the chronological order. It is seen that δ , which is given in the same column, increases with time. When the same standard solution of cupric nitrate was used for preparing the solutions in several series, the same δ was found. x , the concentration of sodium hydroxide added, is given in the second column. Nitric acid is reckoned as sodium hydroxide with negative sign. An asterisk denotes the addition of so much base that a precipitate was formed during the

measurement. In these cases, the potential decreased gradually after having been constant for a time. E , given in the next column, is the electromotive force in volts after subtraction of the asymmetry potential. Therefore, E may be considered the e. m. f. of the combination stated in scheme 2. $-\log(\text{H}^+)$, presented in the fourth column, has been computed by means of

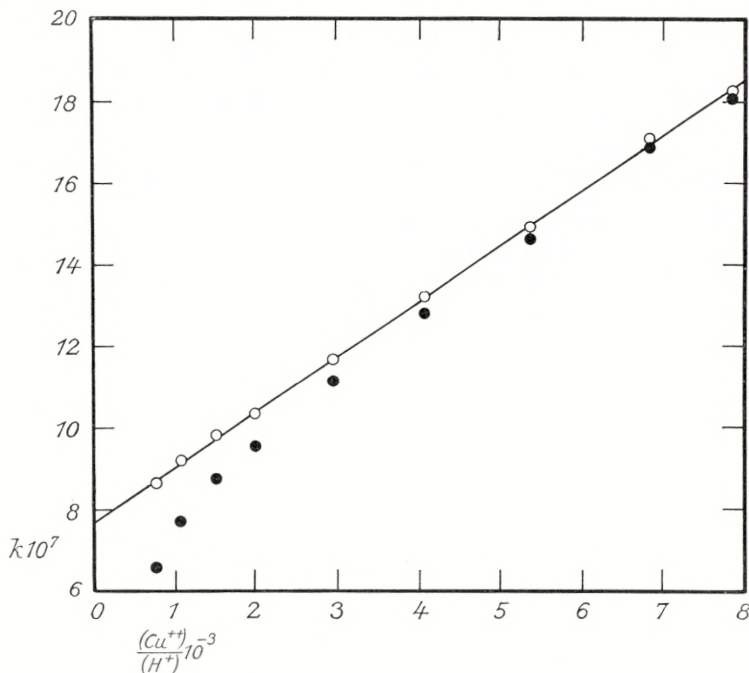


Fig. 3. Series 9: 0.200 m $\text{Cu}(\text{NO}_3)_2$. Solid circles: $\delta = 0$. Open circles: $\delta = 8 \times 10^{-5}$.

equation 4 and the values of A_0 in Table 1. The computations of $\frac{(\text{Cu}^{++})}{(\text{H}^+)}$ and k , given in the next two columns, were carried out as previously explained.

When k is plotted against $\frac{(\text{Cu}^{++})}{(\text{H}^+)}$, the points for each series fall close to a straight line, in agreement with equation 14. The constants $(k_1 + ak_2)$ and k_3 determined from the plot are shown in the fourth and sixth column of Table 3. By means of these values and equation 14, the values of k given in the last column of Table 2 were calculated. The agreement between k found from the measurement and k calculated from formula 14 is satisfactory. Only in the least acid solutions of some of the

Table 2.

	$x \cdot 10^4$	E	$-\log(\text{H}^+)$	$\frac{(\text{Cu}^{++})}{(\text{H}^+)} 10^{-3}$	$k \cdot 10^7$ found	$k \cdot 10^7$ formula
Series 12 0.9994 m $\text{Cu}(\text{NO}_3)_2$ $d = 18 \times 10^{-5}$	-10.23	0.0342	2.900	0.793	5.25	5.25
	- 5.09	0.0424	3.042	1.100	5.27	5.28
	0.00	0.0509	3.190	1.546	5.35	5.32
	3.39	0.0567	3.290	1.945	5.31	5.36
	7.43	0.0626	3.392	2.459	5.42	5.41
	12.56	0.0693	3.508	3.208	5.46	5.48
	18.57	0.0754	3.614	4.093	5.59	5.57
	*25.50	0.0793	3.681	4.771	(6.18)	5.64
Series 11 0.5031 m $\text{Cu}(\text{NO}_3)_2$ $d = 16 \times 10^{-5}$	-10.19	0.0494	2.979	0.479	2.32	2.35
	- 4.085	0.0658	3.263	0.921	2.36	2.40
	- 2.034	0.0725	3.379	1.203	2.46	2.43
	0.000	0.0795	3.500	1.589	2.50	2.47
	1.390	0.0842	3.581	1.914	2.52	2.51
	3.768	0.0913	3.704	2.539	2.58	2.58
	7.038	0.0990	3.838	3.454	2.70	2.69
	11.70	0.1072	3.980	4.783	2.84	2.84
*18.66	0.1155	4.123	6.636	3.06	3.06	
Series 9 0.2000 m $\text{Cu}(\text{NO}_3)_2$ $d = 8 \times 10^{-5}$	- 2.081	0.0897	3.587	0.772	0.864	0.872
	- 1.036	0.0980	3.730	1.073	0.920	0.913
	0.000	0.1066	3.879	1.512	0.981	0.973
	0.932	0.1138	4.004	2.015	1.034	1.041
	2.624	0.1235	4.172	2.963	1.169	1.170
	4.688	0.1314	4.309	4.057	1.320	1.319
	7.465	0.1385	4.432	5.376	1.491	1.498
	11.19	0.1446	4.537	6.829	1.710	1.696
13.87	0.1481	4.598	7.848	1.826	1.834	
*16.32						
Series 5 0.1000 m $\text{Cu}(\text{NO}_3)_2$ $d = 3 \times 10^{-5}$	- 2.058	0.0944	3.653	0.450	0.434	0.433
	- 1.033	0.1069	3.870	0.741	0.468	0.477
	- 0.520	0.1145	4.001	1.001	0.507	0.516
	0.000	0.1218	4.128	1.341	0.578	0.567
	0.704	0.1307	4.282	1.911	0.659	0.652
	1.641	0.1388	4.422	2.635	0.778	0.760
	2.797	0.1460	4.547	3.508	0.889	0.891
	4.243	0.1519	4.649	4.430	1.018	1.029
	5.854	0.1564	4.727	5.291	1.151	1.158
	9.305	0.1630	4.841	6.852	1.387	1.391
*12.54	0.1672	4.914	8.077	1.578	1.574	

Table 2 (continued).

	$x \cdot 10^4$	E	$-\log(\text{H}^+)$	$\frac{(\text{Cu}^{++})}{(\text{H}^+)} \cdot 10^{-3}$	$k \cdot 10^7$ found	$k \cdot 10^7$ formula
Series 6 0.05000 m $\text{Cu}(\text{NO}_3)_2$ + 0.05000 m $\text{Ba}(\text{NO}_3)_2$ $d = 3 \times 10^{-5}$	- 1.028	0.1113	3.946	0.442	0.270	0.268
	- 0.518	0.1228	4.145	0.698	0.306	0.306
	0.000	0.1356	4.367	1.163	0.383	0.376
	0.479	0.1454	4.536	1.714	0.459	0.458
	1.180	0.1544	4.692	2.452	0.571	0.569
	2.122	0.1621	4.826	3.330	0.688	0.700
	3.293	0.1678	4.924	4.164	0.825	0.825
	4.663	0.1726	5.007	5.024	0.953	0.954
7.014	0.1782	5.104	6.250	1.140	1.137	
Series 7 0.02000 m $\text{Cu}(\text{NO}_3)_2$ + 0.08000 m $\text{Ba}(\text{NO}_3)_2$ $d = 3 \times 10^{-5}$	- 0.518	0.1286	4.245	0.351	0.162	0.165
	- 0.210	0.1423	4.483	0.608	0.206	0.203
	0.000	0.1522	4.654	0.900	0.253	0.247
	0.421	0.1667	4.905	1.602	0.345	0.352
	0.935	0.1756	5.059	2.277	0.452	0.453
	1.403	0.1807	5.148	2.788	0.531	0.529
	2.336	0.1876	5.267	3.650	0.658	0.658
	4.144	0.1954	5.402	4.933	0.851	0.850
Series 10 0.05000 m $\text{Cu}(\text{NO}_3)_2$ $d = 8 \times 10^{-5}$	- 1.036	0.1115	3.952	0.447	0.271	0.280
	- 0.518	0.1231	4.153	0.711	0.316	0.323
	- 0.262	0.1292	4.259	0.907	0.363	0.355
	0.000	0.1353	4.364	1.155	0.409	0.395
	0.476	0.1448	4.529	1.687	0.482	0.481
	1.178	0.1537	4.683	2.401	0.595	0.596
	2.339	0.1623	4.832	3.376	0.749	0.754
	3.970	0.1694	4.955	4.465	0.925	0.930
	6.262	0.1756	5.062	5.684	1.126	1.127
9.367	0.1811	5.158	7.043	1.349	1.346	
*11.67	0.184	5.208	7.868	(1.500)	1.479	
Series 8 0.02000 m $\text{Cu}(\text{NO}_3)_2$ $d = 3 \times 10^{-5}$	- 0.520	0.1270	4.240	0.347	0.176	0.178
	- 0.262	0.1377	4.425	0.532	0.225	0.215
	0.000	0.1499	4.636	0.864	0.274	0.271
	0.377	0.1623	4.851	1.415	0.370	0.371
	0.836	0.1710	5.002	1.998	0.471	0.476
	1.401	0.1775	5.114	2.580	0.575	0.581
	2.102	0.1828	5.206	3.177	0.684	0.689
	3.040	0.1877	5.291	3.844	0.807	0.809
	4.204	0.1920	5.365	4.531	0.939	0.933
	5.599	0.1960	5.435	5.282	1.070	1.069
	8.178	0.2010	5.521	6.352	(1.296)	1.262
*9.799	0.2036	5.566	6.985	(1.411)	1.376	

Table 2 (continued).

	$x \cdot 10^4$	E	$-\log(\text{H}^+)$	$\frac{(\text{Cu}^{++})}{(\text{H}^+)} \cdot 10^{-3}$	$k \cdot 10^7$ found	$k \cdot 10^7$ formula
Series 1 0.00991 m $\text{Cu}(\text{NO}_3)_2$ $\delta = 3 \times 10^{-5}$	-0.414	0.1324	4.351	0.222	0.157	0.152
	-0.209	0.1450	4.569	0.367	0.174	0.181
	0.000	0.1593	4.817	0.649	0.240	0.236
	0.237	0.1716	5.030	1.058	0.315	0.316
	0.582	0.1815	5.202	1.565	0.414	0.414
	1.161	0.1906	5.359	2.236	0.540	0.545
	1.753	0.1959	5.451	2.747	0.653	0.645
	2.322	0.2000	5.522	3.216	0.733	0.737
Series 2 0.004985 m $\text{Cu}(\text{NO}_3)_2$ + 0.00493 m $\text{Ba}(\text{NO}_3)_2$ $\delta = 3 \times 10^{-5}$	-0.139	0.1553	4.748	0.279	0.147	0.139
	-0.0564	0.1649	4.914	0.408	0.164	0.164
	0.000	0.1710	5.020	0.521	0.186	0.186
	0.158	0.1839	5.243	0.868	0.250	0.253
	0.393	0.1940	5.418	1.294	0.334	0.336
	0.781	0.2034	5.581	1.868	0.433	0.447
	1.171	0.2085	5.669	2.269	0.526	0.525
	1.550	0.2123	5.735	2.621	0.600	0.593
Series 3 0.001980 m $\text{Cu}(\text{NO}_3)_2$ + 0.00793 m $\text{Ba}(\text{NO}_3)_2$ $\delta = 3 \times 10^{-5}$	-0.0423	0.1765	5.115	0.258	0.136	0.129
	0.0000	0.1850	5.262	0.361	0.153	0.150
	0.1331	0.2028	5.570	0.730	0.220	0.222
	0.2918	0.2128	5.744	1.080	0.287	0.290
	0.5278	0.2211	5.887	1.484	0.364	0.367
	0.8178	0.2273	5.995	1.874	0.442	0.446
	1.162	0.2322	6.080	2.238	0.523	0.517
	1.779	0.2389	6.196	2.826	0.632	0.632
	2.330	0.2430	6.267	3.226	(0.725)	0.710
3.497	0.2498	6.385	3.951	(0.887)	0.852	
Series 4 0.001001 m $\text{Cu}(\text{NO}_3)_2$ + 0.00891 m $\text{Ba}(\text{NO}_3)_2$ $\delta = 3 \times 10^{-5}$	-0.0521	0.1799	5.174	0.149	0.102	0.104
	-0.0211	0.1891	5.333	0.215	0.119	0.117
	0.0000	0.1954	5.442	0.276	0.132	0.129
	0.0884	0.2138	5.761	0.571	0.186	0.187
	0.1802	0.2231	5.922	0.820	0.235	0.235
	0.2918	0.2298	6.038	1.060	0.284	0.282
	0.5231	0.2384	6.187	1.458	0.363	0.360
	0.8161	0.2456	6.312	1.883	0.436	0.443
	1.392	0.2530	6.440	2.371	(0.589)	0.539
	2.334	0.2624	6.603	3.074	(0.761)	0.676
*3.488	0.264	6.63	2.782	(1.256)	0.619	
*4.617	0.268	6.70	2.695	(1.715)	0.602	

Table 2 (continued).

	$x \cdot 10^4$	E	$-\log(\text{H}^+)$	$\frac{(\text{Cu}^{++})}{(\text{H}^+)} \cdot 10^{-3}$	$k \cdot 10^7$ found	$k \cdot 10^7$ formula
Series 15 0.005007 m $\text{Cu}(\text{NO}_3)_2$ $d = 18 \times 10^{-5}$	-0.1027	0.1600	4.843	0.349	0.143	0.153
	0.0000	0.1711	5.035	0.542	0.187	0.192
	0.1451	0.1823	5.229	0.845	0.253	0.254
	0.3855	0.1924	5.404	1.258	0.345	0.337
	0.7067	0.2001	5.538	1.702	0.437	0.428
	1.172	0.2071	5.659	2.227	0.540	0.534
	1.890	0.2139	5.777	2.880	0.666	0.667
	2.930	0.2203	5.888	3.638	0.812	0.821
	4.675	0.2272	6.007	4.605	1.020	1.017
	7.026	0.2335	6.116	5.609	(1.258)	1.221
	9.364	0.2383	6.199	6.417	(1.464)	1.385
*11.71	0.2403	6.234	6.553	(1.793)	1.412	
Series 14 0.002005 m $\text{Cu}(\text{NO}_3)_2$ $d = 18 \times 10^{-5}$	-0.1034	0.1635	4.918	0.166	0.127	0.136
	-0.0515	0.1721	5.067	0.233	0.162	0.150
	0.0000	0.1817	5.233	0.342	0.182	0.174
	0.1197	0.1977	5.510	0.644	0.240	0.238
	0.2936	0.2093	5.711	1.014	0.313	0.318
	0.5458	0.2180	5.862	1.418	0.398	0.405
	0.8330	0.2238	5.962	1.760	0.481	0.478
	1.167	0.2290	6.052	2.126	0.555	0.557
	1.708	0.2346	6.149	2.580	0.667	0.665
	2.335	0.2394	6.232	3.019	(0.777)	0.749
	3.507	0.2464	6.354	3.727	(0.944)	0.901
	4.679	0.2510	6.433	4.154	(1.129)	0.993
	7.017	0.2527	6.463	3.776	(1.862)	0.912
9.335	0.2544	6.492	3.320	(2.817)	0.814	
*13.62	0.2537	6.480	1.931	(7.059)	0.515	
Series 13 0.001007 m $\text{Cu}(\text{NO}_3)_2$ $d = 18 \times 10^{-5}$	-0.0515	0.1769	5.161	0.146	0.132	0.134
	-0.0211	0.1851	5.303	0.202	0.151	0.148
	0.0000	0.1907	5.400	0.252	0.163	0.159
	0.1056	0.2106	5.745	0.553	0.227	0.231
	0.2327	0.2213	5.930	0.836	0.294	0.297
	0.4369	0.2304	6.088	1.177	0.380	0.378
	0.7384	0.2387	6.231	1.585	0.471	0.474
	1.167	0.2459	6.356	2.016	0.582	0.576
	1.759	0.2529	6.477	2.492	(0.709)	0.688
	2.334	0.2578	6.562	2.818	(0.831)	0.765
	3.258	0.2604	6.586	2.625	(1.244)	0.720
	4.672	0.2633	6.657	2.445	(1.914)	0.677
	5.850	0.2650	6.687	2.039	(2.873)	0.581
	7.001	0.2658	6.701	1.532	(4.574)	0.462
8.248	0.2653	6.692				
10.07	0.2657	6.699				
*15.12	0.2725	6.817				

series, values of k considerably greater than those corresponding to formula 14 are found. Such values are placed in parentheses. Especially in the most dilute solutions, which have the smallest hydrogen ion concentrations, the discrepancy may be due to an acid dissociation of the cupric ion which exceeds that expressed in the schemes 5 to 7. Thus, an ion of the composition $\text{Cu}_2\text{O}(\text{OH})^+$ may possibly be formed. The discrepancy may also be due to a beginning precipitation. This question has not been examined more closely.

We now consider Table 3 where $(k_1 + ak_2)$ and k_3 found from the measurements are enumerated. With changing a , we may expect a variation of k_1 , k_2 , and k_3 owing to the change of the ionic strength μ . $(k_1 + ak_2)$ varies also as a direct consequence of the change of a . Within each of the two groups comprising series 5—7 and series 1—4, a varies, while μ is kept constant by replacing the cupric nitrate by barium nitrate. As may be seen from the table, k_3 is constant within each group, while $(k_1 + ak_2)$ varies linearly with a . From the latter relation we find

$$k_1 = 4.8 \times 10^{-9} \text{ and } k_2 = 3.16 \times 10^{-7}, \text{ when } \mu = 0.300,$$

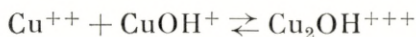
$$k_1 = 7.6 \times 10^{-9} \text{ and } k_2 = 2.1 \times 10^{-7}, \text{ when } \mu = 0.030.$$

Table 3.

Comparison between the mass action constants found from the measurements and those calculated from formulae 16—18.

Series no.	$\text{Cu}(\text{NO}_3)_2$ molarity	$\text{Ba}(\text{NO}_3)_2$ molarity	$(k_1 + ak_2) 10^7$ found	$(k_1 + ak_2) 10^7$ formulae	$k_3 10^{12}$ found	$k_3 10^{12}$ formula
12	0.9994	...	5.17	5.17	4.9	5.9
11	0.5031	...	2.29	2.29	5.8	6.3
9	0.2000	...	0.767	0.774	6.8	6.9
5	0.1000	...	0.366	0.364	7.48	7.45
6	0.05000	0.05000	0.202	0.206	7.48	7.45
7	0.02000	0.08000	0.112	0.111	7.48	7.45
10	0.05000	...	0.208	0.193	8.08	8.09
8	0.02000	...	0.115	0.114	9.03	9.02
1	0.00991	...	0.109	0.097	9.76	9.73
2	0.004985	0.00493	0.085	0.086	9.70	9.73
3	0.001980	0.00793	0.079	0.080	9.78	9.73
4	0.001001	0.00891	0.075	0.078	9.78	9.73
15	0.005007	...	0.082	0.093	10.2	10.4
14	0.002005	...	0.100	0.095	10.8	11.1
13	0.001007	...	0.100	0.098	11.8	11.5

In the other cases, where μ changes from series to series, it is more difficult to estimate the separate values of k_1 and k_2 . However, it follows from the equilibrium



that an increase in the concentration of cupric salt will favour the ion $\text{Cu}_2\text{OH}^{+++}$ at the expense of CuOH^+ . In more concentrated solutions (when $a > 0.1$), the concentration of the ion CuOH^+ is so small that k_2 may be calculated with sufficient accuracy from $(k_1 + ak_2)$, if k_1 can be estimated roughly. On the other hand, in the more dilute solutions (when $a < 0.01$), the ion $\text{Cu}_2\text{OH}^{+++}$ is of so little importance that k_1 may be calculated from $(k_1 + ak_2)$ and a very rough value of k_2 .

We have now found k_1 , when $0.1 \geq a \geq 0.001$; k_2 , when $1 \geq a \geq 0.01$; and k_3 in the entire interval $1 \geq a \geq 0.001$. The values found for the three equilibrium constants may be expressed by means of the following formulae which, for very dilute solution, agree with the requirements of the DEBYE-HÜCKEL law:

$$-\log k_1 = 7.97 + \frac{0.996 \sqrt{\mu}}{1 + 1.0 \sqrt{\mu}} \quad (16)$$

$$-\log k_2 = 6.819 - \frac{0.996 \sqrt{\mu}}{1 + 1.3 \sqrt{\mu}} \quad (17)$$

$$-\log k_3 = 10.890 + \frac{0.996 \sqrt{\mu}}{1 + 2.36 \sqrt{\mu}} \quad (18)$$

The formulae contain two empirical constants each, while the third constant 0.996 is deduced from the Debye-Hückel theory. The values of $(k_1 + ak_2)$ calculated by means of formulae 16—17 are given in the fifth column of Table 3, while k_3 calculated from formula 18 is to be found in the last column of the table. A comparison of the calculated values with those obtained from the measurements shows good agreement. Only the values of k_3 for the two most concentrated solutions deviate appreciably from those calculated on the basis of formula 18.

If we extrapolate to the ionic strength 0 by means of the formulae 16—18, we obtain

$$k_1^0 = 1.07 \times 10^{-8},$$

$$k_2^0 = 1.52 \times 10^{-7},$$

and

$$k_3^0 = 1.29 \times 10^{-11}.$$

We have seen that the assumption of the three equilibria 5—7 offers a satisfactory explanation for the hydrogen ion concentration of the cupric nitrate solutions. The existence of the three equilibria may be proved more directly in the following way: It has already been shown that the experimental values of k , defined by equation 15, satisfy equation 14, when the hydrogen and the cupric ion concentrations vary, if only the ionic strength is kept constant. k_1 , k_2 , and k_3 are constants, but they are not assumed to be equilibrium constants; it follows from this result that

$$(\text{H}^+) + x + a\delta = k_1 \frac{(\text{Cu}^{++})}{(\text{H}^+)} + k_2 \frac{(\text{Cu}^{++})^2}{(\text{H}^+)} + 2k_3 \frac{(\text{Cu}^{++})^2}{(\text{H}^+)^2},$$

when (H^+) and (Cu^{++}) vary, while μ is constant. The only possible explanation of this result is that hydrogen ions are produced by the three equilibria 5—7. The empirical constants k_1 , k_2 , and k_3 are the mass action constants of the equilibria.

The equilibrium constant k_1 is the acid strength of the hydrated cupric ion. At infinite dilution, it is 1.07×10^{-8} . According to J. BJERRUM⁽¹³⁾, the cupric ion in aqueous solution is a planar tetraquo ion which, through a more or less stable intermediate form, is in equilibrium with an octahedral hex-aquo ion.

From the equilibria 6 and 7 we derive

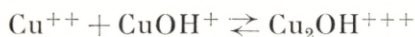


The mass action constant $\frac{k_3}{k_2}$ of this equilibrium is the acid strength of the ion $\text{Cu}_2\text{OH}^{+++}$. At infinite dilution, it is 8.5×10^{-5} . In the ion $\text{Cu}_2\text{OH}^{+++}$, the two copper atoms are bound together by the oxygen atom of the hydroxyl group. To the copper atoms are furthermore bound a certain number of water molecules which, as usual, have not been written. It is impossible to say whether the ion $\text{Cu}_2\text{OH}^{+++}$ splits off the hydrogen ion

from the hydroxyl group or from one of the water molecules. The question whether the ion Cu_2O^{++} contains a naked oxygen bridge, or a hydroxyl bridge and a hydroxyl group directly bound to one of the copper atoms, is therefore left open. The acid strength of the trivalent ion $\text{Cu}_2\text{OH}^{+++}$ is about 8000 times that of the divalent ion Cu^{++} . A change in this direction might be expected on account of the electrostatic effect of the positive charges.

Owing to the fact that the maximum co-ordination number of oxygen is 3, the ion $\text{Cu}_2\text{OH}^{+++}$ cannot take up another hydrogen ion. A dimeric cupric ion Cu_2^{++++} (omitting the water of hydration), therefore, does not exist.

From the equilibria 5 and 6 we derive



with the equilibrium constant

$$\frac{k_2}{k_1} = \frac{(\text{Cu}_2\text{OH}^{+++})}{(\text{Cu}^{++})(\text{CuOH}^+)}$$

which becomes equal to 14.2 at infinite dilution.

Table 4 contains the equilibrium constants k_1 , k_2 , and k_3 for cupric nitrate solutions, calculated from the formulae 16—18, at round values of the concentration α molar. The fifth

Table 4.

The equilibrium constants, hydrogen ion concentrations, and degrees of "hydrolysis" α for pure aqueous solutions of cupric nitrate (α molar), calculated by means of the formulae 16—18.

α	$k_1 10^9$	$k_2 10^7$	$k_3 10^{12}$	$-\log(\text{H}^+)$	pH	$\alpha 10^3$
1.000	2.51	5.15	5.90	3.137	2.85	1.43
0.500	3.02	4.49	6.25	3.456	3.36	1.34
0.200	3.98	3.67	6.87	3.856	3.85	1.22
0.100	4.79	3.16	7.45	4.125	4.13	1.17
0.0500	5.62	2.74	8.09	4.362	4.36	1.18
0.0200	6.76	2.32	9.02	4.634	4.62	1.36
0.0100	7.59	2.10	9.73	4.823	4.79	1.64
0.00500	8.32	1.93	10.35	5.004	4.95	2.08
0.00200	9.12	1.78	11.1	5.234	5.17	2.98
0.00100	9.55	1.71	11.5	5.405	5.33	3.99
0	10.7	1.52	12.9			

column shows $-\log(\text{H}^+)$ for a pure a molar cupric nitrate solution, calculated from the constants given in the table.

In their papers, most of the earlier investigators give the SØRENSEN pH of the cupric salt solution. Generally, the liquid-liquid potential is partly eliminated by means of a bridge solution, such as concentrated potassium chloride or ammonium nitrate, and the remaining liquid-liquid potential is neglected. The results of the present measurements may be expressed in the same way if, instead of formula 4, the following formula is used

$$\text{pH} = 2.022 + 17.32 E, \quad (19)$$

2.022 being the pH of the reference solution: 0.01000 m HCl + 0.0900 m NaCl. If E is eliminated from the equations 4 and 19, we obtain

$$\text{pH} = -\log(\text{H}^+) + 2.022 - A_0.$$

The pH values given in the sixth column of Table 4 have been calculated from $-\log(\text{H}^+)$ by means of this formula. When these values are plotted against $\log a$, the curve of Fig. 1 is obtained.

The degrees of reaction ("hydrolysis") of the cupric ions according to each of the three equilibria 5—7 are

$$\alpha_1 = \frac{(\text{CuOH}^+)}{a} = \frac{k_1(\text{Cu}^{++})}{a(\text{H}^+)},$$

$$\alpha_2 = \frac{2(\text{Cu}_2\text{OH}^{+++})}{a} = \frac{2k_2(\text{Cu}^{++})^2}{a(\text{H}^+)},$$

and

$$\alpha_3 = \frac{2(\text{Cu}_2\text{O}^{++})}{a} = \frac{2k_3(\text{Cu}^{++})^2}{a(\text{H}^+)^2},$$

respectively. The total degree of reaction is $\alpha = \alpha_1 + \alpha_2 + \alpha_3$. The degrees of reaction have been calculated for pure cupric nitrate solutions by means of the equilibrium constants and the hydrogen ion concentrations given in Table 4. α is to be found in the last column of Table 4. When α_1 , α_2 , α_3 , and α are plotted against $\log a$, the four curves of Fig. 4 are obtained. The figure shows how the relative importance of the three equilibria varies with the concentration of the cupric nitrate. Roughly speaking, the concentrations of the three ions are of

the same order of magnitude, when $0.1 > a > 0.01$. In more concentrated solutions, the ion $\text{Cu}_2\text{OH}^{+++}$ gains more and more predominance, while the ion Cu_2O^{++} and, first and foremost, the ion CuOH^+ gradually disappear. In solutions more dilute

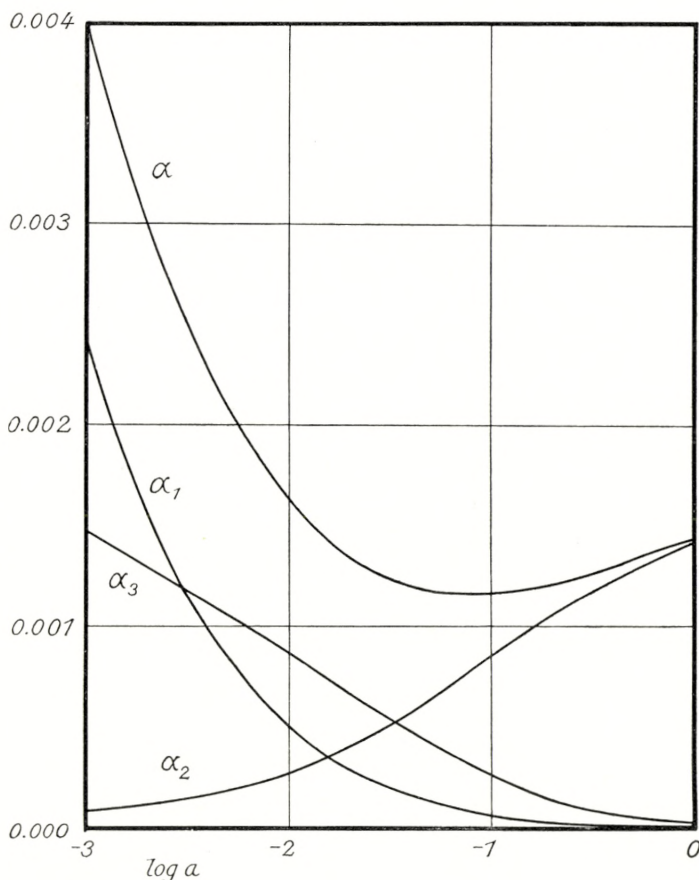


Fig. 4. The degrees of reaction ("hydrolysis") of the cupric ion in an a molar solution of cupric nitrate.

$$\alpha_1 = (\text{CuOH}^+)/a \quad \alpha_2 = 2(\text{Cu}_2\text{OH}^{+++})/a \quad \alpha_3 = 2(\text{Cu}_2\text{O}^{++})/a$$

$$\alpha = \alpha_1 + \alpha_2 + \alpha_3.$$

than 0.01 molar, the form $\text{Cu}_2\text{OH}^{+++}$ gradually disappears. At sufficiently high dilution, only the monomeric ion CuOH^+ will remain.

I wish to express my thanks to the head of the laboratory, Professor NIELS BJERRUM, for advice and kind interest in my work.

Summary.

The hydrogen ion concentration of aqueous solutions of cupric nitrate was measured at 18.0° C. by means of the glass electrode. The range of concentrations examined extended from 0.001 to 1 molar cupric nitrate. The examination included both aqueous solutions and solutions to which small concentrations of nitric acid or sodium hydroxide had been added.

The measurements show that the following equilibria take place



and



The mass action constants of the equilibria have been computed.

From the Chemical Laboratory
of the Royal Veterinary and Agricultural College.
Copenhagen, Denmark.

References.

- (1) G. CARRARA and G. B. VESPIGNANI, *Gazz. chim. ital.* **30** (1900) 35.
 - (2) G. KULLGREN, *Z. physik. Chem.* **85** (1913) 466.
 - (3) J. B. O'SULLIVAN, *Trans. Faraday Soc.* **21** (1925) 319.
 - (4) H. B. FRIEDMAN and J. A. STOKES, *J. Amer. Chem. Soc.* **61** (1939) 118.
 - (5) M. QUINTIN, *J. chim. phys.* **24** (1927) 712.
 - (6) V. ČUPR and J. ŠIRŮČEK, *Collection Czechoslovak Chem. Commun.* **9** (1937) 68.
 - (7) W. S. HUGHES, *J. Chem. Soc.* (1928) 491.
 - (8) J. A. CRANSTON and H. F. BROWN, *J. Roy. Tech. Coll. (Glasgow)* **4** (1937) 54.
 - (9) G. JANDER and H. MÖHR, *Z. physik. Chem. A* **190** (1942) 81.
 - (10) J. BJERRUM, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* **XI** (1931) 5.
 - (11) S. KÜHNEL HAGEN, *Z. anal. Chem.* **117** (1939) 26, *Kemisk Maanedssblad* **23** (1942) 38.
 - (12) D. A. MACINNES and M. DOLE, *J. Amer. Chem. Soc.* **52** (1930) 29.
 - (13) J. BJERRUM, *Metal Ammine Formation in Aqueous Solution. Dissertation.* Copenhagen. (1941) 107.
-

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 8

DETERMINATION OF
PARTICLE SIZE BY THE X-RAY
POWDER METHOD

BY

AKSEL TOVBORG JENSEN



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

INTRODUCTION

It is a well-known fact that the lines of an X-ray powder diagram are broad and diffuse when the elementary particles of the powder are very small¹⁾. Several physicists, especially SCHERRER²⁾³⁾, LAUE⁴⁾⁵⁾, and BRILL⁶⁾⁷⁾⁸⁾⁹⁾, have devised methods for the determination of particle dimensions in a powder from the "line breadth" of its X-ray diagram. The new method to be described in the following is more accurate than SCHERRER's and, compared with the methods described by LAUE and BRILL, it has the advantage that diagrams taken with the ordinary technique (relatively short exposures) may be employed. Unlike BRILL's methods, it does not require powder specimens of no, or of infinitely large absorption; neither very narrow slits in the primary beam, nor a very divergent primary beam are needed. The method is applicable in cases where the absorption in the specimen is incalculable as, for instance, when the Seemann-Bohlin focussing camera is applied, or when specimens in glass capillaries are used in the ordinary Debye camera.

The method is based upon a comparison of the line breadths in two diagrams; the first, a diagram of the powder whose particle size is to be determined, the second, a diagram of the same chemical compound in a suitably large-grained state. Otherwise, the two specimens should be identical with respect to dimensions, density, etc., and both diagrams should be taken with the same arrangement of apparatus.

The ideal case.

The following formula has been proposed for the breadth of a Debye line in a strongly idealized case,

$$b_{hkl} = \frac{k \cdot \lambda}{t \cdot \cos \theta}, \quad \text{I}$$

where θ is the glancing angle, b is the breadth of the line defined as the angular range between two points where the intensity has fallen to half of its maximum value, λ is the wave length, t is the linear extension of the reflecting crystal, and k a constant. If λ and t are measured in the same units and b in radians, the numerical value of k is about 0.90^1 .

The formula I is valid

a) if the primary beam is parallel, infinitely narrow, and reflected by a single crystal whose linear extension is t , perpendicular to the reflecting plane (hkl);

b) for a line in a powder diagram under ideal experimental conditions, i. e. parallel primary beam, infinitely thin specimen, monodispersed powder, if

1) the crystals belong to the cubic system and are developed as cubes with a side t (LAUE⁴, SELJAKOW¹⁰),

2) the crystals are non-cubic, but are limited by the face (hkl) and are of a thickness t perpendicular to (hkl) (LAUE⁴). A large crystal under condition a) and a powder of fairly large particles under conditions b) will give infinitely narrow reflections.

The actual case.

For practical reasons, formula I is not applicable directly. It is impracticable to use specimens so thin and a primary beam so faintly divergent that the breadth of a line from a powder of large particles becomes infinitely thin compared with b_{hkl} in formula I. In order to secure a suitably short time of exposure, specimen dimensions and slit width should not be smaller than sufficient to ensure the resolution of the two lines that lie closest to each other in the diagram. In this case, the lines obtained are not very narrow. The intensity distribution in the lines depends in an intricate way on the dimensions of the slit and of the specimen and on the absorption in the specimen. Such an intensity distribution $I = f(\theta)$ (vide fig. 1) results from a sum of a large number of very narrow rectangles. A rectangle between $\theta = \theta_1$ and $\theta = \theta_1 + \Delta\theta_1$ is the non-absorbed part of the reflected intensity from a large number of small crystals which are oriented in such a way that the plane hkl is in reflecting

¹ SCHERRER²⁾³⁾ gives 0.94, SELJAKOW¹⁰⁾ 0.92, LAUE⁴⁾ 0.90, and W. L. BRAGG¹⁾ 0.89.

position, and so situated that the reflections fall between $\theta = \theta_1$ and $\theta = \theta_1 + \Delta\theta_1$ on the film. The height of the rectangle depends upon the degree of absorption of the primary and the reflected beam in the powder specimen. Let us imagine the large particles divided until every new elementary particle gives a reflection of finite breadth. The intensity which fell between θ_1

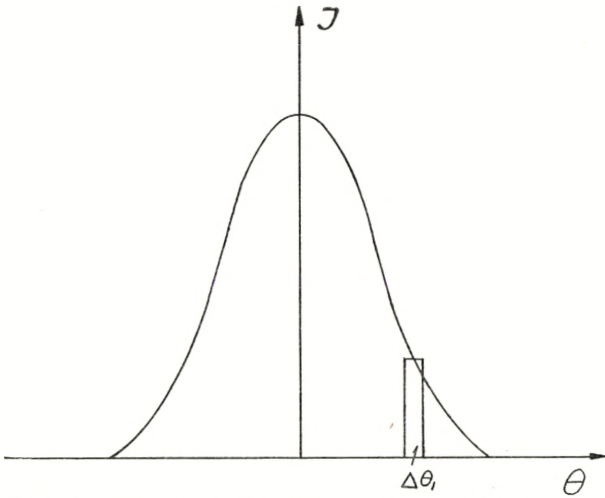


Fig. 1. Intensity distribution in a powder diagram line and one of the rectangles of which the curve is made up.

and $\theta_1 + \Delta\theta_1$ is now smeared out and no longer forms a narrow rectangle but a rounded peak with the intensity distribution $i(\theta_1, \theta)$. θ_1 is the middle of the peak, and θ the point where the intensity is $i(\theta_1, \theta)$. The breadth b of this peak obeys equation I.

We assume that the area below the curve $i(\theta_1, \theta)$ is equal to the area of the rectangle. The intensity in the entire line is also smeared out, and a new intensity distribution function $I^1(\theta)$ is obtained. $I^1(\theta)$ is somewhat broader and lower than $I(\theta)$.

Provided that $I(\theta)$ and $i(\theta_1, \theta)$ are Gaussian error functions

$$I(\theta) = I_0 \cdot e^{-H^2 \theta^2}$$

and

$$i(\theta_1, \theta) = i(\theta_1, \theta) \cdot e^{-h^2(\theta - \theta_1)^2},$$

it is not difficult to prove¹ that $I^1(\theta)$ is also a Gaussian function.

¹ The proof is not given here since its result is incorporated in a number of textbooks on mathematics.

Then, the breadths (B_1 , B and b) of the three curves are connected by the equation

$$B^2 = B_1^2 + b^2. \quad \text{II}$$

Hence $b = b_{hkl}$, which could not be obtained in direct experiments, may be found from equation II.

Formula II is different from a formula put forward without proof by SCHERRER²⁾³⁾ which in our notation reads $B_1 = B + b$.¹

The assumptions involved.

The intensity distribution in an X-ray reflection from a crystal consisting of N successive atomic planes obeys the equation

$$I/I_0 = \frac{\sin^2 \varphi}{\varphi^2}$$

where

$$\varphi = \frac{2 \pi N d \varepsilon \cos \theta}{\lambda} \quad (\text{BRAGG}^1);$$

in our previous notation

$$I/I_0 = \frac{\sin^2 \varphi}{\varphi^2}, \quad \text{III}$$

where $\varphi = c \cdot \theta$ and c is a constant.

Certainly, these equations do not constitute a Gaussian function, but the deviation from a Gaussian function is negligible in this connection, as may be seen from fig. 2 which shows the intensity function III and a Gaussian function of the same height and breadth.

MURDOCK¹¹⁾ examined the intensity distribution function also considering the influence of the crystal shape in a few simple cases. He obtained curves which all lay extraordinarily close to a Gaussian curve.

The small differences between intensity functions for differently developed crystallites (cubes, octahedra) were of the same magnitude as the differences between the intensity functions obtained

¹ (Note added after completion of the paper.)

A similar method for the determination of particle size could not be found in the literature available. However, in a private communication Professor GUNNAR HÄGG called my attention to the fact that the method here described and the argument applied in a number of papers on the breadth of spectral lines have several points of resemblance. Vide e. g. A. H. COMPTON and S. K. ALLISON, X-rays in theory and experiment, (London 1935) p. 740 etc.; BURGER and VAN CITTERT, Zs. f. Physik **44** (1927) 50.

and a Gaussian function of the same height and breadth. Since the shape of crystallites in a powder to be examined by X-rays is always unknown, and since their monodispersity is often questionable, the assumption of a Gaussian function for the distribution $i(\theta_1, \theta)$ is well warranted.

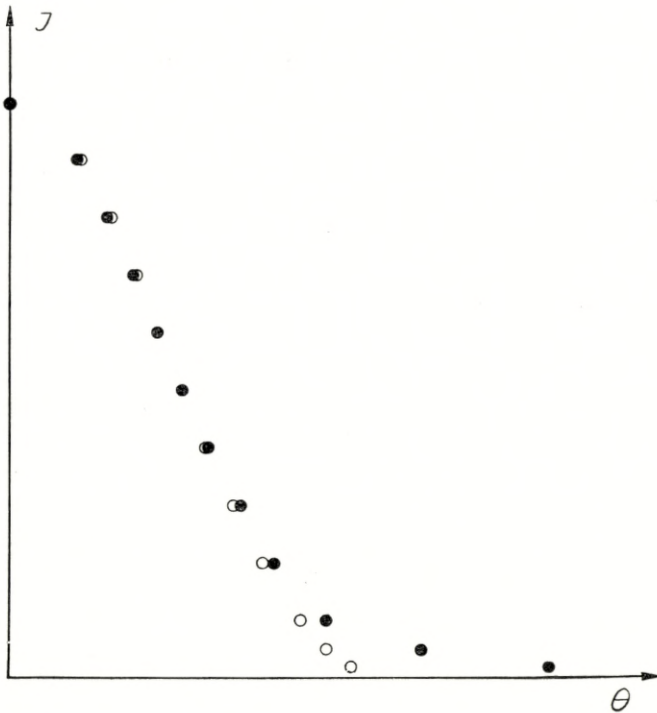


Fig. 2. Theoretical intensity distribution in a spectral line • and the corresponding Gaussian error curve ○.

Fig. 3 shows how closely the intensity distribution in a line of a powder diagram approximates a Gaussian function. Fig. 3 represents a photometer tracing of three lines in a diagram of CaCl_2 of medium particle size. The small circles lie on a Gaussian error curve of the same height and breadth as the second of these lines.

To the discussion following fig. 1 may be added: We presume that all particles of the powder are alike. The rectangle of breadth $\Delta\theta$ is a sum of terms $\tau \cdot f(\varphi, \omega, \alpha)$, where τ is the energy which would have been reflected from one particle if all other particles were absent and if no absorption of the primary or the reflected

beam had occurred. f is a coefficient of absorption including the absorption in the primary as well as in the reflected beam. (φ, ω, z) are the cylindrical coordinates in the powder rod. The area A_1 of the rectangle is

$$A_1 = \Delta\theta_1 \cdot I(\theta_1) = \sum r \cdot f(\varphi, \omega, z) = r \cdot \sum f(\varphi, \omega, z).$$

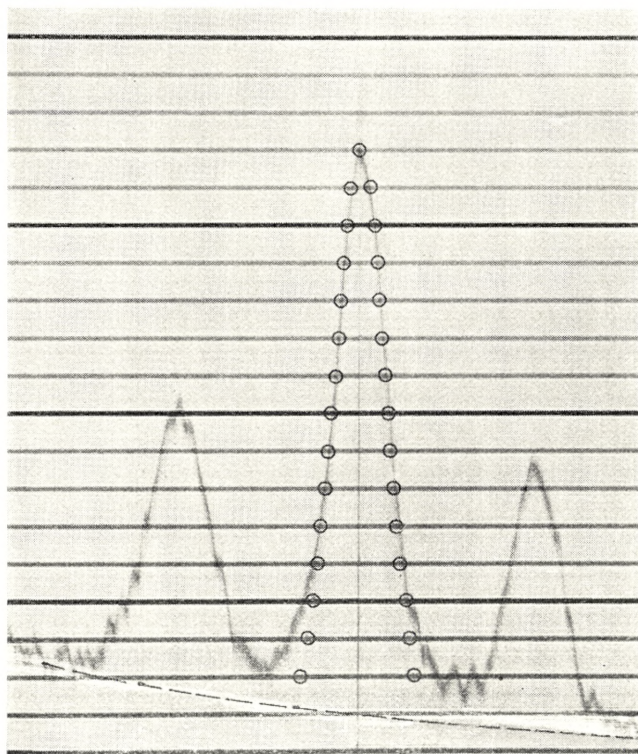


Fig. 3. Intensity distribution in powder diagram lines. The circles lie on a Gaussian error curve of the same height and breadth as the line.

Σ designates that the summation is to be extended over all particles which have the plane (hkl) in reflecting position, and the reflections of which fall within $\Delta\theta$ on the film.

If any particle is divided n times in every direction, the reflected energy becomes

$$A_2 = r' \Sigma f(\varphi, \omega, z).$$

The number of terms has been multiplied by n^3 , but the function f remains unchanged, since, apart from the intensity

distribution in the primary beam, it depends only on the size and the chemical composition of the specimen, and not on the particle size.

In other words, the same function should be summed up in n^3 as many points as previously. If $\tau = n^3 \tau'$, A_1 and A_2 are equal, as assumed above. Even if τ is different from $n^3 \tau'$, formula II $B^{1^2} = B^2 + b^2$ holds. In this case, $n^3 \tau' / \tau = \text{constant}$, hence $A_1 / A_2 = \text{constant}$. $I^1(\theta)$ will have another height than in the case $A_1 = A_2$, but it will show the same breadth.

Summary.

A method for X-ray determination of particle size is described. The breadth B^1 of a line hkl in a powder diagram of the substance whose particle size is to be determined, and the breadth B of the same line hkl in a "normal" diagram of the same chemical compound are measured. For the "normal" diagram, the substance is suitably large-grained, and both diagrams are taken with the same set of apparatus. $b = \sqrt{B^{1^2} - B^2}$ is the breadth with which the line hkl would appear in a diagram of the microcrystalline substance under ideal conditions (parallel primary beam, infinitely thin specimen, etc.). From b , the particle thickness t is computed by the formula

$$b = \frac{k \cdot \lambda}{t \cdot \cos \theta}.$$

The physical conditions for the validity of the formula $b = \sqrt{B^{1^2} - B^2}$ are discussed briefly.

I am indebted to the head of the laboratory, Professor NIELS BJERRUM, for the kind interest he has taken in this study, to TH. SIGURGEIRSSON, M. Sc. for discussions, and to "Teknisk Kemisk Fønd" for a grant.

*Chemical Laboratory,
Royal Veterinary and Agricultural College,
Copenhagen.*

References.

- 1) W. L. BRAGG, "The Crystalline State" I, London 1933, 187—191.
 - 2) P. SCHERRER, Göttinger Nachrichten 1918, 98.
 - 3) P. SCHERRER, Zsigmondy, Kolloidchemie, 3. Ed., Lpzg. 1920, 387.
 - 4) M. v. LAUE, Z. Krist. 64 (1926) 115.
 - 5) M. v. LAUE, Z. Krist. 82 (1932) 127.
 - 6) R. BRILL, Z. Krist. 68 (1928) 387.
 - 7) R. BRILL and H. PELZER, Z. Krist. 72 (1929) 398.
 - 8) R. BRILL and H. PELZER, Z. Krist. 74 (1930) 663.
 - 9) R. BRILL and H. PELZER, Z. Techn. Physik 10 (1929) 663.
 - 10) N. SELJAKOW, Zs. f. Physik 31 (1925) 439.
 - 11) C. C. MURDOCK, Phys. Rev. 35 (1930) 8.
-

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 9

EIN NEUES VERFAHREN
ZUR MESSUNG AKUSTISCHER
IMPEDANZEN

VON

V. THORSEN



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark
Bianco Lunos Bogtrykkeri A/S

EINLEITUNG

Die Messung der schallabsorbierenden Eigenschaften von Stoffen hat in den letzten Jahren auf Grund der grossen Bedeutung, die der Begriff Schalldämpfung erlangt hat, erneutes Interesse erfahren. Es sind verschiedene Verfahren zur Beschaffung der erforderlichen technischen Hilfsmittel vorgeschlagen und angewandt worden, um der Lösung der Probleme auf sicherer mathematischer Grundlage näher zu kommen. Die moderne Elektroakustik war hierfür von grosser Bedeutung, und die Zusammenarbeit mit der Elektrotechnik hat neue Gesichtspunkte erschlossen, die gute Ergebnisse gezeitigt haben.

Bei den älteren Messungen von Absorptionskoeffizienten war die Phasenänderung an der Oberfläche des Stoffes nicht berücksichtigt worden. Ein klarer Überblick wurde erst mit Einführung der akustischen Impedanz $Z = \frac{p}{v}$ erzielt, wobei p der Schalldruck und v die Partikelgeschwindigkeit (Schnelle) ist, die, in komplexer Form gegeben, Absorption und Phasenänderung des Stoffes eindeutig bestimmt. Eine Übersicht über Methoden zur Impedanzmessung findet sich im Handb. d. Exp. Phys. Bd. XVII, 2, S. 70. Eine weitere, neue Methode ist die Schustersche Brücke¹, eine akustische Wheatstonesche Brücke, die in recht einfacher Weise — leider aber nur innerhalb eines ziemlich engen Wellenlängenbereiches, nämlich etwa 400—1600 Hz. — gute Messungen auf rein akustischer Grundlage gestattet. Eine gewisse Erweiterung der Methode wurde von MENZEL² gegeben. Im folgenden sollen einige kritische Bemerkungen zur Schusterschen Brücke gemacht werden.

¹ K. SCHUSTER: Akust. Zs. 4, 253, 1939 (mit Literaturhinweisen auf frühere Arbeiten).

² W. MENZEL: Akust. Zs. 5, 257, 1940.

Das neue Verfahren.

Die Ausarbeitung des neuen Verfahrens fusst auf einem Vergleich von akustischen mit elektrischen (Wechselstrom-) Kreisen.

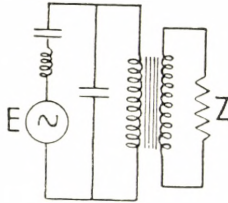
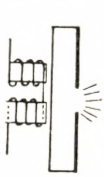


Fig. 1.

Ein gewöhnliches Telephone, das Schallenergie in den Raum strahlt, ist ein akustischer Transformator, dessen akustisches Bild und elektrisches Analogon in Fig. 1 wiedergegeben sind. Die schwingende Membran ist der Generator, der unter Berücksichtigung seiner Masse und Elastizität (Selbstinduktion

und Kapazität) die Luft durch den Vorraum (die primäre Wicklung) in Bewegung setzt (Stromstärke). In der Öffnung (der sekundären Wicklung) erlangen die Partikeln grössere Geschwindigkeit, die durch das Verhältnis von dem wirksamen Membranareal zu dem Areal der Öffnung (Übertragungsverhältnis) bestimmt ist. Die Öffnung strahlt Energie in den umgebenden Raum entsprechend ihrer akustischen Impedanz aus (der Sekundärstrom gibt Energie in einem äusseren Kreis ab). Der reelle Teil der Impedanz, der Strahlungswiderstand, bestimmt die ausgestrahlte Energiemenge, der imaginäre Teil bestimmt die Phase. Ein etwaiger Verlust im Vorraum, z. B. durch Reibung oder dgl., vermindert die Wirkung und ist durch einen über die primäre Wicklung des Transformators geschunteten Verlustkondensator symbolisiert.

Im folgenden wird das Telephone der Bequemlichkeit halber nicht als Transformator, sondern lediglich als akustischer

Generator betrachtet, wobei der Generator dann als der Luftpfropf angesehen werden kann, der wie ein Stempel in der Öffnung schwingt (Kolbenmembran).

Wird das Telephone vor der Öffnung mit einem Rohr versehen, dessen Durchmesser gleich dem Öffnungsdurchmesser ist, so ergeben sich die obenstehenden Bilder (Fig. 2).



Fig. 2.

Der akustische Generator arbeitet auf einer äusseren Impedanz, bestehend aus dem rein imaginären Teil des Rohres (Z_1) und dem reellen Teil R , dem Strahlungswiderstand. Danach ergibt sich

$$Z = R + Z_1 = R + iq_1$$

$$|Z|^2 = R^2 + q_1^2.$$

Der Effekt Eff wird

$$\text{Eff} = \frac{R}{R^2 + q_1^2} \cdot E^2,$$

wobei E die elektromotorische Kraft ist.

Der Effekt erreicht ein Maximum für $q_1 = 0$ und hat dann den Wert

$$\text{Eff}_{\max} = \frac{1}{R} \cdot E^2.$$

Die Grösse von R finden wir auf Grund folgender Betrachtungen: Der Zusammenhang zwischen Absorptionskoeffizient a , Phasenunterschied \mathcal{J} und akustischer Impedanz $Z = w + iq$ ist bekanntlich gegeben durch die Gleichungen

$$a = \frac{4w}{(w+1)^2 + q^2} \quad \text{und} \quad \text{tg } \mathcal{J} = \frac{2q}{w^2 + q^2 - 1},$$

wobei jedoch zu bemerken ist, dass w und q die relativen Widerstandswerte sind, d. h. die wirklichen Widerstandswerte dividiert durch den Widerstand der Luft ρc . Werden diese Gleichungen etwas umgeschrieben, so zeigt es sich, dass sie zwei Systeme von Kreisen mit dem Absorptionskoeffizienten bzw. der Phase als Parameter darstellen. Man erhält

$$\left. \begin{aligned} & \left(w - \frac{2-a}{a} \right)^2 + q^2 = \frac{4(1-a)}{a^2} \\ & \left(\text{Mittelpunkt: } \left(\frac{2-a}{a}, 0 \right) \text{ Radius: } \frac{2}{a} \sqrt{1-a} \right), \end{aligned} \right\} \quad (1)$$

$$\left. \begin{aligned} w^2 + (q - \cot \vartheta)^2 &= \frac{1}{\sin^2 \vartheta} \\ \left(\text{Mittelpunkt: } (0, \cot \vartheta), \text{ Radius: } \frac{1}{\sin \vartheta} \right). \end{aligned} \right\} \quad (2)$$

System (1) schneidet die w -Achse in zwei Punkten, w_1 und w_2 , so dass $w_1 \cdot w_2 = 1$. w_1 entspricht der Phase $\vartheta = 0$, w_2 der Phase $\vartheta = 180^\circ$. $a = 0\%$ gibt die q -Achse selbst. System (2) schneidet die q -Achse in zwei Punkten, $q_1 = \cot \frac{\vartheta}{2}$ und $q_2 = -\operatorname{tg} \frac{\vartheta}{2}$, entsprechend den Phasen 90° bzw. -90° (270°). $\vartheta = 0$ gibt die w -Achse vom Punkt $(1, 0) \rightarrow \infty$, $\vartheta = 180^\circ$ gibt die w -Achse von $(1, 0)$ bis $(0, 0)$. Alle Kreise im System (2) gehen durch $(1, 0)$, der ausserdem ein 100%iger Absorption entsprechender, singulärer Punkt im System (1) ist. Die Schnittpunkte zwischen (1) und (2) ergeben zwei Gruppen von Werten für w und q . w wird stets positiv; das Vorzeichen von q entscheidet, ob der Blindwiderstand induktiv (+) oder kapazitiv (−) ist.

$q_1 = 0$ bedeutet $\vartheta = 180^\circ$, woraus sich ergibt, dass

$$\operatorname{Eff}_{\max} = \frac{1}{w_2} \cdot E^2,$$

wo w_2 der kleinste Wert der Schnittpunkte des Absorptionskreises mit der w -Achse ist.

Es ist jedoch nicht von vornherein klar, ob der Blindwiderstand, wie im obigem Bilde, induktiv oder kapazitiv, oder eventuell ein Gemisch von beiden ist. Man kann sich deshalb ebenso gut ein Bild wie folgendes denken. Hierbei hat man

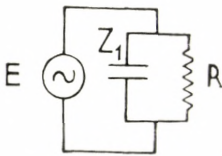


Fig. 3.

$$\begin{aligned} \frac{1}{Z} &= \frac{1}{R} + \frac{1}{Z_1} = \frac{1}{R} + \frac{1}{iq_1} \\ Z &= \frac{Rq_1^2}{R^2 + q_1^2} + i \cdot \frac{R^2q_1}{R^2 + q_1^2}. \end{aligned}$$

Ferner

$$|Z|^2 = \frac{R^2q_1^2}{R^2 + q_1^2} \quad \text{und} \quad \operatorname{Eff} = \frac{Rq_1^2}{R^2 + q_1^2} \cdot \frac{E^2}{|Z|^2} = \frac{1}{R} \cdot E^2;$$

dies ist der gleiche Ausdruck wie oben, R hat aber einen anderen Wert. Wenn $q_1 = \infty$, was $\vartheta = 0$ bedeutet, so erhält man

$$\text{Eff}_{\min} = \frac{1}{w_1} \cdot E^2,$$

wobei w_1 der grösste Wert der Schnittpunkte des Absorptionskreises mit der w -Achse ist.

Variiert man die Rohrlänge vor dem Telephon und misst man jedesmal die Ausstrahlung, so muss sich eine Kurve mit einem Maximum und einem Minimum ergeben, ähnlich der in Fig. 4 wiedergegebenen. Das Minimum liegt nahe bei der Rohrlänge $l = \frac{\lambda}{4}$, das Maximum nahe bei $l = \frac{\lambda}{2}$. Bildet man den Decibelunterschied zwischen der maximalen und der minimalen Ausstrahlung und rechnet in Energie (F) um, so erhält man

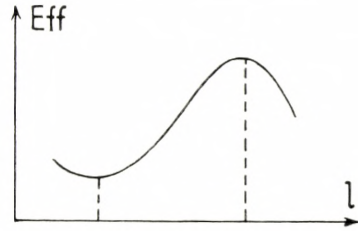


Fig. 4.

$$F = \frac{\text{Eff}_{\max}}{\text{Eff}_{\min}} = \frac{1}{w_2} \cdot E^2 : \frac{1}{w_1} \cdot E^2 = \frac{w_1}{w_2} = w_1^2.$$

Hiermit ist der Strahlungswiderstand w_1 ($= \sqrt{F}$) bestimmt, und der Absorptionskoeffizient ergibt sich aus der Gleichung

$$a = \frac{4 w_1}{(w_1 + 1)^2}.$$

Wenn a klein ist, kann man folgendermassen annähern: Die beiden Werte von w werden als Schnittpunkte des Kreises

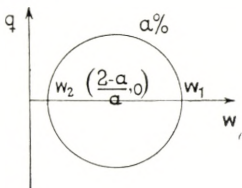


Fig. 5.

$$\left(x - \frac{2-a}{a}\right)^2 + y^2 = \frac{4(1-a)}{a^2}$$

mit der w -Achse abgelesen.

Man erhält

$$w_1 = \frac{2-a + 2\sqrt{1-a}}{a}$$

$$w_2 = \frac{2-a - 2\sqrt{1-a}}{a}.$$

Man sieht erstens, dass $w_1 \cdot w_2 = 1$. Ferner beträgt w_1 nahezu das Zweifache des Radius, also

$$w_1 = 2 \cdot \frac{\sqrt{1-a}}{a}.$$

Wird dies in $F = w_1^2$ eingesetzt, so erhält man

$$a = \frac{4}{F} (\sqrt{F} - 2).$$

Aus vorstehendem erhellt, dass es möglich ist, Strahlungswiderstände experimentell zu bestimmen. Wenn man dazu eine Schustersche Brücke benutzt, so ergeben sich rein prinzipielle, messtechnische Schwierigkeiten, so dass man früher eigentlich auf eine Berechnung allein angewiesen war. Im folgenden soll gezeigt werden, wie die experimentell bestimmten Werte mit den auf theoretischem Wege errechneten übereinstimmen; zuerst sei jedoch eine Erweiterung der Theorie durch strengere Forderungen eingehender beschrieben.

Erweiterung der Theorie.

Im vorstehenden ist eine andere Dämpfung als Strahlung, z. B. Reibung im Rohr, nicht berücksichtigt. Das ist indessen — jedenfalls bei höheren Frequenzen — notwendig, um so mehr, als man bei den Messungen variable Rohrlängen benutzt und infolgedessen mit verschiedener Rohrdämpfung zu rechnen hat. Es wird deshalb ein Verlustwiderstand in das elektrische Vergleichsbild eingeführt.

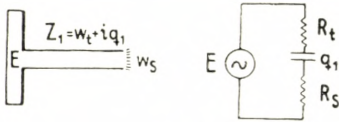


Fig. 6.

Man erhält

$$Z = R_t + R_s + i q_1; \quad |Z|^2 = (R_t + R_s)^2 + q_1^2,$$

$$\begin{aligned} \text{Eff}_{\max} &= \frac{R_t + R_s}{(R_t + R_s)^2 + q_1^2} \cdot E^2 \\ &= \frac{R_t}{(R_t + R_s)^2 + q_1^2} \cdot E^2 + \frac{R_s}{(R_t + R_s)^2 + q_1^2} \cdot E^2. \end{aligned}$$

Wir beschäftigen uns hier lediglich mit dem Ausstrahlungseffekt und sehen, dass dieser, wie oben, Maximum für $q_1 = 0$, $\vartheta_1 = 180^\circ$ wird.

$$\text{Eff}_{\max} = \frac{R_s}{(R_t + R_s)^2} \cdot E^2.$$

Analog erhält man für den Minimumwert der ausgestrahlten Energie

$$\text{Eff}_{\min} = \frac{R_s}{(R_t + R_s)^2} \cdot E^2,$$

oder, in die akustischen Bezeichnungen übertragen,

$$\text{Eff}_{\max} = \frac{w_s}{(w_t + w_s)^2} \cdot E^2 \quad \text{und} \quad \text{Eff}_{\min} = \frac{W_s}{(w_t + W_s)^2} \cdot E^2,$$

wo W die grossen Werte der Schnittpunkte des Absorptionskreises mit der w -Achse, und w die kleinen Werte derselben Grösse bezeichnet. Der Verlustwiderstand ist stets mit den übrigen Widerständen serriegeschaltet (die Verlustdämpfung lässt sich nie umgehen), so dass der gleiche Wert, abgesehen von dem Unterschied auf Grund der verschiedenen Rohrlängen, in beiden Ausdrücken enthalten ist. (Er wird mit einem kleinen Buchstaben bezeichnet, da er im allgemeinen klein ist.)

Man hat also

$$F = \frac{\text{Eff}_{\max}}{\text{Eff}_{\min}} = \frac{w_s (w_t + W_s)^2}{W_s (w_t + w_s)^2}.$$

Wenn $w_t \ll W_s$, was im allgemeinen in hohem Grad der Fall ist, und da $w_s \cdot W_s = 1$, kann dieser Ausdruck auf

$$F = \frac{1}{(w_t + w_s)^2}$$

reduziert werden, woraus sich ergibt, dass

$$w_t + w_s = \frac{1}{\sqrt{F}}.$$

Während man nach der einfachen Theorie

$$w_s = \frac{1}{\sqrt{F}}$$

fand, ist der wahre Strahlungswiderstand

$$w_s = \frac{1}{\sqrt{F}} - w_t.$$

Die Kenntnis von w_t beruht auf einer Anzahl in den letzten Jahren angestellter Untersuchungen¹, aus denen hervorgeht, dass der Absorptionskoeffizient für ein Rohr von der Länge l

$$a = 1 - e^{-2\beta l}.$$

An Hand sehr sorgfältiger Untersuchungen und in Übereinstimmung mit theoretischen Betrachtungen wurde ermittelt, dass

$$\beta = 3,3 \cdot 10^{-5} \frac{\sqrt{\nu}}{r} (1 + 0,0012 t),$$

wobei ν die Frequenz, r der Radius des Rohres und t die Temperatur ist. Hiernach kann man w_t für die beiden in Betracht kommenden Rohrlängen berechnen, nämlich

$$l = \frac{\lambda}{4} \quad \text{und} \quad l = \frac{\lambda}{2}.$$

In dem Ausdruck für F sind die beiden w_t in Zähler und Nenner etwas verschieden, denn w_t des Zählers bezieht sich auf $l = \frac{\lambda}{4}$, w_t des Nenners dagegen auf $l = \frac{\lambda}{2}$; w_t für $l = \frac{\lambda}{4}$ ist gewöhnlich mehrere hundertmal kleiner als W_s , so dass es, wie gesagt, im Zähler ausgelassen werden kann.

Das Reibungsproblem lässt sich indessen auch noch auf andere Weise lösen. Wenn der Reibungs-Absorptionskoeffizient für das Rohr

$$a = 1 - e^{-2\beta l},$$

¹ E. WAETZMANN & L. KEIBS: Ann. d. Phys. **22**, 247, 1935. H. OBERST: Akust. Zs. **2**, 76, 1937.

kann man in Wirklichkeit die w_s -Werte nicht als Schnittpunkte der Iso-Absorptionskreise mit der w -Achse ablesen, wohl aber auf der Kurvenschar mit der Gleichung

$$\left(x - \frac{2-a}{a}\right)^2 + y^2 = \frac{4(1-a)}{a^2},$$

die nicht etwa einer Kreisschar, sondern einer Spirale entspricht; wird darin $y = 0$ gesetzt und ist $a = 1 - e^{-2\beta l}$, so erhält man

$$\left(x - \frac{1 + e^{-2\beta l}}{1 - e^{-2\beta l}}\right)^2 = \frac{4e^{-2\beta l}}{(1 - e^{-2\beta l})^2},$$

woraus

$$x - \frac{1 + e^{-2\beta l}}{1 - e^{-2\beta l}} = \pm \frac{2e^{-\beta l}}{1 - e^{-2\beta l}}.$$

Hieraus findet man

$$x_1 = \frac{(1 + e^{-\beta l_1})^2}{1 - e^{-2\beta l_1}} \quad \text{und} \quad x_2 = \frac{(1 - e^{-\beta l_2})^2}{1 - e^{-2\beta l_2}},$$

wo l_1 und l_2 die beiden Rohrlängen sind, durch die die Impedanz des Rohres reell wird.

Da

$$e^{-\beta l} = 1 - \frac{\beta l}{1!} + \frac{(\beta l)^2}{2!} - \dots$$

kann man, wenn $\beta \cdot l$ klein ist, Glieder der 2ten und höherer Potenzen von $\beta \cdot l$ weglassen und erhält dann

$$x_1 = \frac{2}{\beta l_1} \quad \text{und} \quad x_2 = \frac{\beta l_2}{2}.$$

Fig. 7 zeigt die Impedanz des Rohres als Funktion seiner Länge. Für $l = \frac{\lambda}{4}$ erhält man $w_1 = x_1$, für $l = \frac{\lambda}{2}$ wird $w_2 = x_2$. Würde man einen Iso-Absorptionskreis zur Bestimmung des Strahlungswiderstandes verwenden, so wäre das ein Fehler, denn mit zunehmender Rohrlänge erhöhen sich die weniger als 1 betragenden w_s -Werte um x_2 und die grossen w_s -Werte um $\frac{1}{x_1}$ (nicht um x_1 , da w_s und w_t in Serie geschaltet sind und der grosse x_1 -Wert nur vorkommt, wenn w_t und der induktive

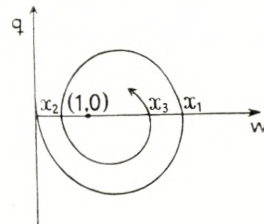


Fig. 7. Die Spirale beginnt in (0,0), geht zunächst nach negativen q -Werten (kapazitiver Widerstand, Elastizität) und dreht sich weiter, bis sie für ein unendlich langes Rohr in (1,0) endet.

Widerstand des Rohres als parallel verbunden betrachtet werden). Um w_s zu ermitteln, muss also von w_2 , das man aus $w_2 = \frac{1}{\sqrt{F}}$ bestimmt, x_2 abgezogen werden.

Demnach

$$w_2 = w_s + w_l = \frac{1}{\sqrt{F}}.$$

Messung einer beliebigen, unbekanntem Impedanz.

Es soll jetzt der Fall betrachtet werden, dass das Rohr mit einem Widerstand, und zwar z. B. durch eine mit einem Loch versehene Platte, abgeschlossen wird. Das akustische Bild und

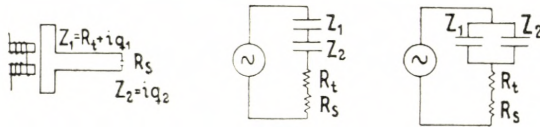


Fig. 8.

die beiden entsprechenden elektrischen Bilder sind in Fig. 8 wiedergegeben. R_t ist der Verlustwiderstand des Rohres, R_s sein Strahlungswiderstand.

Für Fall 1 (Reihenverbindung) hat man dann

$$Z = R_t + R_s + Z_1 + Z_2 = R_t + R_s + i(q_1 + q_2)$$

$$|Z|^2 = (R_t + R_s)^2 + (q_1 + q_2)^2$$

$$\text{Eff} = \frac{R_t + R_s}{(R_t + R_s)^2 + (q_1 + q_2)^2} \cdot E^2$$

$$= \frac{R_t}{(R_t + R_s)^2 + (q_1 + q_2)^2} \cdot E^2 + \frac{R_s}{(R_t + R_s)^2 + (q_1 + q_2)^2} \cdot E^2.$$

Die Ausstrahlung erreicht ein Maximum, wenn

$$q_1 + q_2 = 0$$

oder

$$\vartheta_1 + \vartheta_2 = (2p + 1)\pi$$

und wird dann

$$\text{Eff}_{\max} = \frac{w_s}{(w_s + w_l)^2} \cdot E^2.$$

Für Fall 2 (Parallelverbindung) erhält man auf die gleiche Weise

$$Z = R_t + R_s + \frac{Z_1 Z_2}{Z_1 + Z_2} = R_t + R_s + i \frac{q_1 q_2}{q_1 + q_2}$$

$$\text{Eff} = \frac{R_t + R_s}{(R_t + R_s)^2 + \frac{q_1^2 q_2^2}{(q_1 + q_2)^2}} \cdot E^2$$

$$= \frac{R_t}{(R_t + R_s)^2 + \frac{q_1^2 q_2^2}{(q_1 + q_2)^2}} \cdot E^2 + \frac{R_s}{(R_t + R_s)^2 + \frac{q_1^2 q_2^2}{(q_1 + q_2)^2}} \cdot E^2.$$

Die Ausstrahlung hat ein Minimum für

$$q_1 + q_2 = \infty$$

oder

$$\mathcal{J}_1 + \mathcal{J}_2 = 0$$

und der Minimumwert wird

$$\text{Eff}_{\min} = \frac{W_s}{(w_t + W_s)^2} \cdot E^2.$$

Man erhält also genau dieselbe Ausstrahlungskurve wie vorher, die Lage von Maximum und Minimum hat sich aber infolge des Vorhandenseins der festen Platte verschoben, und auch die Maximal- und Minimalwerte für die Ausstrahlung sind anders (kleiner) geworden, da sich der Strahlungswiderstand des Hauptrohres verändert hat. Betrachtet man, wie oben, das Verhältnis F von Maximum- zu Minimeffekt, so hat man

$$F = \frac{w_s}{W_s} \cdot \frac{(w_t + W_s)^2}{(w_t + w_s)^2},$$

was auch, wenn $w_t \ll W_s$ in den einfacheren Ausdruck

$$F = \frac{1}{(w_t + w_s)^2}$$

übergeht.

Man könnte sich schliesslich eine dritte Verbindungsart denken, wie sie in Fig. 8a wiedergegeben ist.

Hier erhält man

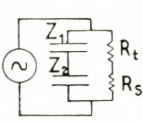


Fig. 8 a.

$$\frac{1}{Z} = \frac{1}{Z_1 + Z_2} + \frac{1}{R_t + R_s}; \quad Z = \frac{(R_t + R_s) \cdot (Z_1 + Z_2)}{R_t + R_s + Z_1 + Z_2}$$

$$Z = \frac{(R_t + R_s) \cdot (q_1 + q_2)^2}{(R_t + R_s)^2 + (q_1 + q_2)^2} + i \frac{(R_t + R_s)^2 \cdot (q_1 + q_2)}{(R_t + R_s)^2 + (q_1 + q_2)^2}$$

wo

$$|Z|^2 = \frac{(R_t + R_s)^2 \cdot (q_1 + q_2)^2}{(R_t + R_s)^2 + (q_1 + q_2)^2}$$

Der Ausstrahlungseffekt wird somit

$$\text{Eff} = \frac{R_s (q_1 + q_2)^2}{(R_t + R_s)^2 + (q_1 + q_2)^2} \cdot \frac{E^2}{|Z|^2} = \frac{R_s}{(R_t + R_s)^2} \cdot E^2,$$

d. h. Fall 1 und Fall 2 völlig entsprechend.

F muss so gross sein, dass w_s positiv wird, also

$$\frac{1}{\sqrt{F}} - w_t > 0.$$

Wenn kein w_t da wäre, könnte das Maximum theoretisch $= \infty$ und das Minimum $= 0$ sein, nämlich für $w_s = 0$, $W_s = \infty$. Sobald ein w_t da ist, muss das Maximum niedriger werden (und das Minimum höher), da aus dem Rohr nicht mehr als die vom Telephon stammende Energie kommen kann minus derjenigen, die im Rohr durch Reibung usw. (w_l) verloren gegangen ist.

Bei der Ableitung der vorstehenden Formeln für die ausgestrahlte Energie war es am praktischsten, Z_1 als aus Strahlungswiderstand, Verlustwiderstand und Blindwiderstand bestehend und Z_2 als reinen Blindwiderstand zu betrachten. In Wirklichkeit könnte man den Strahlungswiderstand natürlich ebenso gut mit Z_2 verknüpfen, dies um so mehr, als er als Loch in einer Platte zutage tritt. Bei der Messung der Impedanz einer solchen Platte mit einer Schusterschen Brücke würde man auf rein prinzipielle Schwierigkeiten stossen, die sich besonders stark geltend machen, wenn der Absorptionskoeffizient des Loches klein ist. Setzt man nämlich die Impedanz Z_2 vor das Ende der Brücke und schafft Gleichgewicht durch Verschiebung von Filz- und Stempelrohr, so weiss man in Wirklichkeit nur, dass die Impedanz in den beiden Symmetriequer-

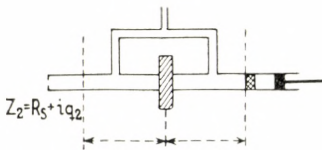


Fig. 9.

schnitten die gleiche ist; während man die Phase bei Z_2 durch einfache Rechnung ermitteln kann, wird bei der Absorptionsbestimmung die Absorption im Rohrstück von der linken punktierten Linie bis Z_2 mitgerechnet. Denkt man sich dieses Rohrstück verlängert oder verkürzt, so muss das Filzrohr um ein ebenso grosses (symmetrisches) Stück verschoben werden, so dass man stets die gleiche (und richtige) Phase und die gleiche (aber falsche) Absorption findet. Die wirkliche Absorption bei Z_2 ist daher erst nach Abzug der Verlustabsorption im Rohr zu bestimmen. Das bedeutet, dass die Verlustdämpfung in der vorstehend entwickelten Methode nicht berücksichtigt ist, was jedoch, wie die Messungen selbst ergeben haben, sehr grosse Fehler nach sich zieht (s. Näheres im folgenden). Hierzu kommt dann noch, dass man bei der von Schuster angegebenen, rein akustischen Eichungsmethode des variablen Widerstandes Absorptionswerte nur für Rohre von Längen, die Multipla von $\frac{\lambda}{2}$ sind, erlangen kann, d. h., dass man für Rohrlängen zwischen 0 und $\frac{\lambda}{2}$ gar keine Absorptionswerte bekommt; und gerade solche Werte sollten zur Korrektur in dem vorstehend besprochenen Falle benutzt werden. Diesen Nachteil vermeidet man bei elektro-akustischer Eichung des Widerstandes; die Korrektur ist aber in allen Fällen auszuführen.

Man kann die Eingangsimpedanz dagegen mit ziemlich guter Genauigkeit mit der Schusterschen Brücke messen, wenn sie vor das Telephon gesetzt wird, wie auf Fig. 10 zu sehen ist. Hierbei bestimmt man die Impedanz (Z) des Gesamtsystems, und die oft grossen Absorptionswerte, jedenfalls, wenn ϑ nahezu 180° ist, brauchen in dem Rohrstück der Schusterschen Brücke kaum für Verlustabsorption korrigiert zu werden.

Soll die Impedanz eines gegebenen Stoffes gemessen werden, so wird derselbe als Abschlussplatte vor die Mündung eines Rohres gebracht, das vor das Telephon gesetzt wird. Das Telephon als akustischer Generator arbeitet nun auf einer äusseren Impedanz, deren Strahlungswiderstand der reelle Teil der Impedanz des Stoffes und deren Verlustwiderstand die Dämpfung des Rohres ist, während der Blindwiderstand die Summe der Blindwiderstände des Rohres und des Stoffes ist. Aus der Verschiebung, die das Ausstrahlungsmaximum im Verhältnis zum Maximum für ein freistrahrendes Rohr erhält, wird der Phasensprung an der Oberfläche des Stoffes (der Blindwiderstand des Stoffes)

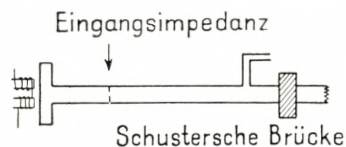


Fig. 10.

berechnet, da $\vartheta_1 + \vartheta_2 = (2p + 1)\pi$. Das Verhältnis von Maximum- zu Minimumwerten (F) ist

$$F = \frac{1}{(w_t + w_s)^2} \quad (\text{wenn } w_t \ll W_s),$$

sonst

$$F = \frac{w_s}{W_s} \frac{(w_t + W_s)^2}{(w_t + w_s)^2}.$$

Die unbekannte Impedanz wird alsdann

$$Z_2 = w_2 + iq_2,$$

wobei w_2 und q_2 die w - bzw. q -Werte der Schnittpunkte des Absorptionskreises (w_s) mit der der w -Achse parallelen Geraden sind, und zwar $q = \cot \frac{\vartheta_2}{2}$.

Anwendung auf verwandten Gebieten.

A. Die dämpfende Wirkung rauher Stoffe.

Im vorstehenden wurde die Grösse β für die Dämpfung von Schallwellen in einem Rohr nach der Kirchhoff-Helmholtzschen Formel

$$\beta = \frac{1}{c \cdot r} \sqrt{\frac{\mu}{\rho}} \pi \nu \left[1 + \sqrt{\frac{5}{2}} \left(\sqrt{z} - \frac{1}{\sqrt{z}} \right) \right]$$

benutzt, wo c die Schallgeschwindigkeit, r der Rohrradius, ν die Frequenz, ρ das spez. Gewicht, μ der Koeffizient für innere Reibung und $z = \frac{c_p}{c_v}$ ist. Die Gültigkeit dieser Formel ist in letzter

Zeit einer grossen Anzahl systematischer Untersuchungen unterzogen worden, wobei sich herausgestellt hat, dass man in Wirklichkeit mit 15 % höheren Werten rechnen muss als man direkt findet. Für ein glattes Messingrohr wird man β daher jetzt als sicher bekannt annehmen dürfen. In den vorstehend entwickelten Formeln sind die Verlustdämpfung und die Strahlungsdämpfung enthalten. Wenn man den Strahlungswiderstand kennt, kann man also aus dem Ausdruck $F = \frac{1}{(w_t + w_s)^2}$ die Verlustdämpfung ermitteln und sehen, wie diese vom Wandmaterial abhängt.

Man kann ein glattes Messingrohr geradezu als Normale benutzen, seine Ausstrahlung messen und daraus, da β bekannt ist, den Strahlungswiderstand ermitteln. Man macht danach ein Rohr aus dem Stoff, dessen dämpfende Wirkung untersucht werden soll, und der den gleichen Strahlungswiderstand hat (oder man bekleidet das Innere eines Messingrohres mit dem Stoff), und kann danach aus der Änderung der Ausstrahlung die vom Stoff verursachte Änderung der Wanddämpfung berechnen.

Damit haben wir ein Verfahren, um die reibungsdämpfende Wirkung eines Stoffes gegenüber Schall zu messen.

B. Das künstliche Ohr.

Mit der Entwicklung elektrischer Hörapparate hat auch die Herstellung eines künstlichen Ohres grosses Interesse erlangt. Der erste derartige Versuch, der jedoch zum Teil auf falschen Voraussetzungen beruht, wurde von GERLACH¹ angestellt. Als brauchbare Modelle werden heute die Konstruktionen von WEST² und von INGLIS, GRAY & JENKINS³ betrachtet. In beiden Modellen wird versucht, mit Hilfe eines Kondensatormikrophons den Schalldruck zu messen, der in einer kleinen Kammer vor einem akustischen Widerstande entsteht. Die kleine Kammer entspricht dem Raum zwischen einem Telephon und dem Ohr, der akustische Widerstand dem Ohr. Dieser Widerstand kann mit Hilfe elektro-akustischer Eichung derart gewählt werden, dass er dem mittleren Widerstandswert normaler Ohren entspricht.

Der Raum vor dem Modell hat ferner eine Impedanz, die der mittleren Impedanz der Räume vor normalen Ohren entspricht. Beide Modelle sind demnach »Mittelohren«; Änderungen, die bei schwerhörigen Ohren vorkommen können und die vielleicht gerade mit veränderten Impedanzeigenschaften des schwingenden Systems in Verbindung stehen, werden also nicht berücksichtigt. Die Eigenschaften normaler Ohren variieren aber auch beträchtlich. Die meisten Ohren haben negativen Blindwiderstand, d. h. das schwingende System ist elastisch (der Blindwiderstand kapazitiv), bisweilen wird er aber positiv ge-

¹ E. GERLACH: Wiss. Veröff. Siemens-Konz. **3**, 139, 1923/24.

² W. WEST: Post Office Electr. Eng. Journ. **22**, 260, 1929/30.

³ H. A. INGLIS, C. H. G. GRAY u. R. T. JENKINS: Bell Syst. Techn. Journ. **11**, 293, 1932.

funden. Es muss also eine Zunahme der Masse vorliegen (der Blindwiderstand induktiv). Auf Grund dieser Verschiedenheit, die von grossem Einfluss auf die Energieabsorption des Ohres ist, dürfte die Konstruktion eines »Mittelohres« von verhältnismässig geringer Bedeutung sein. Das Ohrmodell soll dazu dienen, den Energietransport durch das Trommelfell zum innern Ohr bei einem bestimmten Schalldruck vor dem Ohr zu berechnen.

Die scharfe Formulierung dieses Problems ist WAETZMANN¹ zu verdanken. Eine direkte Messung des Energietransportes ist vorderhand ausgeschlossen, da ein Eingriff in das Ohr zwecks Anbringung von Messapparaten nicht möglich ist. Dagegen kann man die von einem Telephon ausgehende, durch das Trommelfell zugeführte Energie messen, wenn man das Telephon als akustischen Generator mit einer äusseren Impedanz, die der des Ohres entspricht, arbeiten lässt, und die in der Impedanz in Gestalt von Strahlung abgesetzte Energiemenge misst. Dies ist nach den vorstehenden Ausführungen eben dadurch möglich, dass man als Strahlungswiderstände eine Anzahl Löcher in einer Platte und als Blindwiderstände eine variable Rohrlänge benutzt. Mit Hilfe einer so einfachen Einrichtung sind natürlich nur wenige diskrete Werte für die Absorptionskoeffizienten zu erlangen; es muss aber möglich sein, eine Art Irisblende herzustellen, die eine fortgesetzte Variation der Lochgrösse gestattet. Als kontinuierlich variabler Blindwiderstand könnte, ebenso wie bei der Schusterschen Brücke, ein Rohr, das in einem anderen verschiebbar ist, dienen.

Um zu messen, welche Energie ein elektrischer Hörapparat an das Ohr abgibt, könnte folgendes Verfahren angewandt werden:

1. Die freie Ausstrahlung des Hörapparates wird für alle gewünschten Frequenzen² gemessen.

2. Die Impedanz des Ohres wird für die gleichen Frequenzen mit einer Schusterschen Brücke (unter Berücksichtigung der erforderlichen Korrekturen) gemessen.

3. Die variable akustische Impedanz wird für jede Frequenz gleich der Impedanz des Ohres eingestellt und

4. vor das Telephon gebracht, wonach die Ausstrahlung gemessen wird.

¹ E. WAETZMANN & L. KEIBS: Akust. Zs. **1**, 3, 1936.

² V. THORSEN: Akust. Zs. **3**, 218, 1938.

Ist der Abstand zwischen Telephon und Messapparat (Kondensatormikrophon) in den Fällen 1 und 4 der gleiche, so erfahren wir, welcher Bruchteil der Energie, die das Telephon (frei) ausstrahlen kann, im Ohr wirklich absorbiert wird. Nimmt man an, dass die Schwellenwertkurve eines Ohres für gewisse Frequenzen — oder Frequenzgebiete — eine Anzahl Decibel höher ist, so muss es mittels einer Verstärkeranordnung möglich sein, so viel Energie von den verschiedenen Frequenzen zuzuführen, dass der Gehörverlust ersetzt wird. Möglicherweise wird es auch erforderlich, andere Frequenzen, die der Hörapparat besonders kräftig verstärkt (z. B. durch Resonanz), abzuschwächen; das ist ebenfalls wichtig, da sonst eine Überdeckungswirkung entstehen könnte. Anders ausgedrückt: es wird bezweckt, dem System Hörapparat plus Ohr eine normale Schwellenwertkurve zu geben. Dies ist zwar allgemein anerkannt, bisher aber noch nie verwirklicht worden; ein Verfahren nach den hier skizzierten Richtlinien wird natürlich sehr viel experimentelle Arbeit erfordern.

Experimentelle Ergebnisse.

Die Versuche wurden mit einem speziell hierfür gebauten Telephon ausgeführt. Es bestand aus einem dicken Messingzylinder mit runder Ausbohrung, worin das Magnetsystem eingeführt und festgespannt war. In angemessenem Abstand von den Magneten wurde die Membran mit Siegellack befestigt, so dass sie wie eine längs des Randes eingespannte (sonst aber spannungsfreie) Platte schwang. Das Ohrstück war durch eine 3 mm dicke Bleiplatte ersetzt, die zentrisch 6 mm gebohrt war. Die Vorderplatte des Telephons wurde mit einem 1,3 cm langen Messingrohr von 6 mm lichter Weite versehen, so dass die Gesamtlänge des vor dem Telephon befindlichen Rohres 1,7 cm beträgt. Mit einem Stückchen Gummischlauch wurden nun 6 mm Messingrohre verschiedener Länge mit dem kurzen Rohrstück verbunden, indem die Rohre im Gummischlauch ganz dicht zusammengefügt wurden; die Ausstrahlung wurde dann als Funktion der Gesamtrohrlänge gemessen. Die Erfahrung lehrt, dass diese Art des Zusammenfügens vortrefflich ist, denn dabei

erhält man eine ebene fortschreitende Welle im ganzen Rohr; es ist aber notwendig, dass die zusammengefügtten Rohre von genau demselben inneren Durchmesser sind. Derartige Stücke

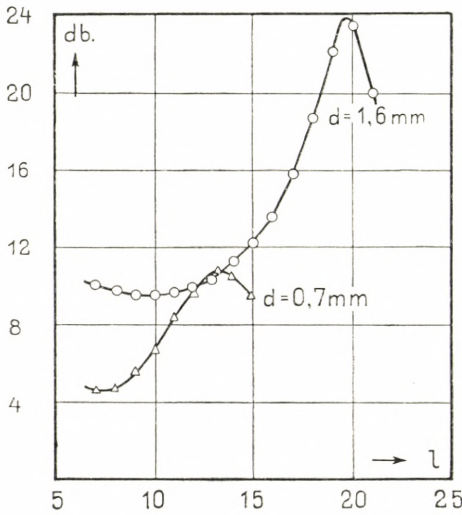


Fig. 11.

Es wurde die Ausstrahlung von der natürlichen Rohröffnung gemessen und ferner die Ausstrahlung, wenn die Öffnung mit Platten abgeschlossen war, worin sich runde Löcher von verschiedenem Durchmesser befanden. In Fig. 11 sind Messungen bei 2 Lochgrößen, nämlich 1,6 und 0,7 mm, graphisch darge-

müssen unter gewöhnlichen, gezogenen Messingrohren ausgesucht werden. Das Telephone wurde in dem schallisolierten und schallgedämpften Raum dem Kondensatormikrophon gegenüber so angebracht, dass die Mündung des Rohres stets 20,0 cm von dem Mikrophon entfernt war. Im übrigen wurden die Messungen in der in einer früheren Abhandlung¹ beschriebenen Weise ausgeführt. Die Versuche wurden für die Wellenlänge $\lambda_{20^\circ} = 44,3$ cm, $n = 768$ Hz durchgeführt.

Tabelle 1.

d	l	ϑ	ϑ_s
6,0	21,8	-185	-185
1,6	19,4	-226	-232
1,3	18,3	-242	-249
1,1	17,3	-259	-259
0,8	15,0	-296	-294
0,7	13,4	-322	-307
0,6	12,6	-335	-316
0,5	12,0	-345	-334

Tabelle 2.

d in mm	db	F	w_1	w_2	w_s	α
6,0	21,8	151,5	12,3	0,081	0,036	13,5
1,6	15,5	35,5	5,95	0,168	0,123	39
1,3	12,6	18,2	4,28	0,233	0,188	53,5
1,1	11,0	12,6	3,55	0,282	0,237	62
0,8	7,7	5,9	2,43	0,411	0,366	80
0,7	6,2	4,2	2,04	0,490	0,445	85
0,6	6,2	4,2	2,04	0,490	0,445	85
0,5	8,3	6,8	2,60	0,385	0,340	75

¹ V. THORSEN: loc. cit.

stellt. Sie sollen gleichzeitig zur Beurteilung der Genauigkeit dienen, die befriedigend ist. Die Messpunkte liegen gut auf glatten Kurven mit ausgeprägten Maxima und Minima. Für die Lage der Maxima wurden vorstehende Werte ermittelt (Tab. 1). In der 3. Kolonne sind die nach der

$$\text{Formel } \vartheta = -180^\circ - \frac{22,15 - l}{\frac{\lambda}{2}} \cdot 360^\circ$$

berechneten Phasenwinkel ϑ aufgeführt. Die letzte Kolonne zeigt die mit einer geeichten Schusterschen Brücke gemessenen Phasenwinkel (ϑ_s). Die Übereinstimmung ist ziemlich befriedigend.

Das Verhältnis F von Maximum- zu Minimumausstrahlung für die verschiedenen Löcher ist in Tab. 2 zusammengestellt. Diese Tabelle enthält ausserdem die w -Werte

($w_1 = \sqrt{F}$ $w_2 = \frac{1}{\sqrt{F}}$); die 3. Kolonne ist F , nämlich der Antilogarithmus zu $\frac{1}{10}$ mal dem Decibelunterschied zwischen maximaler und minimaler Ausstrahlung (2. Kolonne). In der letzten Kolonne schliesslich ist der Absorptionskoeffizient des reellen Strahlungswiderstandes (in %) verzeichnet. w_l kann aus $\frac{\beta l}{2}$ berechnet werden und würde Werte zwischen 0,035 und 0,050 ergeben; es wurde aber ein experimentell bestimmter Mittelwert, $w_l = 0,045$, benutzt. Danach erhält man w_s als $w_s = w_2 - w_l$.

Ferner wurde die nachstehende Messreihe zum Vergleich durchgeführt. Mit Hilfe der Schusterschen Brücke wurden die Absorptionskoeffizienten von einer Anzahl Löcher gemessen, wenn sie in der in Fig. 10 dargestellten Weise als Ausgangsimpedanzen in einer variablen Rohrlänge vor dem Telephon angebracht waren, also nach dem Telephon zu gesehen. Dadurch erhält man eine Anzahl schön verlaufender Kurven mit ausgeprägten Maxima und Minima. Zur Veranschaulichung des Kurvenverlaufes und der Genauigkeit sind die Ergebnisse der Messungen der Löcher mit $d = 1,1$ bzw. $0,7$ mm in Fig. 12 wiedergegeben. Man findet einen variablen Absorptionskoeffizienten,

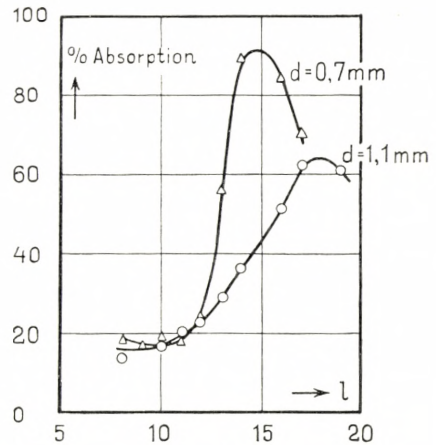


Fig. 12.

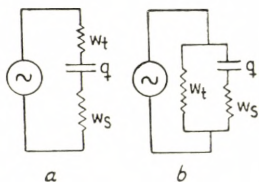


Fig. 13.

der für $\vartheta = 180^\circ$ ein Maximum und für $\vartheta = 0^\circ$ ein Minimum aufweist, und zwar für die gleichen ϑ -Werte, die eine Maximum- bzw. Minimausstrahlung ergeben. Das elektrische Vergleichsbild erhellt aus Fig. 13, *a* und *b*. Der erste Fall, $\vartheta = 180^\circ$, entspricht Fig. 13 *a*, wo $q = 0$, der zweite, $\vartheta = 0^\circ$, Fig. 13 *b*, wo $q = \infty$. Es gelingt also,

$$a_{\max} = w_t + w_s$$

und

$$a_{\min} = w_t$$

zu bestimmen. Dieser letztere Wert für w_t wurde in der vorstehenden Tab. 2 zur Bestimmung von w_s benutzt.

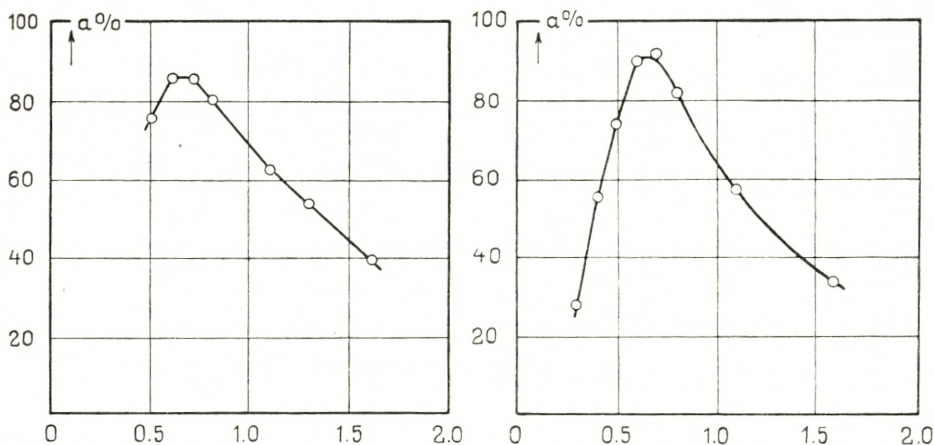


Fig. 14.

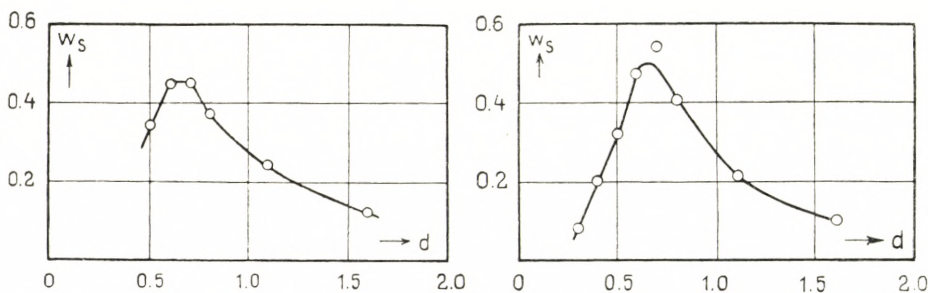


Fig. 15.

Auf Grundlage sämtlicher Messungen wurde Fig. 14 gezeichnet; diese Figur stellt den Absorptionskoeffizienten (angegeben in %) einer vor dem Telefon angebrachten Öffnung als Funktion des Durchmessers der Öffnung dar; die Werte wurden teils mit Hilfe der Ausstrahlungsmethode, teils mit Hilfe der Schusterschen Brücke gemessen, wobei im letzteren Falle von vorn nach dem Telefon zu gemessen wurde. Die Übereinstimmung ist gut.

Schliesslich zeigt Fig. 15 den »reinen« Strahlungswiderstand als Funktion des Durchmessers der Öffnung, a nach der Ausstrahlungsmethode, b mit der Schusterschen Brücke gemessen. Aus Fig. 14 und 15 ist ersichtlich, dass der Absorptionskoeffizient wie auch der Strahlungswiderstand ein Maximum für $d = \text{ca. } 0,7 \text{ mm}$ hat.

Das Vorhandensein eines solchen Maximums ist aber auch unschwer verständlich. Wenn das Loch genau der Öffnung des Rohres entspricht, ist sowohl w als auch q klein. Mit der allmählichen Abnahme des Lochdurchmessers nehmen w und q zu, w wächst anfänglich aber stärker als q , da letzteres gleich $\cot \frac{\vartheta}{2}$ ist und die Kotangente in der Nähe von 90° nur langsam variiert. Bei sehr kleinen Öffnungen wird q dagegen so gross, dass die Linie $q = \text{konst.}$ die den grossen a -Werten entsprechenden Absorptionskreise gar nicht schneidet. Bei einem gewissen mittleren Wert von q findet sich ein Absorptionskreis, den $q = \text{konst.}$ eben tangiert und der daher den grössten Wert ergibt, den w erlangen kann. Dies ist in Fig. 16 veranschaulicht.

Sucht man den Absorptionskoeffizienten der Endfläche, wenn diese nicht mit dem Rohr verbunden ist, so ergibt sich aus den vorstehenden Ausführungen, dass

$$a = \frac{4w}{(w+1)^2 + q^2}.$$

Diese Grösse durchläuft als Funktion des Lochdurchmessers eine Fig. 14 genau entsprechende Kurve. Wenn nämlich w und q

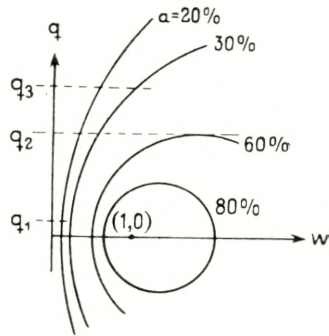


Fig. 16.

klein sind, ist a klein; wenn $d \rightarrow 0$, wird $q \rightarrow \infty$ und $w \rightarrow 0$, so dass a wieder gegen 0 geht. Irgendwo dazwischen hat a ein Maximum, und zwar für den w -Wert, für den der Einfluss von w auf q nicht mehr dominierend ist.

Aus den in den Tab. 1 und 2 angegebenen Werten lässt sich der wahre Absorptionskoeffizient der Öffnung ermitteln. So wird z. B. für $d = 0,8$ mm, wo $w = 0,37$ und $q = \cot\left(-\frac{296^\circ}{2}\right) = 1,60$,

$$a = \frac{4 \cdot 0,37}{1,37^2 + 1,60^2} = 0,34 \text{ oder } 34\%$$

ermittelt. Mit Hilfe der Schusterschen Brücke erhält man bei direkter Messung $a_{0,8} = 35\%$.



DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 10

ÜBER DIE ENERGIEAUSSTRAHLUNG
VON EINEM AKUSTISCHEN
WIDERSTAND

VON

V. THORSEN



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

EINLEITUNG

Eine frühere Abhandlung¹ enthielt die Beschreibung der Ausstrahlung von einem akustischen Widerstand bestehend aus einem Rohr von variabler Länge, das von einem Widerstand, z. B. einer Platte mit rundem Loch, abgeschlossen war. An Hand einer elektrischen Analogie wurde gezeigt, dass ein solches vor einem Telephon angebrachtes System, das als akustischer Transformator wirkt, die Berechnung der von der Rohrleitung ausgestrahlten Energie als Funktion der Impedanz des Rohres und der Endfläche ermöglicht, und dass umgekehrt die Messung der Energieausstrahlung eine Berechnung der Rohrimpedanz gestattet, und zwar für die Ausgangsimpedanz des Rohres wie auch für den Verlustwiderstand im Innern des Rohres. Dadurch erhielt man Methoden 1. zur Messung akustischer Impedanzen, 2. zur Messung der Dämpfung rauher Stoffe, 3. zur Konstruktion eines künstlichen Ohres.

Im folgenden soll die theoretische Grundlage der angewandten Methoden dargelegt werden, wobei eine mathematische Analyse der benutzten akustischen Rohrleitungen durchgeführt wird. Dadurch wird die Richtigkeit des benutzten Verfahrens in allen Einzelheiten bewiesen und die auf den reibungslosen Fall (mit Reflexions- und Phasenkreisen) gewöhnlich angewandte Betrachtungsweise dahin erweitert, dass sie auch Rohre mit Reibungsverlusten umfasst.

1. Die Telegraphengleichung und ihr akustisches Analogon.

Als Grundlage der theoretischen Behandlung wird die »Telegraphengleichung« der Elektrotechnik benutzt, deren Lösung folgende ist:

¹ V. THORSEN: D. Kgl. Danske Vidensk. Selskab, Mat.-fys. Medd. XX, 9.

$$W_a = \frac{U_a}{I_a} = \sqrt{\frac{R+i\omega L}{G+i\omega C}} \cdot \frac{\left(U_e \sqrt{\frac{G+i\omega C}{R+i\omega L}} + I_e \right) e^{l\gamma} + \left(U_e \sqrt{\frac{G+i\omega C}{R+i\omega L}} - I_e \right) e^{-l\gamma}}{\left(U_e \sqrt{\frac{G+i\omega C}{R+i\omega L}} + I_e \right) e^{l\gamma} - \left(U_e \sqrt{\frac{G+i\omega C}{R+i\omega L}} - I_e \right) e^{-l\gamma}}. \quad (1)$$

Hier ist W_a die Eingangsimpedanz, U_a , I_a bzw. U_e und I_e sind die Eingangs- und Ausgangsspannungen bzw. Stromstärken, R , G , L , C ohmscher Widerstand bzw. Ableitung, Selbstinduktion und Kapazität, alles pro Längeneinheit.

Werden die Abkürzungen

$$Z = \sqrt{\frac{R+i\omega L}{G+i\omega C}} \quad (\text{die charakteristische Impedanz}) \quad (2)$$

eingeführt und die Eingangs- und Ausgangsimpedanzen Z_a bzw. Z_e benannt, so kann man (1) auch schreiben:

$$Z_a = Z \frac{\frac{Z_e}{Z} (e^{l\gamma} + e^{-l\gamma}) + (e^{l\gamma} - e^{-l\gamma})}{(e^{l\gamma} + e^{-l\gamma}) + \frac{Z_e}{Z} (e^{l\gamma} - e^{-l\gamma})} \quad (3)$$

oder

$$Z_a = Z \frac{\frac{Z_e}{Z} \cosh l\gamma + \sinh l\gamma}{\cosh l\gamma + \frac{Z_e}{Z} \sinh l\gamma}, \quad (4)$$

wo ferner

$$\gamma = \sqrt{(R+i\omega L)(G+i\omega C)} = \beta + i\alpha. \quad (5)$$

β gibt an, wie sich die Amplitude mit der Länge (l) verändert, α , wie die Phase gleichzeitig variiert.

Für den Übergang von elektrischen zu akustischen Größen gilt folgendes Umsatzschema:

Elektrisch	Akustisch	Dimension
Spannung U Volt	Druck p	$ML^{-1}T^{-2}$
Stromstärke I Amp.	Volumengeschw. u	L^3T^{-1}
Widerstand pro cm $R \cdot$ Ohm	Reibung/cm ² . . . r/A	$ML^{-5}T^{-1}$

Elektrisch	Akustisch	Dimension
Ableitung G Siemens	1: Strömungs- geschw. $A \cdot s$	$M^{-1}L^3T$
Selbstinduktion . L Henry	akust. Masse . . . ρ/A	ML^{-5}
Kapazität C Farad	Nachgibigkeit . . A/pz	$M^{-1}L^3T^2$
Impedanz W Ohm	akust. Impedanz p/u	$ML^{-4}T^{-1}$

Mit Hilfe dieses Schemas kann die Telegraphengleichung dahin erweitert werden, dass sie für die akustische Rohrleitung gilt, und in gewissen Fällen wird es mittels einer mathematischen Behandlung möglich sein, (4) eine solche Form zu geben, dass man eine einfache physikalische Deutung der Ergebnisse erhält.

Der Ausdruck (2) für Z wie auch (5) für γ ist komplex und der Ausdruck (4) für Z_a daher sehr schwer zugänglich. Wenn R und G indessen kleine Grössen sind, kan man Z und γ mit Hilfe von Taylors Formel reihenentwickeln. Man hat

$$f(x+h, y+k) = f(x, y) + \frac{h}{1!} \frac{\partial f}{\partial x} + \frac{k}{1!} \frac{\partial f}{\partial y} + \frac{h^2}{2!} \frac{\partial^2 f}{\partial x^2} + 2 \frac{hk}{2!} \frac{\partial^2 f}{\partial x \partial y} + \frac{k^2}{2!} \frac{\partial^2 f}{\partial y^2} + \dots,$$

woraus

$$f(x, y) = f(0, 0) + \frac{x}{1!} \frac{\partial f}{\partial x} + \frac{y}{1!} \frac{\partial f}{\partial y} + \dots.$$

In diesem Falle ist

$$Z = f(R, G) = \sqrt{\frac{R + i\omega L}{G + i\omega C}},$$

also

$$Z = f(R, G) = f(0, 0) + R \left(\frac{\partial Z}{\partial R} \right)_{R=G=0} + G \left(\frac{\partial Z}{\partial G} \right)_{R=G=0} + \dots.$$

Nun ist

$$\frac{\partial Z}{\partial R} = \frac{1}{2} \sqrt{\frac{G + i\omega C}{R + i\omega L}} \cdot \frac{1}{G + i\omega C} \quad (\text{für } G = R = 0) \rightarrow \frac{1}{2 i\omega C} \sqrt{\frac{C}{L}}$$

und

$$\frac{\partial Z}{\partial G} = -\frac{1}{2} \sqrt{\left(\frac{G + i\omega C}{R + i\omega L} \right)^{-3}} \cdot \frac{1}{R + i\omega L} \quad (\text{für } G = R = 0) \rightarrow -\frac{1}{2 i\omega C} \sqrt{\frac{L}{C}}.$$

Eingesetzt ergibt dies

$$Z = \sqrt{\frac{L}{C}} + \frac{R}{2i\omega C} \sqrt{\frac{C}{L}} - \frac{G}{2i\omega C} \sqrt{\frac{L}{C}}. \quad (6)$$

Auf die gleiche Weise erhält man für $\gamma = \sqrt{(R+i\omega L)(G+i\omega C)}$, wenn R und G nahezu Null sind,

$$\gamma = i\omega\sqrt{LC} + \frac{R}{2} \sqrt{\frac{C}{L}} + \frac{G}{2} \sqrt{\frac{L}{C}}. \quad (7)$$

2. Die reibungsfreie akustische Rohrleitung.

Benutzt man als akustische Rohrleitung ein Messingrohr, so sind seine Wände so steif, dass man mit Sicherheit $G = 0$ hat. Ist das Rohr nicht zu eng, so dass die innere Reibung gleich Null zu betrachten ist, dann ist ausserdem $R = 0$.

Mit Hilfe von (6) und (7) und dem Umsatzenschema erhält man

$$Z = \sqrt{\frac{L}{C}} \quad \text{und} \quad \gamma = i\omega\sqrt{LC}$$

oder, in akustischer Form,

$$Z = \sqrt{\frac{\rho}{A} \cdot \frac{A}{\rho z}} = \frac{\rho c}{A} \quad \text{und} \quad \gamma = i\omega \sqrt{\frac{\rho}{A} \cdot \frac{A}{\rho z}} = i \frac{\omega}{c} = ik,$$

da $\omega = k \cdot c$ und $\rho c^2 = p z$.

Wird dies in (4) eingesetzt, so erhält man

$$Z_a = \rho c \frac{\frac{Z_e}{\rho c} \cosh ikl + \sinh ikl}{\cosh ikl + \frac{Z_e}{\rho c} \sinh ikl}.$$

Bedenkt man, dass

$$\cosh(ix) = \cos x \quad \text{und} \quad \sinh(ix) = i \cdot \sin x,$$

so geht die letzte Gleichung über in

$$Z_a = \rho c \frac{\frac{Z_e}{\rho c} \cos kl + i \sin kl}{\cos kl + i \frac{Z_e}{\rho c} \sin kl}. \quad (8)$$

Dies ist somit die Eingangsimpedanz für ein Rohr von der Länge l , das von der Ausgangsimpedanz Z_e abgeschlossen ist.

Die Gleichung (8) kann auf verschiedene Weise behandelt werden. Führt man den Amplituden-Reflexionskoeffizienten r für Z_e ein und die Phasenänderung bei demselben ϑ , so erhält man

$$r e^{i\vartheta} = \frac{Z_e - Z}{Z_e + Z} = \frac{\frac{Z_e}{\rho c} - 1}{\frac{Z_e}{\rho c} + 1}. \quad (9)$$

Hieraus ergibt sich

$$\frac{Z_e}{\rho c} = \frac{1 + r e^{i\vartheta}}{1 - r e^{i\vartheta}}.$$

Wird dies in (8) eingesetzt, so erhält man, wenn man $\cos kl$ und $\sin kl$ gleichzeitig durch $\frac{e^{ikl} + e^{-ikl}}{2}$ bzw. $\frac{e^{ikl} - e^{-ikl}}{2i}$ ersetzt,

$$\frac{Z_a}{\rho c} = \frac{(1 + r e^{i\vartheta})(e^{ikl} + e^{-ikl}) + (1 - r e^{i\vartheta})(e^{ikl} - e^{-ikl})}{(1 - r e^{i\vartheta})(e^{ikl} + e^{-ikl}) + (1 + r e^{i\vartheta})(e^{ikl} - e^{-ikl})},$$

woraus durch Reduktion

$$\frac{Z_a}{\rho c} = \frac{e^{i(2kl - \vartheta)} + r}{e^{i(2kl - \vartheta)} - r}.$$

Wird im Zähler und Nenner mit $(e^{-i(2kl - \vartheta)} - r)$ multipliziert, und werden \cos und \sin wieder eingeführt, so erhält man

$$Z_a = \rho c \cdot \frac{1 - r^2 - i 2 r \sin(2kl - \vartheta)}{1 + r^2 - 2 r \cos(2kl - \vartheta)}. \quad (10)$$

Diese Gleichung ist zur Untersuchung der Phase gut geeignet.

Man sieht, dass Z_a reell ist, wenn

$$\begin{aligned} \sin(2kl - \vartheta) &= 0, \\ 2kl - \vartheta &= 2p\pi \end{aligned}$$

also

$$l = \frac{\vartheta}{2\pi} \cdot \frac{\lambda}{2} + p \cdot \frac{\lambda}{2}. \quad (10a)$$

Die Rohrlänge l , die den ersten reellen Wert von Z_a ergibt, bestimmt also den Phasensprung bei Z_e .

Für $2kl - \vartheta = 0, 2\pi, 4\pi \dots$ erhält man $\cos(2kl - \vartheta) = +1$ und für $2kl - \vartheta = \pi, 3\pi, 5\pi \dots$ erhält man $\cos(2kl - \vartheta) = -1$, so dass Z_a , das also für diese Werte reell ist,

$$\frac{Z'_a}{qc} = \frac{1+r}{1-r} \quad \text{und} \quad \frac{Z''_a}{qc} = \frac{1-r}{1+r} \quad (10b)$$

wird.

Für $2kl - \vartheta = 0$ ist Z_a Maximum, für $2kl - \vartheta = \pi$ Minimum. Man sieht ferner, dass

$$\frac{Z'_a}{qc} \cdot \frac{Z''_a}{qc} = 1.$$

Die Gleichung gestattet indessen auch, r zu bestimmen. Aus (8) folgt, dass man für $l = 0$

$$Z_a = Z_e$$

hat.

Setzt man also $l = 0$ in (10) ein, so erhält man

$$Z_e = w + iq = qc \frac{1 - r^2 + i2r \sin \vartheta}{1 + r^2 - 2r \cos \vartheta}$$

oder, wenn man

$$\frac{w}{qc} + i \frac{q}{qc} = w_0 + iq_0$$

setzt,

$$w_0 = \frac{1 - r^2}{1 + r^2 - 2r \cos \vartheta} \quad \text{und} \quad q_0 = \frac{2r \sin \vartheta}{1 + r^2 - 2r \cos \vartheta},$$

woraus

$$\cos \vartheta = \frac{w_0(1 + r^2) - (1 - r^2)}{2rw_0} \quad \text{und} \quad \sin \vartheta = \frac{q_0(1 - r^2)}{2rw_0}. \quad (11)$$

Werden die beiden letzten Gleichungen quadriert und addiert, so erhält man

$$[w_0(1+r^2) - (1-r^2)]^2 + q_0^2(1-r^2)^2 = 4r^2w_0^2$$

$$r^4 - 2r^2 \frac{w_0^2 + q_0^2 + 1}{(w_0 + 1)^2 + q_0^2} + \frac{(w_0 - 1)^2 + q_0^2}{(w_0 + 1)^2 + q_0^2} = 0$$

$$r^2 = \frac{w_0^2 + q_0^2 + 1}{(w_0 + 1)^2 + q_0^2} \pm \sqrt{\frac{[w_0^2 + q_0^2 + 1]^2 - [(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}{[(w_0 + 1)^2 + q_0^2]^2}}$$

Der Zähler unter dem Wurzelzeichen wird $4w_0^2$, also

$$r^2 = \frac{(w_0 - 1)^2 + q_0^2}{(w_0 + 1)^2 + q_0^2} \tag{12}$$

Aus (11) erhält man demnach

$$\operatorname{tg} \mathcal{J} = \frac{2q_0}{w_0^2 + q_0^2 - 1} \tag{13}$$

r^2 ist der *Energie-Reflexionskoeffizient*. Wird er mit R bezeichnet, so hat man also

$$r^2 = R \quad \text{und} \quad a = 1 - r^2 = 1 - R,$$

wo a der *Energie-Absorptionskoeffizient* ist.

Die Gleichung (8) kann indessen auch auf andere Weise behandelt werden.

Setzt man

$$\frac{Z_e}{\rho c} = w_0 + iq_0,$$

so erhält man durch Einsetzen in (8)

$$\frac{Z_a}{\rho c} = \frac{(w_0 + iq_0) \cos kl + i \sin kl}{\cos kl + i(w_0 + iq_0) \sin kl} = \frac{w_0 + i \left(q_0 \cos 2kl - \frac{1}{2} [w_0^2 + q_0^2 - 1] \sin 2kl \right)}{(\cos kl - q_0 \sin kl)^2 + w_0^2 \sin^2 kl}$$

Z ist reell, wenn

$$q_0 \cos 2kl - \frac{1}{2} (w_0^2 + q_0^2 - 1) \sin 2kl = 0$$

oder

$$\operatorname{tg} 2kl = \frac{2q_0}{w_0^2 + q_0^2 - 1}$$

Der reelle Wert von $\frac{Z_a}{\rho c}$ ist also

$$\begin{aligned} \left(\frac{Z_a}{\rho c}\right)_{\text{reel}} &= \frac{w_0}{(\cos kl - q_0 \sin kl)^2} \\ &= \frac{w_0}{\frac{1}{2}(1 - \cos 2kl)(w_0^2 + q_0^2) - q_0 \sin 2kl + \frac{1}{2}(1 + \cos 2kl)}. \end{aligned}$$

Aber

$$\cos 2kl = \pm \frac{w_0^2 + q_0^2 - 1}{\sqrt{(w_0^2 + q_0^2 - 1)^2 + 4q_0^2}}$$

und

$$\sin 2kl = \pm \frac{2q_0}{\sqrt{(w_0^2 + q_0^2 - 1)^2 + 4q_0^2}},$$

die, eingesetzt, die reellen Werte $\frac{Z'_a}{\rho c}$ und $\frac{Z''_a}{\rho c}$ ergeben.

$$\begin{aligned} \frac{Z'_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 + \sqrt{(w_0^2 + q_0^2 - 1)^2 + 4q_0^2}}{2w_0} \\ \frac{Z''_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 - \sqrt{(w_0^2 + q_0^2 - 1)^2 + 4q_0^2}}{2w_0} \end{aligned} \quad \left. \vphantom{\frac{Z'_a}{\rho c}} \right\} (14)$$

oder

$$\begin{aligned} \frac{Z'_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 + \sqrt{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}}{2w_0} \\ \frac{Z''_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 - \sqrt{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}}{2w_0}. \end{aligned} \quad \left. \vphantom{\frac{Z'_a}{\rho c}} \right\} (15)$$

Diese beiden Werte sind Schnittpunkte eines Kreises

$$\text{Mittelpunkt: } \left(\frac{w_0^2 + q_0^2 + 1}{2w_0}, 0 \right)$$

$$\text{Radius: } \frac{\sqrt{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}}{2w_0}$$

mit der w -Achse, also mit der Gleichung

$$\left(x - \frac{w_0^2 + q_0^2 + 1}{2w_0}\right)^2 + y^2 = \frac{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}{4w_0^2}.$$

Man hat

$$r^2 = \frac{(w_0 - 1)^2 + q_0^2}{(w_0 + 1)^2 + q_0^2},$$

woraus

$$1 - r = 1 - \sqrt{\frac{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}{4w_0^2}}$$

und

$$1 + r = 1 + \sqrt{\frac{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}{4w_0^2}},$$

woraus wiederum

$$\begin{aligned} \frac{1+r}{1-r} &= \frac{\sqrt{(w_0 + 1)^2 + q_0^2} + \sqrt{(w_0 - 1)^2 + q_0^2}}{\sqrt{(w_0 + 1)^2 + q_0^2} - \sqrt{(w_0 - 1)^2 + q_0^2}} \\ &= \frac{w_0^2 + q_0^2 + 1 + \sqrt{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}}{2w_0}, \end{aligned}$$

und auf die gleiche Weise

$$\frac{1-r}{1+r} = \frac{w_0^2 + q_0^2 + 1 - \sqrt{[(w_0 + 1)^2 + q_0^2][(w_0 - 1)^2 + q_0^2]}}{2w_0}.$$

Die Schnittpunkte des Kreises mit der w -Achse sind also

$$\frac{1+r}{1-r} \quad \text{bzw.} \quad \frac{1-r}{1+r}.$$

Aus (12) und (13) erhellt, dass Iso-Reflexionskurven und Iso-Phasenkurven Kreise sind. Bei einfacher Berechnung erhält man aus (12)

$$w^2 + q^2 - 2w \frac{1+R}{1-R} + 1 = 0$$

oder

$$\left(w - \frac{1+R}{1-R}\right)^2 + q^2 = \frac{4R}{(1-R)^2}, \tag{16}$$

d. h. die Gleichung eines Kreises mit dem Mittelpunkt in $\left(\frac{1+R}{1-R}, 0\right)$ und dem Radius $= \frac{2\sqrt{R}}{1-R}$.

Aus (13) findet man auf dieselbe Weise

$$w^2 + (q - \cot \vartheta)^2 = \frac{1}{\sin^2 \vartheta}, \quad (17)$$

d. h. die Gleichung eines Kreises mit dem Mittelpunkt in $(0, \cot \vartheta)$ und dem Radius $= \frac{1}{\sin \vartheta}$.

Die Kreise (16) und (17) schneiden sich in zwei Punkten, deren Koordinaten w_0 und q_0 die Komponenten von Z_e (und somit von Z_a) sind. Sind w_0 und q_0 bekannt, so können r und ϑ bestimmt werden, und umgekehrt, wenn r und ϑ bekannt sind, können w_0 und q_0 eindeutig bestimmt werden. Z_a durchläuft den Reflexionskreis, wenn l variiert; seine Phase für jeden Wert von l ergibt sich aus (10a), seine Widerstands- und Blind-Komponente w_0 und q_0 aus dem Schnittpunkt des entsprechenden Phasenkreises mit der Geraden der Gleichung $q = \cot \frac{\vartheta}{2}$.

3. Berechnung des Absorptionskoeffizienten.

Wenn das Rohr mit der Eingangsimpedanz Z_a und der Ausgangsimpedanz Z_e vor einem Telephon angebracht wird, so wird der Absorptionskoeffizient von Z_e für die ausgestrahlte Energie bestimmend. Wenn l variiert, wird Z_a einen Iso-Reflexionskreis durchlaufen und für $2kl - \vartheta = 0$ Maximum und für $2kl - \vartheta = \pi$ Minimum sein. Dies ist unmittelbar aus (10) zu entnehmen. Maximum von Z_a besagt Minimum ausgestrahlter Energie, und Minimum von Z_a — Maximum der Ausstrahlung. Die Ausstrahlungskurve erhält somit die in Fig. 1 wiedergegebene Form.

Wird das Verhältnis zwischen den Maximum- und Minimumwerten der Ausstrahlungskurve F benannt, so hat man (10b) gemäss

$$F = \frac{\frac{Z'_a}{\varrho c}}{\frac{Z''_a}{\varrho c}} = \left(\frac{1 + \sqrt{R}}{1 - \sqrt{R}} \right)^2. \quad (18)$$

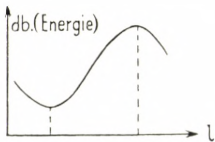


Fig. 1.

Hieraus ergibt sich

$$\sqrt{R} = \frac{\sqrt{F}-1}{\sqrt{F}+1}.$$

Aber $R = 1 - a$, also

$$a = 1 - R = \frac{4\sqrt{F}}{(\sqrt{F}+1)^2}. \quad (19)$$

Ein Vergleich von (18) mit (10 b) ergibt ferner unmittelbar, dass die reelle Komponente von Z_e (der Strahlungswiderstand) gleich \sqrt{F} ist.

Hiermit ist völlige Übereinstimmung mit der in der früheren Abhandlung¹ angegebenen Formel zur Berechnung der Ausstrahlung geschaffen, und da diese Berechnung auf einer elektrischen Analogie fusst, ist die Zulässigkeit der Anwendung dieser Analogie erwiesen.

4. Die akustische Rohrleitung mit Reibung.

Für eine Rohrleitung, die so eng und so lang ist, dass man von der Reibung nicht absehen kann, wird die Lösung von (4) erheblich schwieriger. Man hat jetzt nur $G = 0$ und findet aus (6)

$$Z = \sqrt{\frac{L}{C} + \frac{R}{2i\omega C}} \sqrt{\frac{C}{L}}.$$

Mit Hilfe des Umsatzschemas erhält man

$$Z = \frac{\varrho c}{A} + \frac{r}{A} \cdot \frac{p z}{2iA} \cdot \frac{A}{\varrho c} = \frac{\varrho c}{A} \left(1 - i \frac{r}{2\varrho\omega} \frac{p z}{\varrho c^2} \right) = \frac{\varrho c}{A} \left(1 - i \frac{r}{2\varrho\omega} \right). \quad (20)$$

Ferner aus (7)

$$\gamma = i\omega \sqrt{LC} + \frac{R}{2} \sqrt{\frac{C}{L}},$$

was an Hand des Umsatzschemas

$$\gamma = ik + \frac{r}{2A} \sqrt{\frac{A}{p z} \cdot \frac{A}{\varrho}} = ik + \frac{r}{2\varrho c} \quad (21)$$

ergibt.

¹ loc. cit. S. 3.

Bei direkter Rechnung erhält man

$$Z = \sqrt{\frac{R + i\omega L}{i\omega C}} = \sqrt{\frac{\frac{r}{A} + \frac{i\omega C}{A}}{i\omega \cdot \frac{A}{pz}}} = \frac{1}{A} \sqrt{\frac{(r + i\omega C) pz}{i\omega}},$$

was nach passender Umschreibung übergeht in

$$Z = \frac{qc}{A} \sqrt{1 - i \frac{r}{q\omega}} = \frac{qc}{A} \left(1 - i \frac{r}{2q\omega}\right),$$

wenn $\frac{r}{q\omega}$ klein ist.

Ferner

$$\gamma = \sqrt{(R + i\omega L) \cdot i\omega C} = \sqrt{\left(\frac{r}{A} + i\omega \frac{q}{A}\right) \cdot i\omega \cdot \frac{A}{pz}},$$

was umgeformt werden kann in

$$\gamma = ik \sqrt{1 - i \frac{r}{q\omega}} = ik \left(1 - i \frac{r}{2q\omega}\right) = ik + \frac{r}{2q\omega},$$

wenn $\frac{r}{q\omega}$ klein ist.

Werden die Werte für Z und γ in (4) eingesetzt, so ergibt sich

$$Z_\alpha = qc \cdot \frac{\frac{Z_e}{qc} \cosh \left[ikl \left(1 - i \frac{r}{2q\omega}\right) \right] + \left(1 - i \frac{r}{2q\omega}\right) \sinh \left[ikl \left(1 - i \frac{r}{2q\omega}\right) \right]}{\left(1 - i \frac{r}{2q\omega}\right) \cosh \left[ikl \left(1 - i \frac{r}{2q\omega}\right) \right] + i \frac{Z_e}{qc} \sinh \left[ikl \left(1 - i \frac{r}{2q\omega}\right) \right]}. \quad (22)$$

Dieser Ausdruck, der von ziemlich übersichtlicher Form ist, kann durch passende Behandlung umgewandelt werden in¹

¹ Für die Durchrechnung dieses Ausdrucks möchte ich Herrn stud. mag. B. BUCHMANN-OLSEN danken.

$$Z_a = \rho c \cdot \left. \begin{aligned} & \left[\frac{2 Z_e + r l}{2 \rho c} + \frac{2 \cdot \frac{r^2}{k} Z_e}{(2 \rho c)^3} \right] \cos kl + \frac{r}{2 \rho c} \sin kl + i \left[\left(1 + \frac{2 r l Z_e}{(2 \rho c)^2} \right. \right. \\ & \left. \left. + \frac{2 \cdot \frac{r^3}{k} l Z_e}{(2 \rho c)^4} \right) \sin kl - \frac{2 \cdot \frac{r^2}{k} l}{(2 \rho c)^2} \cos kl \right] \\ & \left(1 + \frac{2 r l Z_e}{(2 \rho c)^2} \right) \cos kl - \frac{2 \cdot \frac{r}{k} Z_e}{(2 \rho c)^2} \sin kl + i \left[\frac{2 Z_e + r l}{2 \rho c} \sin kl \right. \\ & \left. + \frac{2 \cdot \frac{r^2}{k} l Z_e}{(2 \rho c)^3} \cos kl \right] \end{aligned} \right\} (23)$$

worin nur gewöhnliche trigonometrische Funktionen auftreten. Dieser Ausdruck lässt sich noch vereinfachen; denn da $\frac{r}{\rho \omega}$ als eine kleine Grösse vorausgesetzt wurde, kann man von allen Gliedern absehen, die $\frac{r}{\rho \omega}$ in höheren Potenzen als 1 enthalten. Dadurch erhält man

$$Z_a = \rho c \cdot \frac{\left(\frac{Z_e}{\rho c} + \frac{r}{2 \rho c} \cdot l \right) \cos kl + \frac{r}{2 \rho \omega} \sin kl + i \sin kl}{\cos kl - \frac{Z_e}{\rho c} \cdot \frac{r}{2 \rho \omega} \cdot \sin kl + i \left(\frac{Z_e}{\rho c} + \frac{r}{2 \rho c} \cdot l \right) \sin kl}. \quad (24)$$

Dieser Ausdruck enthält einerseits ein »universelles« Reibungsglied $\frac{r}{2 \rho \omega}$, das von Z herrührt, andererseits ein Reibungsglied $\frac{r}{2 \rho c} \cdot l$, das mit der Länge der Rohrleitung wächst.

Bevor diese Gleichung (24) allgemein behandelt werden kann, wollen wir die (12) und (13) entsprechenden Ausdrücke für den Reflexionskoeffizienten $r^2 = R$ und den Phasensprung \mathcal{S} bei Z_e zu bilden versuchen.

Man hat

$$\sqrt{R} \cdot e^{i \mathcal{S}} = \frac{Z_e - Z}{Z_e + Z},$$

wo

$$Z = qc \cdot \left(1 - i \frac{r}{2qc}\right) = qc \cdot (1 - i\eta).$$

Also

$$\sqrt{R} e^{i\vartheta} = \frac{w_0 + iq_0 - (1 - i\eta)}{w_0 + iq_0 + (1 - i\eta)} = \frac{w_0^2 + q_0^2 - 1 + \eta^2}{(w_0 + 1)^2 + (q_0 - \eta)^2} + i \frac{2(q_0 + w_0\eta)}{(w_0 + 1)^2 + (q_0 - \eta)^2}$$

Hieraus erhält man zunächst

$$\operatorname{tg} \vartheta = \frac{2(q_0 + w_0\eta)}{w_0^2 + q_0^2 - 1 + \eta^2} \quad (25)$$

und ferner

$$R = \frac{[w_0^2 + q_0^2 - 1 + \eta^2]^2 + 4[q_0 + w_0\eta]^2}{[(w_0 + 1)^2 + (q_0 - \eta)^2]^2},$$

was nach einiger Umformung zu

$$r^2 = R = \frac{(w_0 - 1)^2 + (q_0 + \eta)^2}{(w_0 + 1)^2 + (q_0 - \eta)^2} \quad (26)$$

führt.

In (25) und auch in (26) ist die Ähnlichkeit mit (12) und (13) auffallend.

Es ist nicht tunlich, (24) in eine (10) entsprechende Form zu bringen. Der reelle Teil des Zählers allein wird so viele Glieder (mehr als 20) enthalten, dass man den Überblick gänzlich verliert und daher keine durchgreifende Reduktion vornehmen kann; dazu kommt dann noch der imaginäre Teil von Zähler und Nenner. Dagegen hat sich das auf S. 9 ff. benutzte Verfahren als brauchbar erwiesen. Der Einfachheit halber wird die Untersuchung jedoch in 2 Teile geteilt: zunächst wird (24) unter der Voraussetzung betrachtet, dass das Glied $\frac{r}{2qc} \cdot l$ (im folgenden mit $\xi \cdot l$ bezeichnet) nicht existiert. Sodann wird das Glied $\xi \cdot l$ wieder in das Endergebnis eingeführt, was sich als ziemlich leicht erweist, wenn die einfachere Form erst gelöst ist.

In (24), ohne das Glied $\xi \cdot l$, setzen wir also

$$\frac{Z_e}{qc} = w_0 + iq_0$$

und erhalten nun

$$\frac{Z_a}{\rho c} = \frac{(w_0 + iq_0) \cos kl + \eta \sin kl + i \sin kl}{\cos kl - \eta (w_0 + iq_0) \sin kl + i (w_0 + iq_0) \sin kl}.$$

Durch Multiplikation in Zähler und Nenner mit der konjugiert komplexen Grösse erhält man

$$\begin{aligned} \frac{Z_a}{\rho c} &= \frac{w_0 - \eta (w_0^2 - q_0^2 - 1) \cos kl \sin kl - w_0 \eta^2 \sin^2 kl}{(\cos kl - w_0 \eta \sin kl - q_0 \sin kl)^2 + (w_0 - q_0 \eta)^2 \sin^2 kl} \\ &+ i \frac{q_0 (\cos^2 kl - \sin^2 kl) - (w_0^2 + q_0^2 - 1) \cos kl \sin kl - 2 (w_0 \eta - q_0 \eta^2) \sin^2 kl}{(\cos kl - w_0 \eta \sin kl - q_0 \sin kl)^2 + (w_0 - q_0 \eta)^2 \sin^2 kl}. \end{aligned}$$

Die Bedingung dafür, dass $\frac{Z_a}{\rho c}$ reell ist, ist demnach

$$q_0 (\cos^2 kl - \sin^2 kl) - (w_0^2 + q_0^2 - 1) \cos kl \sin kl - 2 (w_0 \eta - q_0 \eta^2) \sin^2 kl = 0.$$

Aus dieser Gleichung findet man

$$\operatorname{tg} 2kl = \frac{2 (q_0 + w_0 \eta)}{w_0^2 + q_0^2 - 1}, \tag{27}$$

da die Annäherung gemacht wird, dass Glieder mit η^2 weglassen werden können, was als zulässig zu betrachten ist, da η eine kleine Grösse ist. Dadurch erhält man

$$\cos 2kl = \pm \frac{w_0^2 + q_0^2 - 1}{\sqrt{(w_0^2 + q_0^2 - 1)^2 + 4 (q_0 + w_0 \eta)^2}}$$

und

$$\sin 2kl = \pm \frac{2 (q_0 + w_0 \eta)}{\sqrt{(w_0^2 + q_0^2 - 1)^2 + 4 (q_0 + w_0 \eta)^2}}.$$

Mit Hilfe dieser Ausdrücke kann der reelle Teil des Zählers nach einigem Rechnen folgendermassen geschrieben werden:

$$w_0 - \frac{\eta}{2} (2 q_0 + w_0 \eta),$$

während der Nenner

$$\frac{1}{2} (w_0^2 + q_0^2 + 1) \pm \sqrt{(w_0^2 + q_0^2 - 1)^2 + 4 (q_0 + w_0 \eta)^2}$$

wird.

Macht man ferner den Nenner rational, so lassen sich die beiden reellen Werte der Impedanz $\frac{Z'_a}{\rho c}$ und $\frac{Z''_a}{\rho c}$ folgendermassen schreiben:

$$\left. \begin{aligned} \frac{Z'_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 + \sqrt{(w_0^2 + q_0^2 - 1)^2 + 4(q_0 + w_0\eta)^2}}{2(w_0 - q_0\eta)} \\ \text{und} \\ \frac{Z''_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 - \sqrt{(w_0^2 + q_0^2 - 1)^2 + 4(q_0 + w_0\eta)^2}}{2(w_0 - q_0\eta)}. \end{aligned} \right\} (28)$$

Für die Grösse unter dem Wurzelzeichen in (28) hat man nun

$$\begin{aligned} & (w_0^2 + q_0^2 - 1)^2 + 4(q_0 + w_0\eta)^2 \\ = & [(w_0 + 1)^2 + (q_0 - \eta)^2] [(w_0 - 1)^2 + (q_0 + \eta)^2] + 2\eta^2(w_0^2 + q_0^2 - 1) + \eta^4. \end{aligned}$$

Unter Auslassung höherer Potenzen von η kann (28) daher geschrieben werden:

$$\left. \begin{aligned} \frac{Z'_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 + \sqrt{[(w_0 + 1)^2 + (q_0 - \eta)^2] [(w_0 - 1)^2 + (q_0 + \eta)^2]}}{2(w_0 - q_0\eta)} \\ \text{und} \\ \frac{Z''_a}{\rho c} &= \frac{w_0^2 + q_0^2 + 1 - \sqrt{[(w_0 + 1)^2 + (q_0 - \eta)^2] [(w_0 - 1)^2 + (q_0 + \eta)^2]}}{2(w_0 - q_0\eta)} \end{aligned} \right\} (29)$$

Unter Benutzung des für den reibungsfreien Fall (S. 11) angewandten Verfahrens bildet man nun mit Hilfe von (26) die Grössen

$$\frac{1+r}{1-r} \quad \text{und} \quad \frac{1-r}{1+r},$$

worauf man nach einigem Rechnen (ebenso wie S. 11) zeigen kann, dass

$$\frac{Z'_a}{\rho c} = \frac{1+r}{1-r} \quad \text{und} \quad \frac{Z''_a}{\rho c} = \frac{1-r}{1+r}. \quad (29a)$$

Diese beiden Werte sind Schnittpunkte der Linie $q = -\eta \cdot w$ mit einem Kreise mit

$$\text{Mittelpunkt: } \left(\frac{w_0^2 + q_0^2 + 1}{2(w_0 - q_0\eta)}, -\eta \cdot \frac{w_0^2 + q_0^2 + 1}{2(w_0 - q_0\eta)} \right) \quad (30)$$

und Radius:

$$\sqrt{\frac{[(w_0 + 1)^2 + (q_0 - \eta)^2][(w_0 - 1)^2 + (q_0 + \eta)^2]}{2(w_0 - q_0 \eta)}} \cdot \sqrt{1 + \eta^2}.$$

Die Iso-Reflexionskurven werden aus (26) ermittelt. Man erhält

$$w^2 + q^2 - 2w \frac{1+R}{1-R} + 2q \frac{1+R}{1-R} \cdot \eta + \frac{1-R}{1+R} + \eta^2 = 0$$

oder

$$\left(w - \frac{1+R}{1-R}\right)^2 + \left(q + \eta \frac{1+R}{1-R}\right)^2 = \frac{4R}{(1-R)^2} (1 + \eta)^2. \quad (31)$$

Diese Kreisschar (nämlich, wenn R variiert) enthält natürlich auch den Kreis (30), den Kreis, dessen R w_0 und q_0 entspricht.

Die Iso-Phasenkurven sind auch Kreise. Man erhält nämlich aus (25)

$$w^2 + q^2 - 2w\eta \cot \vartheta - 2q \cot \vartheta - 1 - \eta^2 = 0$$

oder

$$(w - \eta \cot \vartheta)^2 + (q - \cot \vartheta)^2 = \frac{1 + \eta^2}{\sin^2 \vartheta}. \quad (32)$$

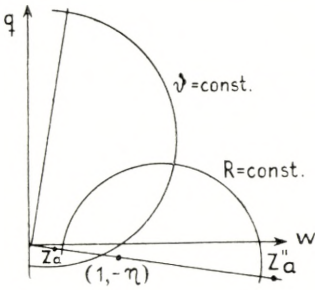


Fig. 2.

Reflexions- und Phasenkreise verlaufen somit wie Fig. 2 zeigt. Die Mittelpunkte liegen nicht, wie im reibungsfreien Falle, auf der w - bzw. q -Achse, sondern auf zwei winkelrecht zueinander befindlichen Linien, die mit den Achsen einen (kleinen) Winkel bilden, dessen Tangente $-\eta$ ist. Der Punkt $(1, 0)$ war in dem reibungsfreien Falle ein $R = 0$ entsprechender, singulärer Punkt, durch den alle

Phasenkreise gingen. Hier ist der entsprechende Punkt $(1, -\eta)$. Dies entspricht also $R = 0$ ($a = 1$) und alle Phasenkreise gehen auch hier durch. Diese letzteren schneiden ausserdem die Linie $q = +\frac{1}{\eta} \cdot w$ in den Punkten

$$\left(\eta \cot \frac{\vartheta}{2}, \cot \frac{\vartheta}{2} \right) \quad \text{und} \quad \left(-\eta \operatorname{tg} \frac{\vartheta}{2}, -\operatorname{tg} \frac{\vartheta}{2} \right).$$

Die Reflexionskreise schneiden die Linie $q = -\eta \cdot w$ an zwei Punkten, deren Koordinaten das Produkt 1 haben und welche Maximum- und Minimumwerte der Eingangsimpedanz Z_a des Rohres sind. Man ersieht hieraus, dass die Maximum- und Minimumwerte von Z_a nicht ganz reell sind, sondern eine kleine Blindkomponente der Grösse $\eta \cdot w$ haben; sie haben aber, wie in (29) und (29a) gezeigt wurde, die exakten Werte $\frac{1+r}{1-r}$ und $\frac{1-r}{1+r}$. Die Ausstrahlungskurve muss also genau den gleichen Verlauf nehmen wie im reibungsfreien Falle, der unter 3 in den Formeln (18) und (19) behandelt wurde. Dies ist indessen nur von sekundärem Interesse, da wir bei der Behandlung von (24) das Glied $\xi \cdot l$ ausgelassen haben. Dasselbe muss nun wieder eingefügt werden. Es zeigt sich indessen, dass $\xi \cdot l$ in der Weise in die Formel eingeht, dass man w unmittelbar durch $w + \xi \cdot l$ ersetzen kann. Aus (29) erhält man

$$\left. \begin{aligned} & \frac{Z'_a}{\varrho c} = \frac{(w_0 + \xi l)^2 + q_0^2 + 1}{2(w_0 + \xi l - q_0 \eta)} \\ & + \frac{\sqrt{[(w_0 + \xi l + 1)^2 + (q_0 - \eta)^2] [(w_0 + \xi l - 1)^2 + (q_0 - \eta)^2]}}{2(w_0 + \xi l - q_0 \eta)} \\ \text{und} \\ & \frac{Z''_a}{\varrho c} = \frac{(w_0 + \xi l)^2 + q_0^2 + 1}{2(w_0 + \xi l - q_0 \eta)} \\ & - \frac{\sqrt{[(w_0 + \xi l + 1)^2 + (q_0 - \eta)^2] [(w_0 + \xi l - 1)^2 + (q_0 + \eta)^2]}}{2(w_0 + \xi l - q_0 \eta)}. \end{aligned} \right\} (33)$$

Wird indessen in (33) $q_0 = -\eta w_0$ eingesetzt, so erhält man

$$\frac{Z'_a}{\varrho c} = w_0 + \xi \cdot l \quad \text{und} \quad \frac{Z''_a}{\varrho c} = \frac{1}{w_0 + \xi \cdot l}, \quad (34)$$

wobei von Gliedern mit höheren Potenzen von η als 1 abgesehen wird.

5. Ausstrahlung von einem akustischen Widerstand mit Reibung.

Fragen wir nach der ausgestrahlten akustischen Energie, so ist das Verfahren im wesentlichen dasselbe wie das unter 3 bei der Berechnung des Absorptionskoeffizienten benutzte. Das Maximum von $\frac{Z_a}{\rho c}$ ist das Minimum ausgestrahlter Energie und umgekehrt. Wird das Verhältnis von Maximum zu Minimum der Ausstrahlungskurve in Energiemassen für F ausgedrückt, so erhält man mit Hilfe von (34)

$$F = \frac{\frac{Z'_a}{\rho c}}{\frac{Z''_a}{\rho c}} = \frac{1}{(w_0 + \xi \cdot l)^2},$$

woraus

$$w_0 + \xi \cdot l = \frac{1}{\sqrt{F}}, \quad (35)$$

und diese Formel steht mit der in der früheren Abhandlung¹ ermittelten in Einklang.

Eine etwas bessere Annäherung erlangt man an Hand der folgenden Betrachtungen. Der Effekt, der sich im akustischen Widerstand absetzt, besteht aus zwei Komponenten, nämlich der sich im Rohr absetzenden (in $\xi \cdot l$) und der von der Öffnung ausgestrahlten (sich in w_0 absetzenden) Komponente. Bei der letzteren hat man für den Maximumwert

$$\text{Eff}'_{\max} = \frac{w'_0}{\left(\frac{Z'_a}{\rho c}\right)^2} \cdot E^2 = \frac{w'_0}{(w'_0 + \xi l_1)^2} \cdot E^2$$

und für den Minimumwert

$$\text{Eff}''_{\min} = \frac{w''_0}{\left(\frac{Z''_a}{\rho c}\right)^2} \cdot E^2 = \frac{w''_0}{(w''_0 + \xi l_2)^2} \cdot E^2,$$

¹ loc. cit. S. 3

wo w'_0 und w''_0 den »grossen« bzw. den »kleinen« Wert von $\frac{Z_a}{\rho c}$ darstellen. Dieselben werden aus (29) gefunden, indem man $q = -\eta \cdot w$ einsetzt. Danach erhält man

$$F = \frac{\text{Eff}_{\max}}{\text{Eff}_{\min}} = \frac{w'_0}{w''_0} \cdot \frac{(w''_0 + \xi l_2)^2}{(w'_0 + \xi l_1)^2}. \quad (36)$$

Hier sind l_1 und l_2 die beiden Rohrlängen, die dem Maximum bzw. Minimum entsprechen. Man sieht, dass, wenn w''_0 gegenüber ξl_2 gross ist, (36) in (35) übergeht, da man ja ausserdem $w'_0 \cdot w''_0 = 1$ hat.

Diese Formel (36) stimmt auch mit der in der früheren Abhandlung gefundenen genau überein, und damit ist die Richtigkeit sämtlicher Formeln für den Zusammenhang zwischen den akustischen Komponenten der verschiedenen Teile und der Ausstrahlung erwiesen. Zugleich hat man an Hand der in Fig. 2 ausgedrückten Betrachtungen ein bequemes Verfahren gewonnen, Probleme betreffend Absorption und Phasensprünge auch in denjenigen Fällen zu behandeln, wo Reibung vorhanden ist. Dieses Verfahren hat bisher nur dann benutzt werden können, wenn kein Verlustwiderstand vorhanden war.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 11

THE DECAY CONSTANT OF RaC'

BY

J. C. JACOBSEN AND TH. SIGURGEIRSSON



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

The determination of the decay constant of RaC' has been the subject of several investigations ^{1, 2, 3, 4, 5}. A considerable accuracy was obtained in the experiments of DUNWORTH⁴ and of ROTBLAT⁵ in which the ROSSI method of detecting coincidences was used. By this method, small time differences are determined by measuring the change occurring in the number of coincidences between two counters, when the resolving power of the circuit is changed.

The applicability of this method is restricted to time intervals from about 10^{-2} to 10^{-6} sec. In ROTBLAT's experiments, the resolving time was changed by changing the grid leak of one of the valves in the ROSSI circuit and, as a lower limit for the resolving time, $6 \cdot 10^{-5}$ sec was found. A lower limit of the order of magnitude 10^{-6} sec is apparently inherent in the use of counters and may be ascribed to a variable time difference between the arrival of the primary particle and the start of the discharge. It is not quite clear how the time difference can attain a value as high as 10^{-6} sec. If it is assumed that the discharge starts when a secondary electron arrives at the central wire, a variable time lag will result if the primary particle passes through the counter at different distances from the central wire. Under ordinary working conditions, the passage of an electron from the wall of the counter will, however, last about 10^{-8} sec, only, so that some other mechanism in the building up of the discharge must be responsible for the limitations of the resolving time actually found.

¹ JACOBSEN, Phil. Mag. XLVII, 23, 1924.

² BARTON, Phil. Mag. VII, 2, 1926.

³ JACOBSEN, Nature **133**, 565, 1934.

⁴ DUNWORTH, Nature **144**, 153, 1939.

⁵ ROTBLAT, Proc. Roy. Soc. **177**, 260, 1940.

⁶ ROTBLAT's paper was not available to us before the present work was completed in the summer of 1942.

The time interval between the emission of a β -particle from RaC and the accompanying α -particle from RaC' is of the order 10^{-3} to 10^{-4} sec. In the present experiments, the Rossi coincidence circuit is used with a constant resolving time, and the impulses from the β - and the α -particles are brought to coincide by delaying the impulse from the β -counter relative to that from the α -counter. The limitation of the method as regards the length of time which can be measured is the same as in the experiments of DUNWORTH and of ROTBLAT; an advantage, although not a fundamental one, is that the decay is observed directly and not by a differentiation as in the experiments mentioned previously.

The experiments.

The experimental arrangement is shown in Fig. 1. To the left of the figure, two counters are shown which are placed close together with their mica windows facing one another. Between the counters, the source is placed, a speck of active deposit from Radium emanation supported on an aluminium foil; the thickness of the foil was 6 mg per cm^2 , that of the windows was 3.5 mg per cm^2 . With this arrangement, one of the counters will count the β -particles from RaB and RaC, but no α -particles; the other one responds to both α - and β -particles. The counters are described in a paper by Miss H. LEVI.¹

With the thickness of the aluminium foil chosen, the β -counter will respond to most of the β -particles from both RaB and RaC; in agreement with this, it was found that the α -counter counted about 50 per cent more than the β -counter.

To the right in Fig. 1 is shown the Rossi stage which needs no further explanation. The resolving time was $7.5 \cdot 10^{-5}$ sec and was kept constant throughout the experiments. The resolving time was determined by counting the number of chance coincidences; the exact knowledge of the resolving time is without influence on the results, if only it is kept constant as will appear later.

The delay of the impulse from the β -counter, which is necessary to obtain coincidences, is effected by the circuits in the middle of Fig. 1; actually, both the β - and the α -kicks are de-

¹ Acta Physiol. Scand. **2**, 311, 1941.

layed by a certain amount, and it is the difference which enters into the measurements. The delay circuits D_1 and D_2 are shown in detail in Fig. 2. This circuit is known as a multivibrator; the mode of operation is well-known, and only for the sake of completeness it may be mentioned here. The circuit contains two valves which are coupled tightly together in such a way that, when the circuit is undisturbed, one valve is always "closed"

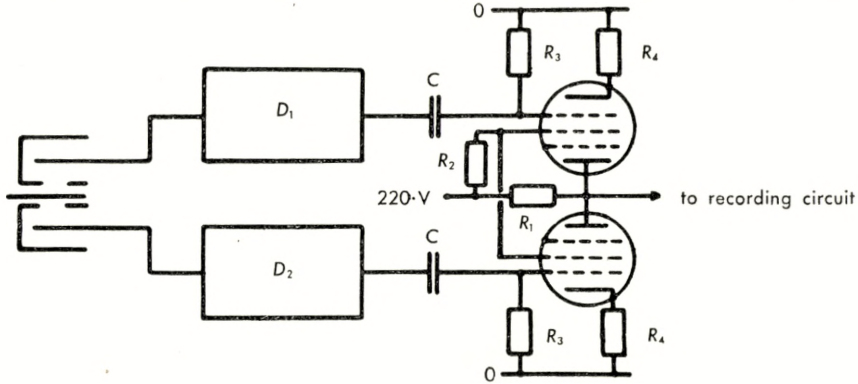


Fig. 1

D_1 D_2 delay circuits
(see fig. 2)

$R_1 = 1.0 M\Omega$
 $R_2 = 0.5 \text{ } \gg$
 $R_3 = 0.1 \text{ } \gg$
 $R_4 = 200 \Omega$
 $C = 25 \text{ cm}$

Valves:
Philips AF 7

and the other valve is "open". A small negative impulse on the grid of the open valve will cause an increase in the anode potential of this valve. This increase is coupled to the grid of the second valve through the condenser C_2 with the result that the anode potential of the second valve decreases. This decrease in turn is transferred to the grid of the first valve, etc. The process continues until the first valve is closed and the second is open. When a stage is reached, where the anode and grid potentials have attained nearly constant values, the circuit quickly returns to its normal state by a process which is the reverse of the one just described. The time interval between the arrival of the primary impulse and the return of the system to its normal state is determined by the condenser C_2 , the anode resistance of the first valve and the grid leak of the second valve and may be varied within wide limits by varying C_2 .

The negative impulses delivered by the delay stages are transferred to the Rossi stage. By introducing a difference in the delay of the impulses from the α - and the β -counter, the Rossi stage records a coincidence when a β - and an α -particle are emitted with a time difference which is determined by the delay circuits. The impulses delivered by the α - and the β -counters were counted separately by scale-of-sixteen circuits; these are not shown in the figure.

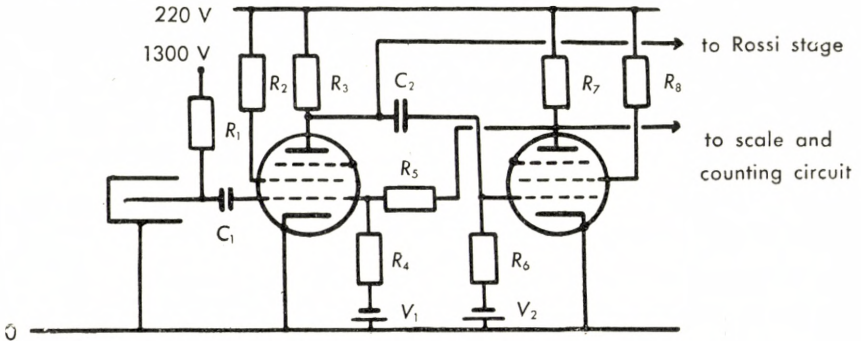


Fig. 2

$R_1 = 1.0 M\Omega$	$R_5 = 1.0 \Omega$	$V_1 = -75 V$
$R_2 = 0.2 \text{ »}$	$R_6 = 0.5 \text{ »}$	$V_2 = -20 V$
$R_3 = 0.2 \text{ »}$	$R_7 = 0.5 \text{ »}$	$C_1 = 50 \text{ cm}$
$R_4 = 0.5 \text{ »}$	$R_8 = 0.2 \text{ »}$	$C_2 = \begin{cases} 300 \text{ cm in the } \alpha\text{-delayer} \\ \text{Variable } > 300 \text{ cm in the } \beta\text{-delayer} \end{cases}$

Both valves: Philips AF 7

The measurements.

The delay circuit connected to the α -counter was kept unchanged during the measurements. By varying the condenser C_2 of the second delay circuit, different time intervals between the emission of the α - and the β -particles, which caused a coincidence, could be introduced. It was found that the delay varied very nearly linearly with the condenser C_2 . For each value of the condenser C_2 , the delay imparted to the β -impulse relative to the α -impulse was measured by connecting the output from the delay circuits through condensers to one pair of plates of a cathode ray oscillograph, the other pair of plates being connected to the usual sweep arrangement. A large number of genuine coincidences was produced by placing the counters close together

with their windows facing one another and irradiating with γ -rays. When a coincidence occurred, the deflections from the counters showed up separately as two peaks on the screen of the oscillograph. By suitably adjusting the sweep frequency, the two peaks could be brought to coincide. When this was the case, the time difference between the two deflections was equal to a whole number of periods of the sweep frequency; the sweep frequency itself was then determined by comparison with a known frequency; actually, a multiple of the frequency of the city mains was used. The only adjustment necessary for the measurement being to bring two peaks on the screen of the oscillograph to coincide, the determination of the delay could easily be made with an accuracy better than 1 per cent.

Among the factors determining the number of delayed coincidences actually counted, the solid angle of the beam of particles entering the counters should be chosen as high as possible. Actually, the counters were placed directly on one another with the source foil clamped between them.

The number of chance coincidences increases with the square of the source strength, while the number of genuine coincidences increases with the first power of the source strength, so that an upper limit is given for the strength of the source which can be used with advantage. In the experiments, the number of particles counted by the α - and the β -counters usually was about 2400 and 1600 per minute at the beginning of the experiment.

If the resolving time of the Rossi stage is τ , coincidences are counted over an interval of time equal to 2τ . The number of coincidences is thus proportional to τ , so that τ should be chosen as high as possible. If, on the other hand, τ is large relative to the half period, details of the decay curve will be lost. In the experiments, τ was chosen equal to $7.5 \cdot 10^{-5}$ sec, except for the measurement of the first point of the decay curve for which τ was $1.2 \cdot 10^{-5}$ sec.

With a given value of τ , the delay of the β -impulse relative to that of the α -impulse must not be made smaller than τ , otherwise the time interval over which the coincidences are counted will be reduced, because the α -particle is always emitted after the β -particle. A further reason to exclude very small time intervals from measurements was that a rather large number

of genuine coincidences was always observed; these no doubt were due to β -particles or secondaries from γ -rays which were scattered from one counter into the other.

For the measurements of the decay constant, the exact value of τ is without influence as long as τ is kept constant.

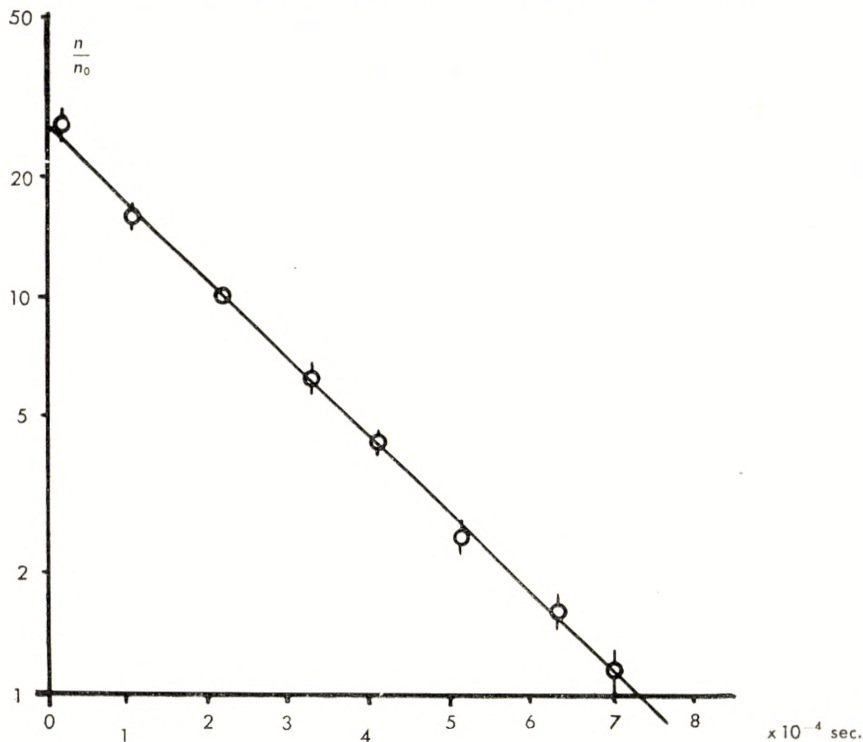


Fig. 3

In Fig. 3 is shown the number of coincidences as a function of the delay of the β -counter relative to that of the α -counter; the vertical lines indicate the statistical errors. For small values of the delay, the accuracy is limited by the necessity of using a small value of the resolving power of the Rossi stage; for large values of the delay, the number of coincidences becomes small. To be independent of the decay of the source, the number of coincidences was counted intermittently with the delay equal to $t_0 = 2.20 \cdot 10^{-4}$ sec and to a variable time interval t . The corresponding number of coincidences being n_0 and n , the ratio $n:n_0$ has been plotted as ordinate in Fig. 3. From the

exponential decay, the half period is found equal to $(1.55 \pm 0.05) \cdot 10^{-4}$ sec, in satisfactory agreement with the value obtained by ROTBLAT.

We now proceed to a comparison between the method used in ROTBLAT's experiments (variation of resolving time) and the method used in the present experiments (delayed coincidences) as regards the statistical error which results from the counting of a given number of particles.

When the resolving time is varied, the number N_1 of particles counted per second with the resolving time T_1 and the number N_2 counted per second with the resolving time T_2 are determined; with delayed coincidences, the number $N = N_2 - N_1$ is determined directly. To obtain definite and simple conditions, it is now assumed that only two points of the decay curve are determined by counting over the time intervals $T_2 - T_1$ and $T'_2 - T'_1$, that these points are separated by some four half periods, that the time intervals $T_2 - T_1$ and $T'_2 - T'_1$ are equal to the half period \mathcal{D} and, finally, that the time interval from $t = 0$ to $t = T_1$ is equal to $1/2 \mathcal{D}$. Other conditions might, of course, be chosen with a corresponding change in the numerical result of the comparison; the conditions just mentioned probably give a fair representation of the actual experiments.

With delayed coincidences, let the number of particles counted per second in the time intervals $T_2 - T_1$ and $T'_2 - T'_1$ be N and $N:16$, respectively; to obtain a relative statistical error ϵ , counting must be performed during a time interval $t_1 = \frac{1}{\epsilon^2 N}$ and $t'_1 = \frac{16}{\epsilon^2 N}$ for the two intervals, respectively, so that nearly the whole time spent in counting is used for the determination of the last point.

When the resolving time is varied, we need similarly only consider the determination of the last point of the decay curve for which the number of particles per second $\frac{N}{16}$ is determined as the difference between two countings, or $\frac{N}{16} = N_1 - N_2$. If N_1 is determined by counting during t_1 seconds, the statistical error is $\sqrt{\frac{N_1}{t_1}}$. A simple consideration shows that N_1 and N_2 are

nearly equal to $3N$; to obtain the same statistical error $\varepsilon \cdot \frac{N}{16}$ as above, t_1 is determined from

$$\varepsilon \cdot \frac{N}{16} = \sqrt{\frac{3N}{t_1}}, \quad t_1 = \frac{3 \cdot 16^2}{\varepsilon^2 N} \quad \text{or} \quad \frac{t_1}{t} = 48.$$

Since the same consideration applies to the determination of N_2 , the total time spent in counting is roughly 100 times longer than with delayed coincidences.

The variation of the resolving time further involves the possibility of a systematic error, as will be shown now. Although the error is probably too small as to be of any influence on the experimental results, an estimate of its magnitude may be of some interest.

When coincidences are counted, the length of the resolving time is not sharply defined. Two impulses may always be recorded as a coincidence if the time interval is smaller than $t - \varepsilon t$, where ε is small, and may never be recorded if the time interval is greater than $t + \varepsilon t$, whereas in the interval $t - \varepsilon t$ to $t + \varepsilon t$, the probability for recording a coincidence may vary between 0 and 1 (Fig. 4). The value of the resolving time, which

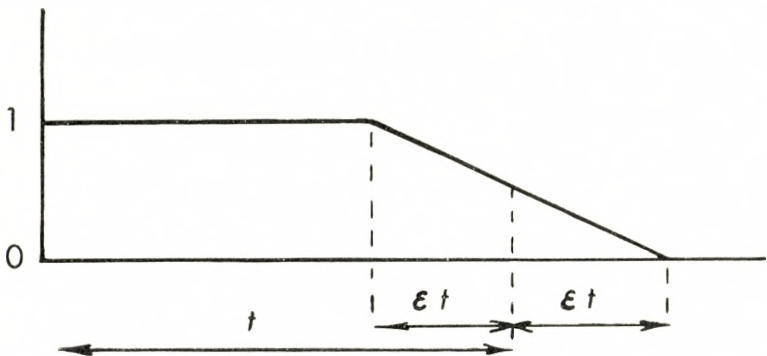


Fig. 4

is determined by counting chance coincidences, corresponds to the area below the curve. The error in question arises from the circumstance that the shape of the "cut off" probably varies with the length of the resolving time. In the measurements with

delayed coincidences, this source of error is not present, because the coincidence stage is left unchanged during the measurements.

The exact shape of the curve in Fig. 4 is of course unknown; a simple assumption is that the slope is rectilinear so that the probability for detecting a coincidence is determined by

$$\begin{aligned} S &= 1 && \text{for } \tau < t - \varepsilon t \\ S &= \frac{t + \varepsilon t - \tau}{2 \varepsilon t} && \text{for } t - \varepsilon t < \tau < t + \varepsilon t \\ S &= 0 && \text{for } \tau > t + \varepsilon t \end{aligned}$$

where τ is the time interval between two impulses.

As a somewhat idealized case, we consider now the value λ^* of the decay constant, which is determined by changing the resolving time t by an amount which is small compared with t itself. If $N(t)$ is the number of coincidences counted with the resolving time t , then

$$\lambda^* = -\frac{d^2 N}{dt^2} : \frac{dN}{dt}.$$

With the resolving time determined as above, we obtain

$$\begin{aligned} N(t) &= \lambda \cdot N(\infty) \int_0^{\infty} e^{-\lambda \tau} S(t, \tau) d\tau = \\ &= \lambda \cdot N(\infty) \left[\int_0^{(1-\varepsilon)t} e^{-\lambda \tau} d\tau + \int_{(1-\varepsilon)t}^{(1+\varepsilon)t} e^{-\lambda \tau} \frac{(1+\varepsilon)t - \tau}{2 \varepsilon t} d\tau \right]; \\ N(t) &= N(\infty) \left[1 - \frac{e^{-x}}{2 \varepsilon x} (e^{\varepsilon x} - e^{-\varepsilon x}) \right], \end{aligned}$$

where $x = \lambda t$.

If εx is small, then

$$\begin{aligned} \frac{dN}{dt} &= \lambda \cdot \frac{dN}{dx} = \lambda \cdot N(\infty) e^{-x} \left(1 - \frac{\varepsilon^2 x}{3} + \frac{\varepsilon^2 x^2}{6} \right) \\ \frac{d^2 N}{dt^2} &= \lambda^2 \frac{d^2 N}{dx^2} = -\lambda^2 \cdot N(\infty) e^{-x} \left(1 + \frac{\varepsilon^2}{3} - \frac{2\varepsilon^2 x}{3} + \frac{\varepsilon^2 x^2}{6} \right) \\ \frac{\lambda^*}{\lambda} &\approx 1 + (1-x) \frac{\varepsilon^2}{3} = 1 + (1-\lambda t) \frac{\varepsilon^2}{3}. \end{aligned}$$

This result shows that, for $\lambda t > 1$, the value of λ^* is smaller than λ , for $\lambda t < 1$, the reverse is the case. The difference, however, is rather insignificant; if, to obtain a quantitative estimate, ε is assumed equal to 0.1, the error in the determination of λ is less than 1 per cent, since λt never differs much from unity with the values of t used in the experiments.

Our thanks are due Professor NIELS BOHR for his continuous interest in our work.

Summary.

The decay constant of RaC' has been determined by a coincidence method. Active deposit from Radium emanation is placed between two counters which are screened in such a way that one counter counts β -particles, only, the other counts both α - and β -particles. By means of a multivibrator, the impulses from the β -counter are delayed relative to the impulses from the α -counter by an amount which is of the same order of magnitude as the mean lifetime of RaC'. Under these circumstances the emission of a β -particle from RaC and the subsequent emission of an α -particle from RaC' are observed as a coincidence. The number of coincidences as a function of the delay of the impulses from the β -counter shows an exponential decay over a time interval of about 4 half periods; for the half period is found $(1.55 \pm 0.05) \cdot 10^{-4}$ sec, in satisfactory agreement with the value obtained previously by ROTBLAT.

The method used in the present experiments is further discussed in relation to the method used by ROTBLAT, partly as regards the statistical errors arising from the counting, and partly as regards a systematic error which is likely to be present in ROTBLAT'S experiments. The result of the discussion is that this systematic error, although undoubtedly present in ROTBLAT'S experiments, is so small that it has hardly been of any influence on the results.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 12

ELECTROMAGNETIC
PROPERTIES OF NUCLEAR SYSTEMS
ON MESON THEORY

BY

C. MØLLER AND L. ROSENFELD



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

CONTENTS

	Page
Introduction	3
Part I: Electromagnetic properties of arbitrary fields	7
1. Gauge-invariance of Lagrangian	7
2. Transition to Hamiltonian and separation of Coulomb field	11
3. Hermitian variables	15
4. System in slowly varying external field	18
Part II: Electromagnetic properties of a system of nucleons and mesons	23
1. Hamiltonian of the total system	23
2. Expression of the charge density in terms of the transformed variables	28
3. Electric dipole and quadrupole moments	33
4. The current density	39
5. Magnetic dipole moment	41
6. Electromagnetic properties of the deuteron	49
Appendix	52
Mean values for the zero-point meson field	52
The exchange and proper current density and related quantities ..	56
Expression of the proper magnetic moments of nucleons for special source distributions	62
References	65

INTRODUCTION

The present paper is intended as a sequel to the authors' previous work "On the field theory of nuclear forces" [1]. In the latter, the question was discussed, whether it would be possible in a well-defined way to delimit a domain in which the meson field theory, in spite of the fundamental difficulties of its quantization, would still yield consistent results. For this purpose, a point of view analogous to the "correspondence" treatment of quantum electrodynamics was adopted and a general prescription could be formulated according to which a consistent interpretation of the formalism may be obtained. For the application of this prescription to the problem of nuclear forces, a first requirement is to separate the "static" part of these forces, for the same reason as the Coulomb forces in electrodynamics must be separated from the radiation field before a correspondence treatment of the latter can be arrived at. It was then shown that such a separation of the static forces could simply be effected by a canonical transformation, and the application of the general criterium mentioned above to the expression obtained in this way for the static forces led to the adoption of the so-called "mixed" meson theory, *i. e.* a mixture of vector and pseudoscalar meson fields with properly chosen intensities of the corresponding nuclear sources [2]. While the consequences of the point of view just sketched for the theory of β -disintegration and meson life-time have been fully discussed [3], its bearing on the electromagnetic properties of nuclear systems remained to be examined: this is the program of the present paper.

The completion of this work, however, has been much delayed, partly due to fortuitous circumstances, but partly also on account of the development of the subject itself. In fact, it was

pointed out in the meantime by one of us [4] that the mixed theory appeared as an elementary (*i. e.* not mixed) form of meson theory in a five-dimensional representation. A subsequent survey of the formal possibilities for five-dimensional meson field theories even showed that the mixed theory was the only combination of four-dimensional types of meson fields possessing the property just mentioned of coinciding with an irreducible five-dimensional type of field [5]. Two possible interpretations of such a five-dimensional representation were discussed; of these, it appeared that the projective interpretation was especially well-suited for the incorporation of the interaction with the electromagnetic field [6]. In this connexion, the main formulae concerning this interaction have already been published [6] and applied to the problem of the photo-effect of the deuteron [7]. On the other hand, this and other problems involving electromagnetic interactions have also been treated by several authors on the basis of different assumptions about meson fields and nuclear forces. It is, therefore, not likely that the present paper will contain any concrete result not already known to the physicists acquainted with the problems concerned. The interest it may nevertheless offer would rather lie in the more methodical aspects of the question, as viewed according to the general lines recalled above. By reason of this circumstance, we have decided also to include, when necessary, already well-known developments in an endeavour to present a rounded-off account of the questions treated.

Special care has been devoted to the more formal side of the theory, such as the systematic use and extension of the "symbolical space" algorithm already introduced in our first paper [1]. Since this system of representation might at first sight appear rather cumbersome when applied to charged meson fields, —in contrast to the more common representation of these fields by complex operators,—a few words on this subject will perhaps not be out of place here. An examination of the calculations developed in the Appendix with the help of this formalism will indeed show that they are not any lengthier or more intricate than those in which complex field variables are used. In most cases, we have even applied, for the sake of a more direct expression of the physical meaning of the quantities

considered, a three-dimensional vector notation in preference to a four- or a five-dimensional tensor notation. This requires, it is true, some practice of vector analysis which otherwise would not be needed and is, of course, not so appropriate to questions involving space and time variables in a symmetrical way (an example of such a case is afforded by the calculation of the magnetic moment of free mesons; cf. p. 42); still, the symbolic space representation of the charged fields proves very convenient even then.

In the first part of the paper, we recall the general method, applicable to any field theory, of taking into account the interaction with the electromagnetic field; on this occasion, we adapt the general formulae to the symbolical space representation, using Hermitian field variables. In the last section of this part, the interaction with a slowly varying external field is more especially considered and the multipole moments suited to the treatment of such a case—in fact, only the first few ones: electric dipole, and quadrupole, and magnetic dipole moments—are introduced. The second part of the paper is chiefly concerned with the application of the canonical transformation separating the static meson fields to the interaction of a system of nucleons and mesons with the electromagnetic field. This amounts to deriving the expression of the charge and current density operators in terms of the transformed variables; from these, the expressions of the multipole moments just mentioned are then easily obtained. The resulting formulae for all such electromagnetic quantities contain, besides the usual terms corresponding to free nucleons and free mesons, and others depending on both nucleon and meson variables, also terms of the so-called “exchange” type, *i. e.* terms which depend only on the nucleon variables in the form of a sum of expressions involving exchange of proton and neutron state between pairs of nucleons; the occurrence of such exchange terms is a typical feature of field theory. There appear also, of course, further contributions of a singular character, which have to be rejected according to the above-mentioned general criterium; it is noteworthy, however, that a singular contribution corresponding to the anomalous magnetic moment of the nucleon may still be retained as being of a different origin, *viz.* due to the fact that a material point

model is adopted for the nucleons*. It is clear that the method of canonical transformation here followed is superior to the usual perturbation method, above all in that it just permits easily to trace, so to say, the origin of the different terms occurring in any formula and so to facilitate their physical interpretation, especially in connexion with our general criterium. Furthermore, it also offers the practical advantage of requiring calculations much simpler and easier to survey than the perturbation theory.

As regards the application of the general scheme so obtained to special problems, we confine ourselves to some brief indications, since a more detailed treatment will be found in other publications. In this respect, we have already mentioned the investigation of the photo-effect of the deuteron by A. PAIS [7]; we wish also to call attention to a forthcoming paper by J. SERPE which, among others, will contain a detailed comparison of the canonical transformation and perturbation methods and a discussion of two singular effects (spreading of electric charge around a nucleon and anomalous magnetic moment of nucleons) providing appropriate examples of the application of our criterium. We should like here to mention that results of calculations by these authors, kindly put by them at our disposal, have been very valuable for the final redaction of this paper.

Empirical evidence, especially the latest information on cosmic-ray mesons, may be considered on the whole to support the mixed theory [2]. Still, in conclusion, we wish to emphasize once more its very serious limitations, already explicitly stated in our previous paper ([1], especially pp. 51—52). In particular, it is out of question that such problems as the scattering of fast mesons [8] could fall within the scope of this or any other form of meson theory as long as the deep-lying problems connected with the universal limiting length (denoted by r_0 in [1]) are not brought nearer to their solution.

* In an analogous way, the static self-energy of a point nucleon may be considered a contribution to its mass.

PART I.

Electromagnetic properties of arbitrary fields.

§ 1. Gauge-invariance of Lagrangian.

Let us consider any field representing "charged" particles, *i. e.* particles capable of interaction with the electromagnetic field. If we assume that the force on such a charged particle due to the electromagnetic field is given by the familiar Lorentz expression, we may, according to an interesting theorem pointed out by RACAŃ [9], conclude that there exists a Lagrangian from which this force can be derived. In fact, to be derivable from a Lagrangian, the expressions of the force components have to satisfy a set of necessary and sufficient conditions established by HELMHOLTZ [10]; and RACAŃ has shown that, if we also assume the force to be independent of the acceleration of the particle, these conditions just reduce to stating that it takes the form of the Lorentz expression.

We may thus start from a Lagrangian density

$$\mathcal{L}(Q_\omega, Q_{\omega|i}; Q_\omega^\dagger, Q_{\omega|i}^\dagger; A_l, A_{l|i})$$

depending on the variables Q_ω , Q_ω^\dagger , A_l and their derivatives $F_{|i} \equiv \frac{\partial F}{\partial x^i}$ ($x^i \equiv x, y, z, ct$); while the A_l denote the components of the electromagnetic potential (vector potential $\vec{A} \equiv A_1, A_2, A_3 = A^1, A^2, A^3$; scalar potential $B \equiv A^4 = -A_4$), the Q_ω represent the quantized variables defining the various types of charged particles considered, Q_ω^\dagger their Hermitian conjugates, numbered in an arbitrary way by the index ω . As is well-known [11]*,

* It will be noticed that the following analysis is entirely analogous to that used for the derivation of the energy-momentum tensor from the invariance of the Lagrangian for arbitrary coordinate transformations [12]. The possibility of extending this method to any group, and in particular to the gauge-transformation group, has been pointed out to us by Dr. J. PODOLANSKI.

the form of this Lagrangian giving rise to the required expression of the Lorentz force may simply be obtained by postulating the invariance of the Lagrangian for the so-called gauge-transformations of the form

$$\begin{cases} Q'_\omega = e \frac{i e_\omega \alpha}{\hbar c} Q_\omega \\ A'_l = A_l + \alpha_{|l}, \end{cases} \quad (1)$$

in which α is an arbitrary c -number function and the constants e_ω , referring to the different kinds of particles considered, are to be identified with the charges of these particles. In fact, the gauge-invariance of any functional, such as \mathcal{L} , may be expressed in the form*

$$\begin{aligned} \delta \mathcal{L} \equiv & \sum_\omega \left\{ \frac{\delta \mathcal{L}}{\delta Q_\omega} \delta Q_\omega + \frac{\partial}{\partial x^i} \left(\frac{\partial \mathcal{L}}{\partial Q_{\omega|i}} \delta Q_\omega \right) + \text{conj.} \right\} \\ & + \frac{\delta \mathcal{L}}{\delta A_i} \delta A_i + \frac{\partial}{\partial x^j} \left(\frac{\partial \mathcal{L}}{\partial A_{i|j}} \delta A_i \right) = 0, \end{aligned}$$

where $\frac{\delta \mathcal{L}}{\delta Q}$ represents the "variational derivative" $\frac{\partial \mathcal{L}}{\partial Q} - \frac{\partial}{\partial x^i} \left(\frac{\partial \mathcal{L}}{\partial Q_{|i}} \right)$.

Putting

$$s^i = \frac{1}{i} \sum_\omega \frac{e_\omega}{\hbar c} \left[\frac{\partial \mathcal{L}}{\partial Q_{\omega|i}} Q_\omega - Q_\omega^\dagger \frac{\partial \mathcal{L}}{\partial Q_{\omega^\dagger|i}} \right], \quad (2)$$

we then get, for an infinitesimal value of the parameter α ,

$$\begin{aligned} \delta \mathcal{L} \equiv & \alpha \left\{ \sum_\omega i \frac{e_\omega}{\hbar c} \frac{\delta \mathcal{L}}{\delta Q_\omega} Q_\omega + \text{conj.} - \frac{\partial s^i}{\partial x^i} \right\} \\ & + \alpha_{|i} \left\{ \frac{\delta \mathcal{L}}{\delta A_i} + \frac{\partial}{\partial x^j} \left(\frac{\partial \mathcal{L}}{\partial A_{i|j}} \right) - s^i \right\} + \alpha_{|i|j} \frac{\partial \mathcal{L}}{\partial A_{i|j}} = 0. \end{aligned} \quad (3)$$

In connexion with the possible non-commutability of Q_ω and Q_ω^\dagger , it has here been assumed that in all terms of \mathcal{L} any factor Q_ω^\dagger or $Q_{\omega^\dagger|i}$ is written on the left of any Q_ω or $Q_{\omega|i}$; if $Q_{\omega|i}$ is not commutable with Q_ω , further restrictions must be imposed

* The abbreviation "conj." denotes the Hermitian conjugate of the expression preceding it.

on the form of the Lagrangian if we wish to write down a general expression such as (3) for $\delta \mathcal{L}$. We will make the assumption, fulfilled in all cases of actual interest, that the expression (3) is valid in terms of q -number variables also. We then derive from it the three following conditions for the Lagrangian density \mathcal{L} , viz.:

1° \mathcal{L} depends on the $A_{i|j}$ only through the field components

$$F_{ij} = A_{j|i} - A_{i|j}; \quad (4)$$

2° we have therefore

$$\frac{\delta \mathcal{L}}{\delta A_i} = s^i + \frac{\partial}{\partial x^j} \left(\frac{\partial \mathcal{L}}{\partial F_{ij}} \right); \quad (5)$$

3° we have

$$\frac{\partial s^i}{\partial x^i} = i \sum_{\omega} \frac{e_{\omega}}{\hbar c} \frac{\delta \mathcal{L}}{\delta Q_{\omega}} Q_{\omega} + \text{conj.} \quad (6)$$

Now, the total Lagrangian density of a system of charged particles and electromagnetic field includes, besides the term hitherto considered which refers to the particles and their interaction with the electromagnetic field, a term

$$\mathcal{L}_f = -\frac{1}{4} F^{ij} F_{ij} \quad (7)$$

describing the field itself. We may then write the field equations for the material variables Q_{ω} , Q_{ω}^{\dagger} in the form

$$\frac{\delta \mathcal{L}}{\delta Q_{\omega}} = 0, \quad \frac{\delta \mathcal{L}}{\delta Q_{\omega}^{\dagger}} = 0 \quad (8)$$

and the electromagnetic field equations in the form

$$\frac{\delta \mathcal{L}_f}{\delta A_i} = -\frac{\delta \mathcal{L}}{\delta A_i}$$

or

$$\frac{\partial F^{ij}}{\partial x^j} = \frac{\delta \mathcal{L}}{\delta A_i}. \quad (9)$$

From the latter it follows that $\frac{\delta \mathcal{L}}{\delta A_i}$ is to be interpreted as the total charge-current density of the system, the general expression of which is thus given by (5). In this formula, s^i , given by (2), is the usual charge-current density, whereas the last term represents a polarization current density, as would arise from an electric and magnetic moment of the particles through which they would directly interact with the electromagnetic field; such direct interaction we shall, however, disregard and we shall, thus, in the following assume that \mathcal{L} only depends on the electromagnetic potential. In any case, since the divergence of the polarization current density automatically vanishes, the conservation law

$$\frac{\partial}{\partial x^i} \frac{\delta \mathcal{L}}{\delta A_i} = 0 \quad \text{or} \quad \frac{\partial s^i}{\partial x^i} = 0 \quad (10)$$

follows from (6) in virtue of the field equations (8).

In order to derive the expression of \mathcal{L} in terms of the A_i , we start from the condition (5) which may be written more simply

$$\frac{\partial \mathcal{L}}{\partial A_i} = s^i \quad (5')$$

and we further require $\mathcal{L}(Q_\omega, Q_{\omega|i}; Q_\omega^\dagger, Q_{\omega|i}^\dagger; A_l)$ to reduce for $A_l \rightarrow 0$ to the form \mathcal{L}° of the Lagrangian density for the particles concerned in the absence of electromagnetic field. It is now easy to show that these two requirements fix the form of \mathcal{L} to*

$$\mathcal{L} = \mathcal{L}^\circ \left(Q_\omega, Q_{\omega|i} - i \frac{e_\omega}{\hbar c} A_i Q_\omega; Q_\omega^\dagger, Q_{\omega|i}^\dagger + i \frac{e_\omega}{\hbar c} A_i Q_\omega^\dagger \right). \quad (11)$$

This expresses, of course, the well-known result that the Lagrangian density for charged particles in interaction with the electromagnetic field is obtained from the Lagrangian for no field by replacing the operator $\frac{\partial}{\partial x^i}$ by $\frac{\partial}{\partial x^i} - i \frac{e_\omega}{\hbar c} A_i$.

* In the first place, the function (11) satisfies the two conditions. On the other hand, if a function of $Q_\omega, Q_{\omega|i}; Q_\omega^\dagger, Q_{\omega|i}^\dagger; A_l$ satisfies (5') and reduces to zero for $A_l \rightarrow 0$, it is readily seen that it must be identically zero.

§ 2. Transition to Hamiltonian and separation of Coulomb field [13].

Let $P_\omega \equiv \frac{\partial \mathcal{L}}{\partial Q_{\omega|4}}$ be the momentum canonically conjugate to Q_ω ; if \mathcal{L} is Hermitian (which we will assume), the momentum canonically conjugate to Q_ω^\dagger is then P_ω^\dagger . From the gauge-invariance of \mathcal{L} it follows immediately that a gauge-transformation (1) transforms the variables P_ω according to the equation

$$P'_\omega = e^{-i \frac{e_\omega}{\hbar c} \alpha} P_\omega.$$

Let further $\mathcal{H}_m^\circ(Q_\omega, \text{grad } Q_\omega, P_\omega^\circ; Q_\omega^\dagger, \text{grad } Q_\omega^\dagger, P_\omega^{\circ\dagger})$ be the Hamiltonian function with no electromagnetic field present, defined in the usual way by the equation

$$\mathcal{H}_m^\circ = \sum_\omega P_\omega^\circ Q_{\omega|4} + \text{conj.} - \mathcal{L}^\circ$$

from which the variables $Q_{\omega|4}$ are understood to be eliminated by means of the equations

$$P_\omega^\circ = \frac{\partial}{\partial Q_{\omega|4}} \mathcal{L}^\circ(Q_\omega, Q_{\omega|i}; Q_\omega^\dagger, Q_{\omega|i}^\dagger).$$

The Hamiltonian with the field present, as derived from the Lagrangian (11), is then easily seen to be*

$$\mathcal{H}_m + \int B \varrho \, dv,$$

where

$$\varrho \equiv s^4 = \frac{1}{i} \sum_\omega \frac{e_\omega}{\hbar c} [P_\omega Q_\omega - Q_\omega^\dagger P_\omega^\dagger]$$

is the charge density and

$$\mathcal{H}_m = \mathcal{H}_m^\circ \left(Q_\omega, \left(\text{grad} - \frac{i e_\omega \vec{A}}{\hbar c} \right) Q_\omega, P_\omega; \dots \dots \right) \quad (12)$$

the gauge-invariant part of the Hamiltonian. Similarly, the Hamiltonian of the pure electromagnetic field (electric field \vec{E} , magnetic field \vec{H}) derived from the Lagrangian (7) is

* $\int \dots \dots \, dv$ denotes a volume integration over the whole space.

$$\mathcal{H}_j - \int B \operatorname{div} \vec{E} dv,$$

with

$$\mathcal{H}_j = \frac{1}{2} \int (\vec{E}^2 + \vec{H}^2) dv; \quad (12')$$

the momentum canonically conjugate to \vec{A} is $-\vec{E}$, while that canonically conjugate to B is identically zero. The Hamiltonian of the total system may therefore be written

$$\mathcal{H} = \mathcal{H}_j + \mathcal{H}_m - \int B \mathcal{C} dv, \quad (13)$$

with

$$\mathcal{C} = \operatorname{div} \vec{E} - \rho. \quad (13')$$

The Hamiltonian formalism must be completed, in the usual way, by the fundamental commutation rules and the accessory condition

$$\mathcal{C} = 0 \quad (14)$$

to be imposed on the wave-function describing the state of the system. While the variables \vec{A} and $-\vec{E}$ satisfy the ordinary canonical commutation rules, the variable B commutes with all other electromagnetic field quantities.

Let us now apply to the Hamiltonian \mathcal{H} the canonical transformation corresponding to an arbitrary change of phase of the variables Q_ω, P_ω :

$$Q'_\omega = e^{i \frac{e\omega}{\hbar c} \alpha} Q_\omega, \quad P'_\omega = e^{-i \frac{e\omega}{\hbar c} \alpha} P_\omega, \quad (15)$$

whereby α may be any operator independent of Q_ω, P_ω and \vec{E} (but eventually containing the electromagnetic potentials); such a transformation may, in fact, be put into the form*

$$Q'_\omega = \mathcal{O}^{-1} Q_\omega \mathcal{O}, \quad P'_\omega = \mathcal{O}^{-1} P_\omega \mathcal{O}, \quad (15)$$

with

$$\mathcal{O} = e^{\frac{i}{\hbar c} \int \rho \alpha dv}. \quad (16)$$

* This may readily be verified for tensor variables (satisfying canonical commutation rules) as well as for spinor variables (which obey commutation rules corresponding to the exclusion principle).

Expressed in the new (primed) variables, the part \mathcal{H}_m of the Hamiltonian is given by

$$\mathcal{H}_{m(\text{new var.})} = \mathcal{S}' \mathcal{H}'_m \mathcal{S}'^{-1},$$

where \mathcal{H}'_m , in accordance with a general convention made in this paper, denotes the same function of the new variables as the function \mathcal{H}_m of the old variables. Now, the gauge-invariance of \mathcal{H}_m may be stated in the form

$$\mathcal{H}_m(\vec{A}, \dots) = \mathcal{S}^{-1} \mathcal{H}_m(\vec{A} + \text{grad } \alpha, \dots) \mathcal{S},$$

from which it follows, according to (12), that

$$\left. \begin{aligned} & \mathcal{H}_{m(\text{new var.})} = \\ & = \mathcal{H}'_m \left(Q'_\omega, \left(\text{grad} - i \frac{e_\omega}{\hbar c} \vec{A}' - i \frac{e_\omega}{\hbar c} \text{grad } \alpha \right) Q'_\omega, P'_\omega; \dots \right). \end{aligned} \right\} (17)$$

As to the other terms of the Hamiltonian (13), they may be written, if we put

$$\mathcal{S} \vec{E} \mathcal{S}^{-1} = \vec{E} + \vec{E}_0 \tag{18}$$

and

$$\mathcal{Q} = \frac{1}{2} \int \vec{E}_0^2 dv, \tag{19}$$

in the form

$$\mathcal{H}_{l(\text{new var.})} = \mathcal{H}'_l + \frac{1}{2} \int (\vec{E}' \vec{E}'_0 + \vec{E}'_0 \vec{E}') dv + \mathcal{Q}' \tag{20}$$

and

$$- \int B \mathcal{C}_{(\text{new var.})} dv, \quad \text{with} \quad \mathcal{C}_{(\text{new var.})} = \text{div } \vec{E}' + \text{div } \vec{E}'_0 - \rho'. \tag{21}$$

Now, we can choose α in such a way that the field \vec{E}_0 separated from the total field by the canonical transformation \mathcal{S} be just the Coulomb field

$$\left\{ \begin{aligned} \vec{E}_0 &= -\text{grad} \int \rho \varphi_0 dv, & \varphi_0 &= \frac{1}{4\pi r}; \\ \text{div } \vec{E}_0 &= \rho. \end{aligned} \right. \tag{22}$$

We have only to take

$$\alpha = \int \operatorname{div} \vec{A} \cdot \varphi_0 dv, \quad (23)$$

whence

$$\mathcal{E} = e \frac{i}{\hbar c} \int \vec{E}_0 \cdot \vec{A} dv. \quad (24)$$

The new field variable \vec{E}' then represents the transverse part of the electric field; since, according to (23),

$$\operatorname{rot}(-\operatorname{grad} \alpha) = 0, \quad \operatorname{div}(-\operatorname{grad} \alpha) = -\int \operatorname{div} \vec{A} \cdot \Delta \varphi_0 dv = \operatorname{div} \vec{A},$$

we also see that $-\operatorname{grad} \alpha$ may be interpreted as the longitudinal part of the vector-potential \vec{A} , so that in the new Hamiltonian (17) only the transverse part

$$\vec{A}_\perp = \vec{A} + \operatorname{grad} \alpha$$

(or rather \vec{A}'_\perp) occurs, and this variable \vec{A}'_\perp also defines the magnetic field by $\vec{H}' = \operatorname{rot} \vec{A}'_\perp$. On account of the condition (14), the value of B in the Hamiltonian (13) may be chosen arbitrarily; taking for it the Coulomb potential $\int \varrho \varphi_0 dv$, one causes the terms

$$\frac{1}{2} \int (\vec{E}' \vec{E}'_0 + \vec{E}'_0 \vec{E}') dv = \int \vec{E}' \vec{E}'_0 dv$$

and

$$-\int B \mathcal{C}_{(\text{new var.})} dv = -\int B \operatorname{div} \vec{E}' dv$$

to cancel each other. Finally, the term \mathcal{Q}' simply represents (in the primed variables) the Coulomb energy

$$\mathcal{Q}' = \frac{1}{2} \int \varrho(\vec{x}) \varrho(\vec{x}') \varphi_0(|\vec{x} - \vec{x}'|) dv dv',$$

so that we get a Hamiltonian

$$\left. \begin{aligned} & \mathcal{H}_{(\text{new var.})} = \\ & = \mathcal{H}_m^\circ \left(Q'_\omega, \left(\operatorname{grad} - \frac{i e_\omega \vec{A}'_\perp}{\hbar c} \right) Q'_\omega, P'_\omega; \dots \dots \right) + \mathcal{H}'_f + \mathcal{Q}' \end{aligned} \right\} (25)$$

from which the longitudinal parts of electric field and vector potential are entirely eliminated. Also the accessory condition

$$\mathcal{C}_{(\text{new var.})} \equiv \text{div } \vec{E}' = 0 \quad (26)$$

does no more contain the longitudinal field; but it must not be forgotten that \vec{A}' , and not \vec{A}'_{\perp} , is canonically conjugate to $-\vec{E}'$.

§ 3. Hermitian variables.

Chiefly when dealing with meson fields it may be advantageous to use Hermitian variables for the description of the charged particles. As such, we may conveniently take

$$Q_{\mathbf{1}} = \frac{1}{\sqrt{2}}(Q + Q^{\dagger}), \quad Q_{\mathbf{2}} = \frac{1}{i\sqrt{2}}(Q - Q^{\dagger}),$$

so that

$$Q = \frac{1}{\sqrt{2}}(Q_{\mathbf{1}} + iQ_{\mathbf{2}}). \quad (27)$$

Considering these formally as "components" of a "vector" \mathcal{Q} along the directions of orthogonal axes $\mathbf{1}$ and $\mathbf{2}$, we may interpret the phase transformation $Q' = e^{i\alpha} Q$ as a "rotation" of angle α about the symbolical axis $\mathbf{3}$, perpendicular to $\mathbf{1}$ and $\mathbf{2}$. Any (Hermitian) "component" $Q_{\mathbf{3}}$ in the direction of the axis $\mathbf{3}$ can be assumed to represent a neutral particle of the same kind. Taking the vector potential (and all electromagnetic field components) along the axis $\mathbf{3}$:

$$\vec{A} \equiv (0, 0, A),$$

we may then express the gauge-invariant derivatives

$$Q_{|l} - \frac{ie}{\hbar c} A_l Q, \quad Q_{|l}^{\dagger} + \frac{ie}{\hbar c} A_l Q^{\dagger},$$

through which the interaction with the electromagnetic field is introduced into the Hamiltonian, in the form

$$\mathcal{Q}_{|l} - \frac{e}{\hbar c} \mathbf{A}_l \mathbf{\Lambda} \mathcal{Q},$$

the sign $\mathbf{\Lambda}$ denoting a vector product in symbolical space. Calling finally \mathbf{P} the momenta canonically conjugate to \mathcal{Q} , we have

$$P = \frac{1}{\sqrt{2}} (P_{\mathbf{1}} - iP_{\mathbf{2}}),$$

and, since the Lagrangian density \mathcal{L} has been assumed to be Hermitian, $P_{\mathbf{1}}$ and $P_{\mathbf{2}}$ are also Hermitian.

For the electric charge density, we have, according to (2), the general formula*

$$\left. \begin{aligned} \varrho &= \frac{1}{i} \sum_{\omega} \frac{e_{\omega}}{\hbar c} (P_{\omega} Q_{\omega} - Q_{\omega}^{\dagger} P_{\omega}^{\dagger}) \\ &= \sum_{\omega} \frac{e_{\omega}}{\hbar c} \left\{ \frac{1}{2} [\mathbf{P}_{\omega} \wedge \mathbf{Q}_{\omega} - \mathbf{Q}_{\omega} \wedge \mathbf{P}_{\omega}]_{\mathbf{3}} + \right. \\ &\quad \left. + \frac{1}{2i} [[P_{\omega \mathbf{1}}, Q_{\omega \mathbf{1}}]_{-} + [P_{\omega \mathbf{2}}, Q_{\omega \mathbf{2}}]_{-}] \right\}; \end{aligned} \right\} \quad (28)$$

for the current density \vec{I} , analogous formulae hold, \mathbf{P}_{ω} being replaced by $\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_{\omega i}}$ ($i = 1, 2, 3$). In order further to discuss these expressions, we must specify the commutation rules assumed for the \mathbf{P}_{ω} and \mathbf{Q}_{ω} . Let us first consider a field, describing particles of charge e , which obeys the canonical commutation rules

$$[P_{\omega i}(\vec{x}), Q_{\omega' i'}(\vec{x}')]_{-} = \frac{\hbar c}{i} \delta(\vec{x} - \vec{x}') \delta_{\omega\omega'} \delta_{ii'}, \quad (29)$$

the index ω referring to the different components of the field. The last term in ϱ , which occurs for every given state of the field, then would represent an infinite additional charge density of this state; such a term, however, can easily be avoided by a slight modification of the Lagrangian density \mathcal{L} ; we have only to replace \mathcal{L} by half the sum of \mathcal{L} itself and the expression \mathcal{L}' obtained from it by reversing the order of all factors** $Q, Q_{|i}, Q^{\dagger}, Q_{|i}^{\dagger}$. We may, therefore, disregard the infinite "zero-point" charge density altogether and write, in this case,

* We use the notation $[A, B]_{\pm} = AB \pm BA$.

** In fact, if we repeat for \mathcal{L}' the considerations of § 1, we find for $\partial \mathcal{L}'$ and, consequently, for the charge-current density expressions in which the order of factors is automatically reversed. We may further assume that $\frac{\partial \mathcal{L}'}{\partial Q_{\omega i}} = \frac{\partial \mathcal{L}}{\partial Q_{\omega i}}$, a condition which is fulfilled in all cases of actual interest. We then get

$$\left. \begin{aligned} \varrho &= \frac{e}{\hbar c} \sum_{\omega} (\mathbf{P}_{\omega} \wedge \mathbf{Q}_{\omega})_{\mathbf{3}} \\ I_i &= \frac{e}{\hbar c} \sum_{\omega} \left(\frac{\partial \mathcal{L}}{\partial \mathbf{Q}_{\omega|i}} \wedge \mathbf{Q}_{\omega} \right)_{\mathbf{3}} \end{aligned} \right\} \quad (30)$$

As an example of a field of this type, we may of course mention the meson field to be extensively treated in the following.

For the sake of completeness, we shall still briefly examine the case of Dirac particles (electrons or nucleons) described by four component wave-functions satisfying the commutation rules which correspond to the exclusion principle. The decomposition (27) of such variables into Hermitian constituents has been especially discussed, in connexion with the theory of electron pairs, by MAJORANA [14]; it will thus suffice here to write down the main formulae of this theory without entering into details as to their interpretation. The commutation rules may be stated as follows:

$$\begin{aligned} [Q_{\omega\mathbf{1}}(\vec{x}), Q_{\omega'\mathbf{1}}(\vec{x}')]_+ &= [Q_{\omega\mathbf{2}}(\vec{x}), Q_{\omega'\mathbf{2}}(\vec{x}')]_+ = \delta(\vec{x} - \vec{x}') \delta_{\omega\omega'}, \\ \text{all } Q_{\omega\mathbf{1}} &\text{ anticommute with all } Q_{\omega\mathbf{2}}; \end{aligned}$$

the indices $\omega, \omega' (= 1, 2, 3, 4)$ refer to the components of the (spinor) wave-function; we may in the usual way consider the complex wave-function Q or its Hermitian constituents \mathbf{Q} as matrices, with respect to ω , of 4 rows and 1 column. For any Hermitian operator \mathcal{G} , we may then write*

$$\varrho = \frac{1}{i} \sum_{\omega} \frac{e_{\omega}}{\hbar c} \left\{ (P_{\omega} Q_{\omega} - Q_{\omega}^{\dagger} P_{\omega}^{\dagger}) + (Q_{\omega} P_{\omega} - P_{\omega}^{\dagger} Q_{\omega}^{\dagger}) \right\},$$

which eliminates the zero-point charge density.

In this special case, it would also be possible (cf. [15]) to define the charge density by the expression

$$\varrho = \frac{1}{i} \sum_{\omega} \frac{e_{\omega}}{\hbar c} (P_{\omega} Q_{\omega} - P_{\omega}^{\dagger} Q_{\omega}^{\dagger}),$$

which does not contain a zero-point density either; this form, however, has the disadvantage of not being Hermitian by itself, but only on account of the commutation rules.

* \tilde{A} denotes the transposed of A , and A^* its complex conjugate; one has $A^{\dagger} = \tilde{A}^*$.

$$\int Q^\dagger \mathcal{O} Q dv = \\ = \frac{1}{2} \int \tilde{Q}_1 \mathcal{O} Q_1 dv + \frac{1}{2} \int \tilde{Q}_2 \mathcal{O} Q_2 dv + \frac{i}{2} \int \tilde{Q}_1 (\mathcal{O} + \mathcal{O}^*) Q_2 dv;$$

in particular, for a purely imaginary operator, the last term disappears. Choosing for instance, with MAJORANA, a representation of the Dirac matrices for which the Hamiltonian \mathcal{H}_m° is purely imaginary, we get for the total energy of a system of charged (and eventually neutral) Dirac particles the expression

$$\frac{1}{2} \int \tilde{\mathcal{Q}} \mathcal{H}_m^\circ \mathcal{Q} dv$$

used by MAJORANA. Since

$$P_1 = \frac{1}{2} \hbar c \tilde{Q}_2, \quad P_2 = -\frac{1}{2} \hbar c \tilde{Q}_1,$$

we may now write

$$\sum_{\omega=1}^4 [P_{\omega 1}, Q_{\omega 1}]_- = \sum_{\omega=1}^4 [P_{\omega 2}, Q_{\omega 2}]_- = -\frac{1}{2} \hbar c (\tilde{\mathcal{Q}} \wedge \mathcal{Q})_3;$$

the last term of (28) now gives

$$q = ie (\tilde{\mathcal{Q}} \wedge \mathcal{Q})_3,$$

while the first would, on account of the commutation rules, contribute an infinite "zero-point charge density"; but this may again be avoided by taking instead of \mathcal{L} half the difference of \mathcal{L} and \mathcal{L}' . For the current density, we find

$$\vec{I} = ie (\tilde{\mathcal{Q}} \vec{\alpha} \wedge \mathcal{Q})_3,$$

$\vec{\alpha}$ denoting the well-known velocity matrices of the Dirac electron theory. All these formulae coincide, of course, with those of MAJORANA.

§ 4. System in slowly varying external field.

An important class of problems is concerned with the behaviour of a system of charged particles under the action of

an external electromagnetic field. We have then to introduce into the Hamiltonian of the system additional terms expressing the energy of interaction between this field and the particles considered. Let us in this section denote by $\vec{A}, B; \vec{E}, \vec{H}$ the variables of the external field. On account of (5') we may write, for the part of the interaction energy depending linearly on \vec{A} ,

$$-\int \vec{A} \vec{I} dv + \int B \rho dv, \quad (31)$$

where we must take for the charge and current densities \vec{I}, ρ their expressions in absence of any electromagnetic field (for $\vec{A} = 0$): we shall neglect any contribution of higher order in \vec{A} . The expression (31) is not modified by the canonical transformation \mathcal{S}^0 effecting the separation of the Coulomb field, so that, after this separation has been performed, it keeps the same form in terms of the new variables; we may here without danger of confusion omit the primes indicating functions of the new variables. More particularly, we shall consider a "slowly varying" field, *i. e.* assume that the relative variations of the field components over distances of the order of the dimensions of the system and times of the order of the proper periods connected with the system are small compared with unity. In such a case, we may put the expression of the interaction of the system with the external field into a more convenient form depending on the values of the electromagnetic field and its derivatives at some arbitrarily chosen point O and on the successive 2^n -pole moments of the system with respect to this point. The properties of the system to any desired approximation are then easily derived from the expressions of such 2^n -pole moments, so that the problem is essentially reduced to the calculation of these operators.

We shall now carry out the transformation of the interaction operator up to the second order of approximation. If we further assume that the velocities of the particles are small compared with the velocity of light, *i. e.* that the dimensions of the system are small compared with the proper wave-lengths, this approximation involves

$$\left. \begin{array}{l}
 \text{the total charge} \\
 \text{the electric dipole moment} \\
 \text{the magnetic dipole moment} \\
 \text{and} \\
 \text{the electric quadrupole moment*}
 \end{array} \right\} \begin{array}{l}
 e = \int \rho \, dv, \\
 \vec{P} = \int \rho \vec{x} \, dv, \\
 \vec{M} = \frac{1}{2} \int \vec{x} \wedge \vec{I} \, dv, \\
 Q^{ik} = \frac{1}{2} \int \rho x^i x^k \, dv
 \end{array} \quad (32)$$

of the system, the vector \vec{x} being taken from O as origin. For this purpose, we start from the expansions

$$\left\{ \begin{array}{l}
 B = B_0 + \vec{x} \operatorname{grad}_0 B_0 + \frac{1}{2} \vec{x} (\vec{x} \operatorname{grad}_0) \operatorname{grad}_0 B_0 + R \\
 \vec{A} = \vec{A}_0 + (\vec{x} \operatorname{grad}_0) \vec{A}_0 + \vec{R}',
 \end{array} \right.$$

the residuals R, \vec{R}' containing higher derivatives of the potentials with respect to the point O . Remembering further that**

$$\int \vec{I} \, dv = \dot{\vec{P}},$$

we therefore get***

$$\left\{ \begin{array}{l}
 \int B \rho \, dv = e B_0 + \dot{\vec{P}} \operatorname{grad}_0 B_0 + (Q \operatorname{grad}_0) \operatorname{grad}_0 B_0 + R'' \\
 \int \vec{A} \vec{I} \, dv = \dot{\vec{P}} \vec{A}_0 + \int \vec{I} (\vec{x} \operatorname{grad}_0) \vec{A}_0 \, dv + \vec{R}''',
 \end{array} \right.$$

the residuals R'', \vec{R}''' being of higher order of approximation than the second. Now,

$$\begin{aligned}
 (\vec{x} \operatorname{grad}_0) \vec{A}_0 &= -(\vec{A}_0 \operatorname{grad}_0) \vec{x} + \operatorname{grad}_0 (\vec{A}_0 \vec{x}) - \vec{x} \wedge \operatorname{rot}_0 \vec{A}_0 \\
 &= \vec{A}_0 + \operatorname{grad}_0 (\vec{A}_0 \vec{x}) - \vec{x} \wedge \vec{H}_0;
 \end{aligned}$$

and

* See the footnote on p. 21.

** We use the notation $\dot{A} = \frac{\partial A}{\partial x^4}$ (or $\frac{dA}{dx^4}$).

*** If Q is a tensor and \vec{u} a vector, $Q\vec{u}$ represents the vector with components $\sum_k Q^{ik} u_k$.

$$\begin{aligned} \vec{I}(\vec{x} \text{ grad}_0) \vec{A}_0 &= \vec{I} \vec{A}_0 + \vec{x} (\vec{I} \text{ grad}_0) \vec{A}_0 + \vec{A}_0 (\vec{I} \text{ grad}_0) \vec{x} - (\vec{I} \wedge \vec{x}) \vec{H}_0 \\ &= \vec{x} (\vec{I} \text{ grad}_0) \vec{A}_0 + (\vec{x} \wedge \vec{I}) \vec{H}_0; \end{aligned}$$

on the other hand,

$$\vec{I}(\vec{x} \text{ grad}_0) \vec{A}_0 + \vec{x} (\vec{I} \text{ grad}_0) \vec{A}_0 = \vec{I} \text{ grad} \left\{ \vec{x} (\vec{x} \text{ grad}_0) \vec{A}_0 \right\}.$$

We may therefore write

$$\vec{I}(\vec{x} \text{ grad}_0) \vec{A}_0 = \frac{1}{2} (\vec{x} \wedge \vec{I}) \vec{H}_0 + \frac{1}{2} \vec{I} \text{ grad} \left\{ \vec{x} (\vec{x} \text{ grad}_0) \vec{A}_0 \right\},$$

whence

$$\begin{aligned} \int \vec{I}(\vec{x} \text{ grad}_0) \vec{A}_0 dv &= \vec{M} \vec{H}_0 - \frac{1}{2} \int \text{div} \vec{I} \cdot \vec{x} (\vec{x} \text{ grad}_0) \vec{A}_0 dv \\ &= \vec{M} \vec{H}_0 + (\dot{Q} \text{ grad}_0) \vec{A}_0. \end{aligned}$$

With this result, we thus get

$$\begin{aligned} - \int \vec{A} \vec{I} dv &= - \dot{\vec{P}} \vec{A}_0 - \vec{M} \vec{H}_0 - (\dot{Q} \text{ grad}_0) \vec{A}_0 - \vec{R}''' \\ &= \vec{P} \dot{\vec{A}}_0 - \vec{M} \vec{H}_0 + (Q \text{ grad}_0) \dot{\vec{A}}_0 \\ &\quad - \frac{d}{c dt} \left[\vec{P} \vec{A}_0 + (Q \text{ grad}_0) \vec{A}_0 \right] - \vec{R}''' \end{aligned}$$

and, finally,

$$\left. \begin{aligned} & - \int \vec{A} \vec{I} dv + \int B \varrho dv = \\ & = e B_0 - \vec{P} \vec{E}_0 - \vec{M} \vec{H}_0 - (Q \text{ grad}_0) \vec{E}_0 \\ & - \frac{d}{c dt} \left[\vec{P} \vec{A}_0 + (Q \text{ grad}_0) \vec{A}_0 \right] + \dots \dots \dots, \end{aligned} \right\} \quad (33)$$

* It should be pointed out that the usual (and, from a systematic point of view, more rational) definition of the quadrupole moment would be expressed in our notation by $Q^{ik} - \frac{1}{3} \delta^{ik} \sum_l Q^{ll}$. This makes, however, no difference for the interaction term $-(Q \text{ grad}_0) \vec{E}_0$ since, for an external field, $\text{div}_0 \vec{E}_0 = 0$.

the residual terms being of higher order. All terms but the last in formula (33) are familiar expressions, the interpretation of which it is superfluous to recall*. As regards the last term, it follows from its being the time derivative of an operator that it does not give any contribution either to the mean value of the interaction energy in stationary states of the system or to the matrix elements corresponding to transitions between states of the total system with the same energy.

* See the footnote on p. 21.

PART II.

Electromagnetic properties of a system of nucleons and mesons.

In our previous paper [1]*, arguments have been developed for assuming that nucleons produce two independent kinds of meson fields, *viz.* vector and pseudoscalar fields, each consisting of both (positively and negatively) charged and neutral mesons. From the standpoint of this "mixed" theory, we shall now investigate the interaction of a system of nucleons and of such meson fields with the electromagnetic field. In the first place, we shall derive, by a direct application of the general considerations of Part I of the present paper, the expression of the Hamiltonian function of the total system. To this Hamiltonian we shall then apply the canonical transformation effecting, as explained in NF, the separation of the static meson fields produced by the nucleons and we shall briefly discuss the different electromagnetic interaction terms obtained in this way; we shall especially fix our attention on the interaction with an external electromagnetic field.

§ 1. Hamiltonian of the total system.

The expression of the total Hamiltonian, after separation of the Coulomb field, may immediately be obtained from formula (25) of Part I (in which we may, from now on, omit the primes denoting the "new" variables). For the description of the nucleons and meson fields, we shall use the notations of NF, to which we beg the reader to refer for the explanation of their meaning; to begin with, we have to do with the variables

* In the following, this paper will be referred to as NF, its formula (n) as NF(n).

originally introduced there before the separation of the static parts of the meson fields. Treating the nucleons as Dirac particles, *i. e.* attributing to them the Hamiltonian \mathcal{H}_k defined by formula NF (6), we get, according to formulae (25) and (31) of Part I, an additional term of electromagnetic interaction

$$-\int \vec{I}_{\text{nucl}} \cdot \vec{A} dv,$$

where \vec{A} represents the sum of the transversal vector-potential \vec{A}_\perp of the field considered a part of the system and of the vector-potential \vec{A}_{ex} of any "external" field eventually acting on it. The current-density \vec{I}_{nucl} of the nucleons may be expressed, with our choice of variables, by

$$\vec{I}_{\text{nucl}} = e \sum_i \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \vec{\alpha}^{(i)} \delta(\vec{x} - \vec{x}^{(i)}). \quad (1)$$

If there is also an external scalar potential B_{ex} , a further term

$$\int \varrho_{\text{nucl}} \cdot B_{\text{ex}} dv$$

arises, in which the charge density ϱ_{nucl} is given by

$$\varrho_{\text{nucl}} = e \sum_i \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \delta(\vec{x} - \vec{x}^{(i)}). \quad (2)$$

As regards the meson fields, let us start from the Lagrangian density with no electromagnetic field present (which had not been stated explicitly in NF):

$$\mathcal{L}^0 = \left. \begin{aligned} & -\frac{1}{2} \left\{ \kappa^2 (\vec{U}^2 - \mathbf{V}^2) + \vec{G}^2 - \vec{F}^2 \right\} + \vec{M} \vec{U} - \mathbf{N} \mathbf{V} \\ & -\frac{1}{2} \left\{ \kappa^2 \Psi^2 + \vec{\Gamma}^2 - \Phi^2 \right\} + \mathbf{R} \Psi + \frac{1}{2} (\vec{P}^2 - \mathcal{Q}^2), \end{aligned} \right\} \quad (3)$$

in which the variables \mathcal{Q}_ω are \vec{U} , \mathbf{V} and Ψ , while \vec{G} , \vec{F} , $\vec{\Gamma}$, Φ are defined by the second formula NF (2), the first NF (1), NF (22), and the first NF (21), respectively. Following the procedure of Part I, we first see that $-\vec{F}$ and Φ are canonically conjugate

to \vec{U} and Ψ , respectively, while the variable \mathbf{V} , the conjugate momentum of which identically vanishes, must in this case be regarded as defined in terms of the other variables by the accessory Lagrangian equation, generalizing the first NF (2),

$$\kappa^2 \mathbf{V} = -\operatorname{div} \vec{\mathbf{F}} + \mathbf{N} + \frac{e}{\hbar c} \vec{\mathbf{A}} \wedge \vec{\mathbf{F}}. \quad (3')$$

Let us now introduce the notations*

$$\left\{ \begin{array}{l} \hat{\mathbf{V}} = \mathbf{V} + \kappa^{-2} \frac{e}{\hbar c} \vec{\mathbf{A}} \wedge \vec{\mathbf{F}} \\ \hat{\mathbf{G}} = \vec{\mathbf{G}} - \frac{e}{\hbar c} \vec{\mathbf{A}} \wedge \vec{\mathbf{U}} \\ \hat{\mathbf{I}} = \vec{\mathbf{I}} + \frac{e}{\hbar c} \vec{\mathbf{A}} \wedge \Psi, \end{array} \right. \quad (4)$$

the quantities \mathbf{V} , $\vec{\mathbf{G}}$, $\vec{\mathbf{I}}$ being defined by NF (2) and NF (22); denote further by \hat{A} the function A of \mathbf{V} , $\vec{\mathbf{G}}$, $\vec{\mathbf{I}}$, in which these quantities have been replaced by $\hat{\mathbf{V}}$, $\hat{\mathbf{G}}$, $\hat{\mathbf{I}}$, respectively. Since the accessory equation (3') amounts to saying that the variable \mathbf{V} has to be replaced by the quantity $\hat{\mathbf{V}}$ just introduced by the first formula (4), we may with this notation, according to the general results of Part I, write the Hamiltonian of the meson fields, including the interaction with the electromagnetic field, in the form

$$\mathcal{H}_F + \mathcal{H}_\Phi + \int \varrho_{\text{mes}} \cdot B_{\text{ex}} dv;$$

in this formula, \mathcal{H}_F and \mathcal{H}_Φ are defined by NF (7) and NF (26), respectively, while ϱ_{mes} denotes the meson charge density

$$\varrho_{\text{mes}} = \frac{e}{\hbar c} \left\{ -\vec{\mathbf{F}} \wedge \vec{\mathbf{U}} + \Phi \wedge \Psi \right\}_3 \quad (5)$$

as given by formula (30) of Part I. For the meson current density we get

$$\hat{I}_{\text{mes}} = \frac{e}{\hbar c} \left\{ \hat{\mathbf{G}} \wedge \vec{\mathbf{U}} - \vec{\mathbf{F}} \wedge \hat{\mathbf{V}} + \hat{\mathbf{I}} \wedge \Psi \right\}_3. \quad (6)$$

* The sign \wedge denotes a vector product in ordinary as well as in symbolical space.

This expression thus depends on the vector potential \vec{A} ; for $\vec{A} = 0$, it goes over into

$$\vec{I}_{\text{mes}} = \frac{e}{\hbar c} \{ \vec{G} \wedge \vec{U} - \vec{F} \wedge \vec{V} + \vec{\Gamma} \wedge \vec{\Psi} \}_3. \quad (7)$$

Accordingly, the part of the above Hamiltonian representing the interaction energy between meson fields and electromagnetic field consists of a linear term

$$-\int \vec{I}_{\text{mes}} \cdot \vec{A} dv$$

and a quadratic term

$$+ \frac{1}{2} \left(\frac{e}{\hbar c} \right)^2 \int dv \{ (\vec{A} \wedge \vec{U})^2 + (\vec{A} \wedge \vec{\Psi})^2 + \kappa^{-2} (\vec{A} \wedge \vec{F})^2 \}. \quad (8)$$

The latter we will in the following neglect altogether; we shall, therefore, only be concerned with the expression $\vec{I}_{\text{nucl}} + \vec{I}_{\text{mes}}$, keeping in mind, however, that it does not represent the total current density when an electromagnetic field is present.

Summing up, we thus get for the total Hamiltonian

- 1° the energy of the system of nucleons and mesons without electromagnetic field, denoted, as in NF (56), by $\mathcal{H}_k + \mathcal{H}_F + \mathcal{H}_\Phi$,
- 2° the energy of the electromagnetic field \mathcal{H}_f (formula (13) of Part I),
- 3° the total Coulomb energy of protons and charged mesons,
- 4° the interaction energy approximately given by

$$\left. \begin{aligned} & - \int (\vec{I}_{\text{nucl}} + \vec{I}_{\text{mes}}) (\vec{A}_\perp + \vec{A}_{\text{ex}}) dv \\ & + \int (\varrho_{\text{nucl}} + \varrho_{\text{mes}}) B_{\text{ex}} dv, \end{aligned} \right\} \quad (9)$$

the exact expression being obtained by adding to (9) the quadratic term (8).

Now in NF, Part II, § 5, arguments of principle have been developed for applying to the nucleon and meson variables a canonical transformation separating the static part of the meson fields. In fact, the situation in meson theory has been compared

with that in electrodynamics and it has been pointed out that only the transformed variables provide a suitable starting point for a "correspondence" interpretation of the formalism. It further turns out that the transformation in question also brings about an appreciable simplification in the treatment of the electromagnetic properties of nuclear systems.

We have, thus, now to calculate the expression of the above Hamiltonian function in terms of the transformed variables. For the first part, this has been done in NF; the second part is unaffected by the canonical transformation; as regards the two last parts, we have simply to insert for the quantities $\varrho \equiv \varrho_{\text{nucl}} + \varrho_{\text{mes}}$ and $\vec{I} \equiv \vec{I}_{\text{nucl}} + \vec{I}_{\text{mes}}$ their expressions in terms of the new variables. In the following calculations, all symbols will be taken to denote the new variables and functions of the new variables; the old variables and the functions of those variables will then be distinguished by a $\bar{}$ (*e. g.* \bar{A}). Thus, calling \mathcal{S} the canonical transformation operator, we write

$$\bar{A} = \mathcal{S} A \mathcal{S}^{-1}$$

and, according to NF (57, 20, 36),

$$\left. \begin{aligned} \mathcal{S} &= e^{\frac{i}{\hbar c} \mathcal{H}} \\ \mathcal{H} &= \int dv \{ \vec{F}^\circ \cdot \vec{U} - \vec{U}^\circ \cdot \vec{F} + \Psi^\circ \Phi \}; \end{aligned} \right\} (10)$$

we may therefore write, with the notation introduced in NF (68),

$$\bar{A} = A + \sum_{l=1}^{\infty} \frac{1}{l!} \left\{ \frac{i}{\hbar c} \mathcal{H}, A \right\}^l. \tag{11}$$

It is just this formula we have to apply in order to express the quantities

$$\bar{\varrho} \equiv \bar{\varrho}_{\text{nucl}} + \bar{\varrho}_{\text{mes}}, \quad \bar{I} \equiv \bar{I}_{\text{nucl}} + \bar{I}_{\text{mes}}, \tag{12}$$

occurring in the interaction terms of the Hamiltonian, in terms of the new variables. In the following sections, we shall treat the charge density and the current density separately, since the

former has some peculiar properties not exhibited by the latter; we begin with the somewhat simpler case of the charge density.

§ 2. Expression of the charge density in terms of the transformed variables.

The first term in the expansion of type (11) of $\overleftarrow{\varrho} - \varrho$, viz. $\frac{i}{\hbar c} [\mathcal{H}, \varrho]$, we shall denote by $\varrho_{(1)}$. In the first place, we have*, from (10) and (5),

$$\left. \begin{aligned} \varrho_{\times} &\equiv \frac{i}{\hbar c} [\mathcal{H}, \varrho_{\text{mes}}] \\ &= \frac{e}{\hbar c} \left\{ -\vec{F}^{\circ} \wedge \vec{U} + \vec{U}^{\circ} \wedge \vec{F} - \Psi^{\circ} \wedge \Phi \right\}_{\mathbf{3}}. \end{aligned} \right\} \quad (13)$$

The dependence of \mathcal{H} on the variables pertaining to the different nucleons of the system may be expressed by

$$\mathcal{H} = \sum_i \mathbf{r}^{(i)} \int \mathfrak{F}^{(i)} dv; \quad (14)$$

comparing then the expressions (13) and (10) of ϱ_{\times} and \mathcal{H} , we may also write

$$\varrho_{\times} = -\frac{e}{\hbar c} \sum_i \left\{ \mathbf{r}^{(i)} \wedge \int \mathfrak{F}^{(i)} \right\}_{\mathbf{3}} \equiv \sum_i \varrho_{\times}^{(i)}. \quad (15)$$

Passing now on to ϱ_{nucl} , we get from (2)

$$\frac{i}{\hbar c} [\mathcal{H}, \varrho_{\text{nucl}}] = -\frac{e}{2} \sum_i \delta(\vec{x} - \vec{x}^{(i)}) \cdot \frac{i}{\hbar c} [\mathcal{H}, \tau_{\mathbf{3}}^{(i)}]$$

and from (14)

$$\frac{i}{\hbar c} [\mathcal{H}, \tau_{\mathbf{3}}^{(i)}] = -\frac{2}{\hbar c} \left\{ \mathbf{r}^{(i)} \wedge \int \mathfrak{F}^{(i)} dv \right\}_{\mathbf{3}},$$

whence, on account of (15),

$$\frac{i}{\hbar c} [\mathcal{H}, \varrho_{\text{nucl}}] = -\sum_i \delta(\vec{x} - \vec{x}^{(i)}) \int \varrho_{\times}^{(i)} dv. \quad (16)$$

From (13) and (16), the term $\varrho_{(1)}$ is thus found to be

* We write $[A, B]$ for $[A, B]_{-} = AB - BA$.

$$\left. \begin{aligned} \varrho_{(1)} &= \varrho_{\times} - \sum_i \delta(\vec{x} - \vec{x}^{(i)}) \int \varrho_{\times}^{(i)} dv \\ &= \sum_i \left\{ \varrho_{\times}^{(i)} - \delta(\vec{x} - \vec{x}^{(i)}) \int \varrho_{\times}^{(i)}(\vec{x}') dv' \right\}, \end{aligned} \right\} \quad (17)$$

giving by integration

$$\int \varrho_{(1)} dv = 0,$$

in agreement with the invariance property of the total charge of the system. The expression of $\varrho_{(1)}$ may be put into a more symmetrical shape by introducing the function of two sets of space variables

$$\varrho_{\times}(\vec{x}, \vec{x}') \equiv \sum_i \varrho_{\times}^{(i)}(\vec{x}) \delta(\vec{x}' - \vec{x}^{(i)}); \quad (18)$$

we then get

$$\varrho_{(1)}(\vec{x}) = \int dv' [\varrho_{\times}(\vec{x}, \vec{x}') - \varrho_{\times}(\vec{x}', \vec{x})]. \quad (19)$$

Using (15) and the expressions NF (14) and NF (30) of the static meson fields, we may write for $\varrho_{\times}(\vec{x}, \vec{x}')$

$$\left. \begin{aligned} &\varrho_{\times}(\vec{x}, \vec{x}') \\ = -\frac{e}{\hbar c} \left\{ \mathbf{N}(\vec{x}') \wedge \vec{\mathbf{U}}(\vec{x}) - \mathbf{S}(\vec{x}') \wedge \vec{\mathbf{F}}(\vec{x}) + \vec{\mathbf{P}}(\vec{x}') \wedge \Phi(\vec{x}) \right\}_{\mathbf{3}} \\ &\cdot \text{grad}' \varphi(|\vec{x}' - \vec{x}|). \end{aligned} \right\} \quad (20)$$

We see that $\varrho_{(1)}$ is linear in the free meson variables, so that it does not give any contribution to the charge density of a nuclear system in the absence of free mesons; it only gives rise to a peculiar interaction between nucleons and free mesons when an external electrostatic field (*e. g.* the field of atomic electrons) is present.

The next contribution to the expansion of $\vec{\varrho} - \varrho$ is $\frac{i}{2\hbar c} [\mathcal{H}, \varrho_{(1)}]$. It consists of two terms of which the one,

$$\varrho_{(2)} = \sum_i \left\{ \varrho_{\times \times}^{(i)} - \delta(\vec{x} - \vec{x}^{(i)}) \int \varrho_{\times \times}^{(i)}(\vec{x}') dv' \right\}, \quad (21)$$

is obtained by disregarding the non-commutability of the meson field variables and is thus quadratic in these variables, while the other*,

$$\left. \begin{aligned} \varrho_{\text{exch}} &= \frac{e}{2\hbar c} \left\{ \vec{U}^\circ \wedge \vec{F}^\circ - \vec{F}^\circ \wedge \vec{U}^\circ \right\}_{\mathbf{3}} = \frac{e}{\hbar c} \left\{ \overline{\vec{U}^\circ \wedge \vec{F}^\circ} \right\}_{\mathbf{3}} \\ &= -\frac{e}{2\hbar c} \operatorname{div} \left\{ \vec{U}^\circ \wedge \vec{V}^\circ - \vec{V}^\circ \wedge \vec{U}^\circ \right\}_{\mathbf{3}} = -\frac{e}{\hbar c} \operatorname{div} \left\{ \overline{\vec{U}^\circ \wedge \vec{V}^\circ} \right\}_{\mathbf{3}}, \end{aligned} \right\} (22)$$

is independent of the meson field variables. The value (22) of ϱ_{exch} is readily obtained from (13); in fact, the field-independent part of $\frac{i}{2\hbar c} [\mathcal{H}, \varrho_{\times}]$ is just the expression (22) and, since it may be written as a sum of terms relative to the single nucleons, and of which the space integrals vanish, it also represents, according to (15) and (17), the field-independent part of $\frac{i}{2\hbar c} [\mathcal{H}, \varrho_{(1)}]$. While, as indicated in formula (21), the quadratic term $\varrho_{(2)}$ is obviously a sum of contributions from the single nucleons, the field-independent term is, according to (22), of the form $\sum_{i \neq k} \left\{ \mathbf{T}^{(i)} \wedge \mathbf{T}^{(k)} \right\}_{\mathbf{3}} \Phi^{(i,k)}$ and does, therefore, not contain any such contribution, but presents an "exchange" character with regard to the proton and neutron states of the nucleons; we shall call it the "exchange charge density". For a pair of nucleons, exchange operators of the type just considered, $\left\{ \mathbf{T}^{(1)} \wedge \mathbf{T}^{(2)} \right\}_{\mathbf{3}} \Phi^{(1,2)}$, have non-vanishing matrix elements only if the system is a deuteron, and these matrix elements refer to states in which the two particles have exchanged their proton or neutron character; in particular, the mean value of such an operator in any stationary state of the system is zero.

For the $\varrho_{\times \times}^{(i)}$ entering into the expression (21) of the quadratic term $\varrho_{(2)}$, we may write, according to (14) and (15),

$$\varrho_{\times \times}^{(i)}(\vec{x}) = -\frac{e}{\hbar c} \cdot \frac{i}{2\hbar c} \int dv' \left[\mathbf{T}^{(i)} \mathfrak{F}^{(i)}(\vec{x}'); \left\{ \mathbf{T}^{(i)} \wedge \mathfrak{F}^{(i)}(\vec{x}) \right\}_{\mathbf{3}} \right]_-.$$

The commutator in this formula has the following value (the index (i) having been dropped):

* We denote by \overline{AB} the symmetrical combination $\frac{1}{2}(AB + BA)$.

$$[\mathcal{J}'_1, \mathcal{J}_2]_- - [\mathcal{J}'_2, \mathcal{J}_1]_- + i\tau_1 [\mathcal{J}'_3, \mathcal{J}_1]_+ + i\tau_2 [\mathcal{J}'_3, \mathcal{J}_2]_+ - i\tau_3 \{ [\mathcal{J}'_1, \mathcal{J}_1]_+ + [\mathcal{J}'_2, \mathcal{J}_2]_+ \};$$

putting

$$\left\{ \begin{aligned} \mathfrak{F}^{(i)} &\equiv \mathbf{a}^{(i)} + \overset{\rightarrow}{\sigma}^{(i)} \overset{\rightarrow}{\mathbf{b}}^{(i)} \\ \mathbf{a}^{(i)} &\equiv g_1 \vec{f}^{(i)} \vec{U} \\ \overset{\rightarrow}{\mathbf{b}}^{(i)} &\equiv \frac{g_2}{\kappa} \varrho_3^{(i)} \vec{f}^{(i)} \wedge \vec{F} + \frac{f_2}{\kappa} \vec{f}^{(i)} \Phi \\ \vec{f}^{(i)} &\equiv \text{grad}^{(i)} \varphi (|\vec{x} - \vec{x}^{(i)}|), \end{aligned} \right. \quad (23)$$

we get, $\mathbf{a}^{(i)}$ and $\overset{\rightarrow}{\mathbf{b}}^{(i)}$ being commutable,

$$\left\{ \begin{aligned} [\mathcal{J}'_i, \mathcal{J}_j]_- &= 2i\sigma \vec{b}'_i \wedge \vec{b}_j \\ [\mathcal{J}'_i, \mathcal{J}_j]_+ &= 2[a'_i a_j + \vec{b}'_i \vec{b}_j + \sigma(a'_i \vec{b}_j + \vec{b}'_i a_j)], \end{aligned} \right.$$

so that finally

$$\varrho_{\times \times}^{(i)} = \frac{e}{(\hbar c)^2} \int dv' \left\{ \overset{\rightarrow}{\sigma}^{(i)} (\overset{\rightarrow}{\mathbf{b}}^{(i)'} \wedge \overset{\rightarrow}{\mathbf{b}}^{(i)})_{\mathbf{3}} + \tau^{(i)} [a_{\mathbf{3}}^{(i)'} \mathbf{a}^{(i)} + \vec{b}_{\mathbf{3}}^{(i)'} \overset{\rightarrow}{\mathbf{b}}^{(i)} + \overset{\rightarrow}{\sigma}^{(i)} (a_{\mathbf{3}}^{(i)'} \overset{\rightarrow}{\mathbf{b}}^{(i)} + \vec{b}_{\mathbf{3}}^{(i)'} \mathbf{a}^{(i)})] - \tau_{\mathbf{3}}^{(i)} [\mathbf{a}^{(i)'} \mathbf{a}^{(i)} + \overset{\rightarrow}{\mathbf{b}}^{(i)'} \overset{\rightarrow}{\mathbf{b}}^{(i)} + \overset{\rightarrow}{\sigma}^{(i)} (\mathbf{a}^{(i)'} \overset{\rightarrow}{\mathbf{b}}^{(i)} + \overset{\rightarrow}{\mathbf{b}}^{(i)'} \mathbf{a}^{(i)})] \right\}. \quad (24)$$

Owing to the existence of the fluctuating zero-point meson field, this expression gives a contribution to the charge density operator of a nuclear system in the absence of free mesons, namely the mean value of $\varrho_{(2)}$ with respect to the meson variables for the state corresponding to the zero-point meson field. Denoting such a mean value by $\mathfrak{M}_0 \{ \dots \}$ and taking account of the formulae (A7) and (A8) of the Appendix, we directly get from (24)

$$\begin{aligned} \mathfrak{M}_0 \{ \varrho_{\times \times}^{(i)} \} &= -\frac{2e}{(\hbar c)^2} \tau_{\mathbf{3}}^{(i)} \int dv' \mathfrak{M}_0 \{ a_{\mathbf{1}}^{(i)'} a_{\mathbf{1}}^{(i)} + \vec{b}_{\mathbf{1}}^{(i)'} \vec{b}_{\mathbf{1}}^{(i)} \} \\ &= -\frac{2e}{(\hbar c)^2} \tau_{\mathbf{3}}^{(i)} \int dv' \left[g_1^2 \mathfrak{M}_0 \{ (\vec{f}^{(i)} \vec{U}_{\mathbf{1}})_{\vec{x}} (\vec{f}^{(i)} \vec{U}_{\mathbf{1}})_{\vec{x}} \} \right. \\ &\quad + \left(\frac{g_2}{\kappa} \right)^2 \mathfrak{M}_0 \{ (\vec{f}^{(i)} \wedge \vec{F}_{\mathbf{1}})_{\vec{x}} (\vec{f}^{(i)} \wedge \vec{F}_{\mathbf{1}})_{\vec{x}} \} \\ &\quad \left. + \left(\frac{f_2}{\kappa} \right)^2 (\vec{f}^{(i)} \vec{f}^{(i)'}) \mathfrak{M}_0 \{ \Phi_{\mathbf{1}} \Phi'_{\mathbf{1}} \} \right], \end{aligned}$$

according to (23). By means of the formulae (A 12), (A 13), and (A 14) of the Appendix, this takes the form

$$= -e \tau_{\mathbf{3}}^{(i)} \cdot \frac{g_1^2 + 2g_2^2 + f_2^2}{4\pi \hbar c} \cdot \frac{e^{-\kappa r^{(i)}}}{r^{(i)}} \left(\kappa + \frac{1}{r^{(i)}} \right) \frac{1}{2\pi^2 r^{(i)}} K_2(\kappa r^{(i)}), \quad \left. \vphantom{\frac{e^{-\kappa r^{(i)}}}{r^{(i)}}} \right\} \mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\} \quad (25)$$

where $r^{(i)} = |\vec{x} - \vec{x}^{(i)}|$, and K_2 is the Bessel function defined in the Appendix (of course, we have in our theory the relation $f_2^2 = g_2^2$, but we have not made use of it in writing this formula in order to distinguish the contribution from pseudoscalar mesons). Owing to the singularity of $K_2(z)$ for $z = 0$, however, the integral $\int dv \mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\}$ diverges, so that $\mathfrak{M}_0 \left\{ \varrho_{(2)j} \right\}$ is not well-defined. According to the general criterium proposed in NF, Part II, § 5, no significance can therefore be attributed to any effect depending on this term or similar effects arising from a perturbation calculation, and it is *a fortiori* useless to consider further terms in the expansion of $\bar{\varrho} - \varrho$.

In particular, one might think that an additional interaction energy between a nucleon and an electron would be obtained by taking for B_{ex} in $\int dv B_{\text{ex}} \mathfrak{M}_0 \left\{ \varrho_{(2)j} \right\}$ the Coulomb potential

$$B_{\text{ex}} = - \frac{e}{4\pi |\vec{x} - \vec{x}^{(0)}|}$$

of an electron at $\vec{x}^{(0)}$, since it is possible in this case to carry out the integration in such a way as to get a finite result. In fact, taking the position of the nucleon as the origin of polar coordinates and observing that $\mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\}$ is a spherically symmetrical charge distribution, we may write, according to potential theory,

$$\begin{aligned} & \int \frac{1}{4\pi |\vec{x} - \vec{x}^{(0)}|} \mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\} r^2 dr d\Omega \\ &= \frac{1}{R} \int_0^R \mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\} r^2 dr + \int_{Rr}^{\infty} \frac{1}{Rr} \mathfrak{M}_0 \left\{ \varrho_{\times \times}^{(i)} \right\} r^2 dr, \end{aligned}$$

where $R \equiv |\vec{x}^{(i)} - \vec{x}^{(0)}|$; therefore

$$\int B_{\text{ex}} \mathfrak{M}_0 \{q_{(2)}\} dv = -e \int_R^{\infty} \left(\frac{1}{r} - \frac{1}{R} \right) \mathfrak{M}_0 \{q_{\times \times}^{(i)}\} r^2 dr.$$

To this should then be added a term of the same order of magnitude, derived as a second order perturbation from the other couplings between nucleons and mesons; we need not consider it here, however, but refer the reader to the forthcoming paper by J. SERPE mentioned in the Introduction. The total effect has been first pointed out by H. FRÖHLICH, W. HEITLER and B. KAHN [16], but its reality was contested by W. LAMB [17]. From our point of view, it should be clear that the whole effect must be discarded; a more detailed discussion will be found in SERPE's paper.

The conclusions of this section may thus be summarized in the formula

$$\bar{q} = q_{\text{nucl}} + q_{\text{mes}} + q_{(1)} + q_{(2)} + q_{\text{exch}}, \quad (26)$$

the various terms being defined by (2), (5), (17) [with (13), (15)], or (19) [with (20)], (21) [with (24), (23)], and (22), respectively.

§ 3. Electric dipole and quadrupole moments.

With the help of (26), the expressions of the electric dipole and quadrupole moments, defined by formulae (32) of Part I, may be written down immediately in the same form. They comprise 1° terms independent of the meson fields (which we often call, for the sake of brevity, "field-independent terms"), *viz.* a term due to the elementary charges of the protons (q_{nucl}) and an exchange term arising from the coupling through the meson field (as embodied in q_{exch}); 2° a coupling term between nucleons and free mesons, linear in the meson field components (from $q_{(1)}$); 3° terms quadratic in the meson variables, *viz.* the contribution from the free mesons (q_{mes}) and a quadratic coupling term (from $q_{(2)}$).

It will be noticed that the spreading of the charge of a nucleon, which would be due to the interaction with the zero-point meson field, a part of which would be given by $\mathfrak{M}_0 \{q_{(2)}\}$, would possess spherical symmetry and would thus not give rise to any multipole moments (as usually defined; cf. footnote on p. 21);

in fact, it would act only at large distances from the nucleon and the analysis of Part I, § 4 would not be applicable to it*.

The most notable terms characteristic of meson theory are the exchange terms which, using (22), NF (14, 3, 4, 89), and the relation

$$\int \varphi(|\vec{x}^{(i)} - \vec{x}|) \varphi(|\vec{x}^{(k)} - \vec{x}|) \vec{x} dv = \frac{e^{-\kappa|\vec{x}^{(i)} - \vec{x}^{(k)}|}}{8\pi\kappa} \cdot \frac{\vec{x}^{(i)} + \vec{x}^{(k)}}{2},$$

may be written

$$\vec{P}_{\text{exch}} = \frac{e}{\hbar c} \int \left\{ \overline{\vec{U}^\circ \mathbf{A} \mathbf{V}^\circ} \right\}_{\mathbf{3}} dv = \frac{1}{2} \cdot \frac{e}{\hbar c} \cdot \frac{g_1 g_2}{\kappa} \sum_{i \neq k} \left\{ \mathbf{T}^{(i)} \mathbf{A} \mathbf{T}^{(k)} \right\}_{\mathbf{3}} \left[\vec{\sigma}^{(i)} \wedge (\vec{x}^{(i)} - \vec{x}^{(k)}) \right] \varphi(|\vec{x}^{(i)} - \vec{x}^{(k)}|) \quad (27)$$

$$\begin{aligned} Q_{\text{exch}}^{lm} &= \frac{1}{2} \frac{e}{\hbar c} \int \left[\left\{ \overline{\vec{U}^\circ \mathbf{A} \mathbf{V}^\circ} \right\}_{\mathbf{3}}^l x^m + \left\{ \overline{\vec{U}^\circ \mathbf{A} \mathbf{V}^\circ} \right\}_{\mathbf{3}}^m x^l \right] dv \\ &= \frac{1}{8} \cdot \frac{e}{\hbar c} \cdot \frac{g_1 g_2}{\kappa} \sum_{i \neq k} \left\{ \mathbf{T}^{(i)} \mathbf{A} \mathbf{T}^{(k)} \right\}_{\mathbf{3}} \cdot \left[(\vec{x}^{(i)} + \vec{x}^{(k)})^l (\vec{\sigma}^{(i)} \wedge (\vec{x}^{(i)} - \vec{x}^{(k)}))^m \right. \\ &\quad \left. + (\vec{x}^{(i)} + \vec{x}^{(k)})^m (\vec{\sigma}^{(i)} \wedge (\vec{x}^{(i)} - \vec{x}^{(k)}))^l \right] \cdot \varphi(|\vec{x}^{(i)} - \vec{x}^{(k)}|). \end{aligned} \quad (28)$$

Owing to their exchange character, these terms do not give any contribution to the energy of stationary states of nuclear systems, but they may play a role in the calculation of the transition probabilities between such states under the influence of an external electric field. Since the mean value of the charge density ρ_{nucl} in any stationary state is invariant with respect to reflections about the centre of the nucleus, there is no electric dipole moment at all in these states. The interaction energy with an external electric field is thus determined, apart from the total charge, by the quadrupole moment Q_{nucl} . A most important example of such an interaction is met with in the theory of hyperfine structure, the external field in that case being the field of the atomic electrons. It may be shown** that the

* The effect of a (finite) spreading of the nuclear charge on the atomic levels (owing to the electrostatic interaction with the atomic electrons) is discussed by H. CASIMIR, loc. cit. [18], § 5.

** See H. CASIMIR, loc. cit. [18], § 4. The expression (d) which forms the starting point of CASIMIR's investigation is identical with our interaction term $-(Q \text{ grad}_0) \vec{\mathcal{E}}_0$, where $\vec{\mathcal{E}}_0$ is the field due to the atomic electrons at the centre of the nucleus.

quadrupole charge distribution of the nucleus enters into the expression of the interaction energy only through one parameter, called the "quadrupole moment" of the nucleus (in the spectroscopic sense), and conveniently defined as 6 times the mean value of our $Q^{33} - \frac{1}{3} \sum_i Q^{ii}$ for the state with the maximum value of the magnetic quantum number*.

For the calculation of transition probabilities between states of the same total energy, it is equivalent, as noticed at the end of § 4 of Part I, as operator of electric dipole interaction to use $-\vec{A}_0 \vec{P}$ instead of $-\vec{E}_0 \vec{P}$. At an early stage of the development of nuclear theory, it was pointed out by SIEGERT [19] that the choice of the former operator might be advantageous because of a remarkable connexion between the operator** \vec{P} (or $\int \vec{I} dv$) and the exchange potential of the nuclear forces; the exchange part of \vec{P} could then directly be written down when this potential was given. We shall now derive "SIEGERT'S theorem" from the more general point of view of our theory which also involves "velocity-dependent" couplings between the nucleons (*i. e.* terms depending on the velocities of the nucleons). For this purpose (since SIEGERT'S theorem is only concerned with processes involving no free mesons), we have only to consider the field-independent part of \vec{P} ; neglecting accordingly all terms in \vec{P} which are quadratic in the meson field variables, we have only to calculate (cf. (26)) the field-independent part of $\vec{P}_{\text{nucl}} + \vec{P}_{(1)} + \vec{P}_{\text{exch}}$. We begin with \vec{P}_{nucl} ; from

$$\vec{P}_{\text{nucl}} = e \sum_i \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \vec{x}^{(i)}$$

we get

* This quantity has been calculated for the ground state of the deuteron in NF, Part III, § 3 (no restriction being made on the nuclear source constants; cf. [4]). It should be observed, however, that, since relative coordinates were used, the quantity Q there computed (formula (123)) should be divided by 4 to yield the quadrupole moment in the spectroscopic sense.

** In conformity with our previous notations, \vec{I} represents the total current density of the system, including the dependence on the vector-potential.

$$\dot{\vec{P}}_{\text{nucl}} = e \sum_i \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \vec{a}^{(i)} - \frac{e}{2} \sum_i \dot{\tau}_{\mathbf{3}}^{(i)} \vec{x}^{(i)},$$

the field-independent part of which to a sufficient approximation reduces to

$$\mathfrak{M}_0 \left\{ \dot{\vec{P}}_{\text{nucl}} \right\} = \int \vec{I}_{\text{nucl}} dv - \frac{e}{\hbar c} \sum_i \frac{i}{2} [\mathcal{Q}'_n + \mathcal{Q}''_n, \tau_{\mathbf{3}}^{(i)}] \vec{x}^{(i)}. \quad (29)$$

In this formula, \mathcal{Q}'_n is the static potential of the nuclear forces given by NF (65), and \mathcal{Q}''_n the velocity-dependent potential given by NF (84)* or NF (85):

$$\left. \begin{aligned} \mathcal{Q}'_n &= \frac{1}{2} \int (\mathbf{N}' \mathbf{N} + \kappa^2 \mathbf{S}' \mathbf{S}) \varphi dv dv' \\ \mathcal{Q}''_n &= \int (\mathbf{N}' \vec{T} + \vec{S}' \wedge \vec{M} + \vec{P}' \mathbf{R}) \text{grad } \varphi dv dv'. \end{aligned} \right\} \quad (30)$$

All their terms are of the form $\int \mathbf{A}(\vec{x}') \mathbf{B}(\vec{x}) dv dv'$ with

$$\mathbf{A}(\vec{x}) = \sum_i \mathbf{r}^{(i)} a^{(i)} \delta(\vec{x} - \vec{x}^{(i)})$$

and an analogous expression for \mathbf{B} . We have

$$\sum_i \frac{i}{2} [\mathbf{A}(\vec{x}') \mathbf{B}(\vec{x}), \tau_{\mathbf{3}}^{(i)}] \vec{x}^{(i)} = \left\{ \mathbf{A}(\vec{x}') \wedge \mathbf{B}(\vec{x}) \right\}_{\mathbf{3}} (\vec{x} - \vec{x}')$$

and, therefore,

$$\left. \begin{aligned} \vec{J} &\equiv - \frac{e}{\hbar c} \sum_i \frac{i}{2} [\mathcal{Q}'_n, \tau_{\mathbf{3}}^{(i)}] \vec{x}^{(i)} \\ &= - \frac{e}{2\hbar c} \int \left\{ \mathbf{N}' \wedge \mathbf{N} + \kappa^2 \vec{S}' \wedge \vec{S} \right\}_{\mathbf{3}} (\vec{x} - \vec{x}') \varphi dv dv', \end{aligned} \right\} \quad (31)$$

$$\left. \begin{aligned} \vec{J}' &\equiv - \frac{e}{\hbar c} \sum_i \frac{i}{2} [\mathcal{Q}''_n, \tau_{\mathbf{3}}^{(i)}] \vec{x}^{(i)} \\ &= - \frac{e}{\hbar c} \int \left(\left\{ \mathbf{N}' \wedge \vec{T} + \vec{S}' \wedge \vec{M} + \vec{P}' \wedge \mathbf{R} \right\}_{\mathbf{3}} \cdot \text{grad } \varphi \right) (\vec{x} - \vec{x}') dv dv', \end{aligned} \right\} \quad (32)$$

$$\mathfrak{M}_0 \left\{ \dot{\vec{P}}_{\text{nucl}} \right\} = \int \vec{I}_{\text{nucl}} dv + \vec{J} + \vec{J}'. \quad (33)$$

Looking apart from the relatively small term \vec{J}' , formula (33) already embodies the primitive form of SIEGERT's theorem;

* The second formula (30) is equivalent with NF (84), though slightly modified in form.

the comparison of (31) and (30) shows how the exchange part of $\mathfrak{M}_0 \{ \dot{P}_{\text{nuc1}} \}$ (or of $\mathfrak{M}_0 \{ \int \vec{I} dv \}$) may be built up from the static potential \mathcal{Q}_n . We may, however, obtain a more rigorous formulation by taking account of the other terms of $\mathfrak{M}_0 \{ \dot{P} \}$, namely $\mathfrak{M}_0 \{ \dot{P}_{(1)} + \dot{P}_{\text{exch}} \}$. We have from (19)

$$\vec{P}_{(1)} = \int dv dv' \varrho_{\times} (\vec{x}, \vec{x}') (\vec{x} - \vec{x}'),$$

whence, by (20),

$$= -\frac{e}{\hbar c} \int dv dv' \mathfrak{M}_0 \left\{ \left(\mathbf{N}' \wedge \vec{U} - \vec{S}' \wedge \vec{F} + \vec{P}' \wedge \vec{\Phi} \right)_{\mathbf{3}} \text{grad}' \varphi \right\} (\vec{x} - \vec{x}') \left. \vphantom{\int} \right\} \quad (34)$$

The contribution from \dot{P}_{exch} may be put into a quite analogous form by writing, from (22) and NF (9, 14),

$$\begin{aligned} \vec{P}_{\text{exch}} &= \frac{e}{\hbar c} \int dv \left\{ \overline{\vec{U} \wedge \vec{F}^{\circ}} \right\}_{\mathbf{3}} \vec{x} \\ &= \frac{e}{2\hbar c} \int dv dv' \left(\left\{ \vec{S}' \wedge \vec{F}^{\circ} - \mathbf{N}' \wedge \vec{U}^{\circ} \right\}_{\mathbf{3}} \cdot \text{grad}' \varphi \right) (\vec{x} - \vec{x}'), \end{aligned}$$

the replacement of the factor \vec{x} by $(\vec{x} - \vec{x}')$ in the last formula being allowed on account of $\text{rot } \vec{F}^{\circ} = 0, \text{div } \vec{U}^{\circ} = 0$. Thence,

$$\mathfrak{M}_0 \{ \dot{P}_{\text{exch}} \} = -\frac{e}{\hbar c} \int dv dv' \left(\left\{ \mathbf{N}' \wedge \vec{U}^{\circ} - \vec{S}' \wedge \vec{F}^{\circ} \right\}_{\mathbf{3}} \cdot \text{grad}' \varphi \right) (\vec{x} - \vec{x}'), \quad (35)$$

with the notation (slightly extending NF (79))

$$\overset{\circ}{A} = \frac{i}{\hbar c} [\mathcal{H}_n + \mathcal{L}_n, A].$$

Now, the meson field equations* yield to our approximation,

$$\left. \begin{aligned} \mathfrak{M}_0 \{ \dot{\vec{U}} \} + \overset{\circ}{\vec{U}} &= \vec{T} \\ \mathfrak{M}_0 \{ \dot{\vec{F}} \} + \overset{\circ}{\vec{F}} &= -\vec{M} \\ \mathfrak{M}_0 \{ \dot{\vec{\Phi}} \} &= \vec{R}. \end{aligned} \right\} \quad (36)$$

* Cf. the equations NF (87) which must be completed, however, by the contributions from the term \mathcal{L} in NF (86), \mathcal{L} being given by NF (67).

Inserting the values (36) in the sum of (34) and (35), and comparing with (32), we get just

$$\mathfrak{M}_0 \left\{ \vec{P}_{(1)} + \vec{P}_{\text{exch}} \right\} = -\vec{J}. \quad (37)$$

Addition of (33) and (37) then gives the generalized form of SIEGERT'S theorem

$$\mathfrak{M}_0 \left\{ \vec{P} \right\} = \int \vec{I}_{\text{nucl}} dv + \vec{J}. \quad (38)$$

It is remarkable that the exchange part of $\mathfrak{M}_0 \left\{ \vec{P} \right\}$ is completely determined by the static potential, even when due account is taken of the effects of the first order in the velocities of the nucleons. From the primitive form (33) of the theorem, with the term \vec{J} omitted, it had been concluded [20] that it was justified to take as electric dipole interaction the operator $-\vec{E}_0 \vec{P}_{\text{nucl}}$ without any exchange moment. Although this argumentation is not rigorous, we see that the conclusion is nevertheless correct even to the first order of approximation with respect to the nuclear velocities. From the preceding analysis it appears, more precisely, that (since the relevant matrix elements of $-\vec{E}_0 \vec{P}_{(1)}$ or $-\vec{A}_0 \vec{P}_{(1)}$ vanish) the contribution from $-\vec{E}_0 \vec{P}_{\text{exch}}$ can only be of a higher order in the velocities of the nucleons*—at first sight a somewhat surprising result, since the operator itself is velocity-independent—and that this contribution just cancels the term \vec{J} of the same order of magnitude from $-\vec{E}_0 \vec{P}_{\text{nucl}}$. Summarizing, we may say that the transition probabilities between states of the same energy, insofar as electric dipole interaction is concerned, may, to the first order in nuclear velocities, indifferently be computed from the operator $-\vec{E}_0 \vec{P}_{\text{nucl}}$ or from the operator $-\vec{A}_0 \mathfrak{M}_0 \left\{ \vec{P} \right\}$, given by (38) and (31). In the next section, we shall verify by a direct calculation that, as results from the way it has been introduced, \vec{J} is just the "exchange" part of the integral current operator.

* In connexion with the problem of the photo-disintegration of the deuteron, treated from the point of view of the present paper, this point has been discussed in detail by A. PAIS, loc. cit. [7], Appendix.

We may finally remark that SIEGERT's theorem also holds in any meson theory using only one kind of meson field, such as a pure vector or a pure pseudoscalar theory. In these cases, the expression of \vec{J} would involve a term arising from the dipole interaction; in the mixed theory, such terms automatically cancel each other, yielding a simpler form for \vec{J} (cf. the Appendix). It may therefore be said that the simplification of the electrical quantities in the mixed theory has its origin in SIEGERT's theorem on account of the simpler determination of the static potential.

§ 4. The current density.

The expression of the current density in terms of the new variables may again be obtained as an expansion of the type (11). We shall not here carry out a complete calculation, but confine ourselves to the terms which are of practical interest; thus, we shall, in this expansion, neglect all terms which, besides involving some power of the source constants g and f also contain a factor of the order of magnitude of the ratio of the nucleon velocities to the velocity of light. This implies, in the first place, that in all terms of the form

$$\left\{ \frac{i}{\hbar c} \mathcal{K}, \vec{I}_{\text{nucl}} + \vec{I}_{\text{mes}} \right\}^l$$

we shall neglect the contributions arising from \vec{I}_{nucl} . We may then write to a sufficient approximation

$$\left. \begin{aligned} \vec{I} &= \vec{I}_{\text{nucl}} + \vec{I}_{\text{mes}} + \vec{I}_{\times} + \frac{i}{2\hbar c} \left[\mathcal{K}, \vec{I}_{\times} \right], \\ \text{with } \vec{I}_{\times} &= \frac{i}{\hbar c} \left[\mathcal{K}, \vec{I}_{\text{mes}} \right]. \end{aligned} \right\} \quad (39)$$

As regards the dependence on the meson variables, the situation is here slightly complicated by the fact that \vec{I}_{mes} itself, apart from the current density for a system of entirely free mesons,

$$\vec{I}_{\text{free mes}} = \frac{e}{\hbar c} \left\{ \text{rot } \vec{U} \wedge \vec{U} + \kappa^{-2} \vec{F} \wedge \text{div } \vec{F} - \text{grad } \Psi \wedge \Psi \right\}_{\mathbf{3}}, \quad (40)$$

also contains a term, linear in the meson variables, which depends on the nuclear source densities:

$$\vec{I}_{\text{mes}} = \vec{I}_{\text{free mes}} + \frac{e}{\hbar c} \left\{ \vec{S} \wedge \vec{U} - \kappa^{-2} \vec{F} \wedge \mathbf{N} + \vec{P} \wedge \Psi \right\}_{\mathbf{3}}. \quad (41)$$

Therefore, \vec{I}_{\times} as well as $\frac{i}{2\hbar c} [\mathcal{H}, \vec{I}_{\times}]$ will contain terms of degrees 0,1 and 2 in the meson variables; for those of degree zero* we readily get, from (40) and (41),

$$\vec{I}_{(0)} = \frac{e}{\hbar c} \left\{ \overline{\vec{G}^{\circ} \wedge \vec{U}^{\circ}} - \overline{\vec{F}^{\circ} \wedge \mathbf{V}^{\circ}} + \overline{\vec{I}^{\circ} \wedge \Psi^{\circ}} \right\}_{\mathbf{3}}. \quad (42)$$

We may thus write, instead of (39),

$$\vec{I} = \vec{I}_{\text{nucl}} + \vec{I}_{\text{free mes}} + \vec{I}_{(1)} + \vec{I}_{(2)} + \vec{I}_{(0)}, \quad (43)$$

$\vec{I}_{(1)}$ and $\vec{I}_{(2)}$ denoting expressions linear and quadratic, respectively, in the meson variables, which need not be written down explicitly.

The field-independent term $\vec{I}_{(0)}$ is not entirely analogous to the term ϱ_{exch} of formula (26), for the expression on the right in formula (42) gives rise not only to an "exchange" current density \vec{I}_{exch} (involving a sum over all pairs of nucleons), but also to a "proper" current density, \vec{I}_{prop} , consisting of a sum of terms pertaining to the single nucleons. Since we have assumed for the nucleons a theory corresponding to the idealization of material points, the proper current density \vec{I}_{prop} will contain divergent contributions. Such divergences, however, have a quite different origin from those arising from the field-independent part $\mathfrak{M}_0 \left\{ \vec{I}_{(2)} \right\}$ of $\vec{I}_{(2)}$ on account of the existence of the fluctuating zero-field. While, according to our general criterium, we must entirely discard the latter, we may therefore treat the former quite independently and, for instance, prevent their occurrence

* It should be noticed that the separation of the terms of degrees 0 and 2 is not unambiguous, since a commutation of the variables \vec{F}° and \vec{U}° in the latter would, on account of the commutation rules, give rise to a contribution of order zero. An unambiguous definition is first obtained if one adds the condition that the terms in question be Hermitian. Such an Hermitization has been performed in formula (42), as indicated by the symmetrization bars.

by replacing the point distribution of the nuclear sources by some extended distribution. To this question we shall come back in the next section, in connexion with the problem of the proper magnetic moment of nucleons. Here, we shall only point out that the expression of the integral field-independent current $\int \vec{I}_{(0)} dv$ turns out to be identical with the quantity \vec{J} defined by (31); the details of the calculation are reported in the Appendix. Owing to the factor $(\vec{x} - \vec{x}')$ in \vec{J} , no "proper" contribution obviously arises in this case, so that we actually have

$$\int \vec{I}_{(0)} dv = \int \vec{I}_{\text{exch}} dv = \vec{J}. \quad (44)$$

§ 5. Magnetic dipole moment.

The last point we have to discuss is the transformation of the magnetic dipole moment of a system of mesons and nucleons to the new variables. For the sake of completeness, we begin by briefly recalling the well-known situation as regards Dirac particles and free mesons. Consider a Dirac particle of charge e and mass M ; let

$$\vec{v}^{(0)} = \frac{1}{Mc} \left(\vec{p}^{(0)} - e \vec{A}(\vec{x}^{(0)}) \right),$$

$\vec{x}^{(0)}$, $\vec{p}^{(0)}$ being the canonical position and momentum coordinates of the particle (the latter multiplied by c). The evaluation of

$$\frac{d}{dt} \left\{ \varrho_2 \vec{\sigma} \delta(\vec{x} - \vec{x}^{(0)}) \right\} = \frac{1}{i\hbar c} \left[\varrho_2 \vec{\sigma} \delta(\vec{x} - \vec{x}^{(0)}), \mathcal{H}_k \right]$$

with the help of the Hamiltonian

$$\mathcal{H}_k = Mc \left(\alpha \vec{v}^{(0)} + \varrho_3 c \right)$$

leads, after some easy reductions, to the formula

$$\begin{aligned} \vec{I} \equiv e \alpha \delta(\vec{x} - \vec{x}^{(0)}) &= \frac{e}{c} \varrho_3 \cdot \frac{1}{2} \left(\vec{v}^{(0)} \delta(\vec{x} - \vec{x}^{(0)}) + \delta(\vec{x} - \vec{x}^{(0)}) \vec{v}^{(0)} \right) \\ &+ \mu_0 \text{rot} \left\{ \varrho_3 \vec{\sigma} \delta(\vec{x} - \vec{x}^{(0)}) \right\} + \mu_0 \frac{d}{dt} \left\{ \varrho_2 \vec{\sigma} \delta(\vec{x} - \vec{x}^{(0)}) \right\} \end{aligned}$$

for the current density operator; $\mu_0 = \frac{e\hbar}{2Mc}$ denotes the "Bohr magneton" relative to the mass-value M . Hence, the magnetic moment $\vec{M} = \frac{1}{2} \int \vec{x} \wedge \vec{I} dv$ takes the form*

$$\vec{M} = \frac{e}{2c} \varrho_3 \vec{x}^{(0)} \wedge \vec{v}^{(0)} + \mu_0 \varrho_3 \vec{\sigma} + \frac{d}{c dt} \left(\frac{1}{2} \mu_0 \vec{x}^{(0)} \wedge \varrho_3 \vec{\sigma} \right). \quad (45)$$

The first term represents the magnetic moment arising from the motion of the charge ("orbital" moment), the second is the intrinsic magnetic moment connected with the spin. The third term is in any case negligible for particles in a nucleus since it is of higher order in the velocities.

For "free" vector mesons in an electromagnetic field (the word "free" indicating the absence of nuclear sources), a quite analogous decomposition of the current density and magnetic moment is possible. This has been shown by PROCA [21], and somewhat more elaborately by KEMMER [22] from the point of view of the "particle aspect" of meson theory. We repeat KEMMER's derivation by means of the formalism of meson field theory. It will here be convenient to make use of a four-dimensional notation, putting

$$\partial = \frac{\partial}{\partial x^i} - \frac{e}{\hbar c} \mathbf{A}_i \wedge \quad (46)$$

and

$$\mathbf{G}_{ik} = \partial_i \mathbf{U}_k - \partial_k \mathbf{U}_i; \quad (47)$$

thus, we have

$$\left. \begin{aligned} (\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) &= \vec{U}, \quad \mathbf{U}_4 = -\mathbf{U}^4 = -\hat{V}, \\ (\mathbf{G}_{23}, \mathbf{G}_{31}, \mathbf{G}_{12}) &= \hat{\mathbf{G}}, \quad (\mathbf{G}_{14}, \mathbf{G}_{24}, \mathbf{G}_{34}) = -(\mathbf{G}^{14}, \mathbf{G}^{24}, \mathbf{G}^{34}) = \vec{F}. \end{aligned} \right\} (48)$$

The part of the current density (6) corresponding to the vector mesons takes the form

$$\hat{I}^{\text{vect}} = -\frac{e}{\hbar c} \{ \mathbf{G}_{ik} \wedge \mathbf{U}^k \}_3$$

or, according to** (47),

* Use is here made of the formula $\int \vec{x} \wedge \text{rot } \vec{u} dv = 2 \int \vec{u} dv + \text{surface integral}$.

** Use is also made of the relation, easily derived from (46),

$$\partial_i (\mathbf{a} \wedge \mathbf{b}) = (\partial_i \mathbf{a}) \wedge \mathbf{b} + \mathbf{a} \wedge (\partial_i \mathbf{b}).$$

$$\left. \begin{aligned} \hat{I}_i^{\text{vect}} &= -\frac{e}{\hbar c} \left\{ (\partial_i \mathbf{U}_k - \partial_k \mathbf{U}_i) \mathbf{\Lambda} \mathbf{U}^k \right\}_{\mathbf{3}} \\ &= \frac{e}{\hbar c} \left\{ \mathbf{U}^k \mathbf{\Lambda} \partial_i \mathbf{U}_k \right\}_{\mathbf{3}} + \frac{e}{\hbar c} \left\{ \partial_k (\mathbf{U}_i \mathbf{\Lambda} \mathbf{U}^k) \right\}_{\mathbf{3}} \\ &\quad - \frac{e}{\hbar c} \left\{ \mathbf{U}_i \mathbf{\Lambda} \partial_k \mathbf{U}^k \right\}_{\mathbf{3}}. \end{aligned} \right\} \quad (49)$$

From the field equations

$$\kappa^2 \mathbf{U}^i = -\partial_k \mathbf{G}^{ik} \quad (50)$$

we deduce

$$\kappa^2 \partial_i \mathbf{U}^i = -\partial_i \partial_k \mathbf{G}^{ik} = -\frac{1}{2} (\partial_i \partial_k - \partial_k \partial_i) \mathbf{G}^{ik};$$

now, by (46),

$$\partial_i \partial_k - \partial_k \partial_i = -\frac{e}{\hbar c} \mathbf{F}_{ik} \mathbf{\Lambda}; \quad \left(\mathbf{F}_{ik} = \frac{\partial \mathbf{A}_k}{\partial x^i} - \frac{\partial \mathbf{A}_i}{\partial x^k} \right) \quad (51)$$

so that

$$\kappa^2 \partial_i \mathbf{U}^i = \frac{e}{2\hbar c} \mathbf{F}_{ik} \mathbf{\Lambda} \mathbf{G}^{ik} \quad (52)$$

and formula (49) becomes*

$$\left. \begin{aligned} \hat{I}_i^{\text{vect}} &= \frac{e}{\hbar c} \left\{ \mathbf{U}^k \mathbf{\Lambda} \partial_i \mathbf{U}_k \right\}_{\mathbf{3}} + \frac{e}{\hbar c} \frac{\partial}{\partial x^k} \left\{ \mathbf{U}_i \mathbf{\Lambda} \mathbf{U}^k \right\}_{\mathbf{3}} \\ &\quad - \frac{1}{2} \left(\frac{e}{\kappa \hbar c} \right)^2 \left\{ \mathbf{U}_i \mathbf{\Lambda} (\mathbf{F}_{kl} \mathbf{\Lambda} \mathbf{G}^{kl}) \right\}_{\mathbf{3}}. \end{aligned} \right\} \quad (53)$$

The first term represents the ordinary convection current density, including the dependency on the vector potential; the second term, according to its form, is connected with an intrinsic (magnetic and electric) polarization of the vector mesons, while the third arises from a peculiar polarizability of these mesons in an electromagnetic field. It is noteworthy that the intrinsic current density by itself has a vanishing four-divergence. The current density of the pseudoscalar mesons contains, of course, only a convection term which may be written**

* We notice that $\left\{ \partial_i \mathbf{K} \right\}_{\mathbf{3}} = \frac{\partial}{\partial x^i} K_{\mathbf{3}}$.

** We have, according to formula (27) of Part I,

$$\mathbf{a}^\dagger \mathbf{a} \equiv |\mathbf{a}|^2 = \frac{1}{2} (\mathbf{a}_1^2 + \mathbf{a}_2^2).$$

$$\hat{I}^{\text{ps-scal}} = \{\Psi \mathbf{\Lambda} \text{grad } \Psi\}_{\mathbf{3}} - 2 \left(\frac{e}{\hbar c} \right)^2 \vec{A} |\Psi|^2. \quad (54)$$

According to (48), we may write more explicitly for the total current density* (53) and (54)

$$\left. \begin{aligned} \hat{I} = & \frac{e}{\hbar c} \left\{ \sum_{i=1}^3 \mathbf{U}^i \mathbf{\Lambda} \text{grad } \mathbf{U}^i - \hat{\mathbf{V}} \mathbf{\Lambda} \text{grad } \hat{\mathbf{V}} + \Psi \mathbf{\Lambda} \text{grad } \Psi \right\}_{\mathbf{3}} \\ & - 2 \left(\frac{e}{\hbar c} \right)^2 \vec{A} [|\vec{U}|^2 - |\hat{\mathbf{V}}|^2 + |\Psi|^2] \\ & + \frac{e}{2\hbar c} \text{rot} \left\{ \vec{\mathbf{U}} \mathbf{\Lambda} \vec{\mathbf{U}} \right\}_{\mathbf{3}} + \frac{e}{\hbar c} \frac{\partial}{c \partial t} \left\{ \vec{\mathbf{U}} \mathbf{\Lambda} \hat{\mathbf{V}} \right\}_{\mathbf{3}} \\ & - \frac{1}{2} \left(\frac{e}{M_m c^2} \right)^2 \left[\vec{\mathbf{U}}_1 \left(\vec{H} \hat{G}_1 - \vec{E} \vec{F}_1 \right) + \vec{\mathbf{U}}_2 \left(\vec{H} \hat{G}_2 - \vec{E} \vec{F}_2 \right) \right]. \end{aligned} \right\} \quad (55)$$

The magnetic moment is, therefore, (with $\vec{\mathcal{M}} \equiv \vec{x} \mathbf{\Lambda} \text{grad}$)

$$\left. \begin{aligned} \vec{M}_{\text{free mes}} = & \frac{e}{2\hbar c} \int \left\{ \sum_{i=1}^3 \mathbf{U}^i \mathbf{\Lambda} \vec{\mathcal{M}} \mathbf{U}^i - \hat{\mathbf{V}} \mathbf{\Lambda} \vec{\mathcal{M}} \hat{\mathbf{V}} + \Psi \mathbf{\Lambda} \vec{\mathcal{M}} \Psi \right\}_{\mathbf{3}} dv \\ & - 2 \left(\frac{e}{\hbar c} \right)^2 \int (\vec{x} \mathbf{\Lambda} \vec{A}) [|\vec{U}|^2 - |\hat{\mathbf{V}}|^2 + |\Psi|^2] dv \\ & + \frac{e}{2\hbar c} \int \left\{ \vec{\mathbf{U}} \mathbf{\Lambda} \vec{\mathbf{U}} \right\}_{\mathbf{3}} dv + \frac{e}{2\hbar c} \frac{d}{c dt} \int \vec{x} \mathbf{\Lambda} \left\{ \vec{\mathbf{U}} \mathbf{\Lambda} \hat{\mathbf{V}} \right\}_{\mathbf{3}} dv \\ & - \frac{1}{4} \left(\frac{e}{M_m c^2} \right)^2 \int \left[(\vec{x} \mathbf{\Lambda} \vec{\mathbf{U}}_1) \left(\vec{H} \hat{G}_1 - \vec{E} \vec{F}_1 \right) \right. \\ & \left. + (\vec{x} \mathbf{\Lambda} \vec{\mathbf{U}}_2) \left(\vec{H} \hat{G}_2 - \vec{E} \vec{F}_2 \right) \right] dv. \end{aligned} \right\} \quad (56)$$

The two first lines constitute the “orbital” moment, the third line contains the “intrinsic” moment

$$\vec{M}_{\text{intr}} = \frac{e}{2\hbar c} \int \left\{ \vec{\mathbf{U}} \mathbf{\Lambda} \vec{\mathbf{U}} \right\}_{\mathbf{3}} dv \quad (57)$$

and a term involving the commutator of a certain expression with the energy of the free meson field; the last term corresponds

* A still more compendious derivation of this expression in a form similar to (49), but directly embodying also the contribution from the pseudoscalar field, is of course possible with the help of the five-dimensional projective formalism of meson theory.

to an "induced" polarization which we henceforth shall neglect. Comparing the expression (57) of the intrinsic magnetic moment with that of the spin angular momentum (cf. NF (60))

$$-\frac{1}{c} \int \vec{U} \wedge \vec{E} dv,$$

a glance at the formula (A 4) of the Appendix will show that, for a transverse* vector meson of given wave-number \vec{k} and given electric charge, the expectation value of the magnetic moment is

$$\pm \frac{e}{2\hbar} \cdot \frac{1}{\sqrt{k^2 + \kappa^2}}$$

times that of the spin; for slow mesons, this becomes $\pm \frac{e}{2\hbar\kappa} = \pm \frac{e}{2M_m c}$, i. e. the normal value with a "Landé factor" unity [23].

Turning now to the expression (43) of the current density of our system of nucleons and mesons, we shall — leaving aside the contributions from the field-dependent parts $\vec{I}_{(1)} + \vec{I}_{(2)}$ — set up the expression of the magnetic dipole moment of the system. In the first place, we see that the term \vec{I}_{nucl} contributes to this moment a sum of terms of the form (45) (each multiplied by the corresponding factor $\frac{1 - \tau_3}{2}$), while $\vec{I}_{\text{free mes}}$ gives rise to the contribution (56). Here, the symbol $\frac{d}{c dt}$ should, strictly speaking, be understood as indicating $(i\hbar)^{-1}$ times the commutator with the total energy of free nucleons and mesons; but we may with a negligible error replace this by the total energy of the system, including the couplings (of which the most important is the static interaction between the nucleons), since the modifications so introduced involve both the source constants and the nuclear velocities. The terms of the magnetic moment involving $\frac{d}{c dt}$ will then not give any contribution either to the expectation value of the moment in a stationary state of the system or to the

* For a longitudinal meson, the expectation value of spin and magnetic moment is zero. This is also the case for a linearly polarized transverse meson. A well-defined value of the spin ($\pm \hbar$) is obtained for circularly polarized mesons.

matrix elements pertaining to processes in which the total energy is conserved.

From the last term $\vec{I}_{(0)} = \vec{I}_{\text{exch}} + \vec{I}_{\text{prop}}$ of (43) we derive an "exchange magnetic moment" \vec{M}_{exch} and a further contribution \vec{M}_{prop} to the magnetic moments of the single nucleons, which we shall call "proper magnetic moment" [24]. The evaluation of

$$\vec{M}_{(0)} = \frac{1}{2} \int \vec{x} \wedge \vec{I}_{(0)} dv \quad (58)$$

is carried out in the Appendix, the contributions from the vector and from the pseudoscalar meson fields being calculated separately; it again appears that the resulting expression of \vec{M}_{exch} is considerably simpler in the mixed theory than in any other*. Using the notation

$$r^{(ik)} = |\vec{x}^{(i)} - \vec{x}^{(k)}|, \quad \vec{x}_0^{(ik)} = \frac{\vec{x}^{(i)} - \vec{x}^{(k)}}{r^{(ik)}}, \quad (59)$$

it may be written

$$\vec{M}_{\text{exch}} = \frac{e}{4\hbar c} \sum_{i \neq k} \left\{ \begin{aligned} & \left\{ \vec{\mathbf{T}}^{(i)} \wedge \vec{\mathbf{T}}^{(k)} \right\}_{\mathbf{3}} \left\{ \left(\frac{g_2}{\kappa} \right)^2 (\vec{\sigma}^{(i)} \wedge \vec{\sigma}^{(k)}) (1 - 2\kappa r^{(ik)}) \right. \\ & + \left(\frac{g_2}{\kappa} \right)^2 [(\vec{\sigma}^{(i)} \wedge \vec{\sigma}^{(k)}) \cdot \vec{x}_0^{(ik)}] \vec{x}_0^{(ik)} (1 + \kappa r^{(ik)}) \\ & \left. - (g_1^2 + g_2^2 \vec{\sigma}^{(i)} \vec{\sigma}^{(k)}) (\vec{x}^{(i)} \wedge \vec{x}^{(k)}) \right\} \varphi(r^{(ik)}) \end{aligned} \right\} \quad (60)$$

The proper magnetic moment is simply the sum of the contributions from the vector and the pseudoscalar meson fields:

$$\vec{M}_{\text{prop}} = \vec{M}_{\text{prop}}^{\text{vect}} + \vec{M}_{\text{prop}}^{\text{ps-scal}}; \quad (61)$$

the latter are

$$\left. \begin{aligned} \vec{M}_{\text{prop}}^{\text{vect}} &= \frac{e}{4\hbar c} \sum_i \int \left\{ \left\{ \vec{\mathbf{S}}^{(i)}(\vec{x}') \wedge \vec{\mathbf{S}}^{(i)}(\vec{x}'') \right\}_{\mathbf{3}} (1 - \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv'' \right\} \\ \vec{M}_{\text{prop}}^{\text{ps-scal}} &= -\frac{e}{4\hbar c} \sum_i \int \left\{ \left\{ \left\{ \vec{\mathbf{P}}^{(i)}(\vec{x}') \wedge \vec{\mathbf{P}}^{(i)}(\vec{x}'') \right\}_{\mathbf{3}} \kappa |\vec{x}' - \vec{x}''| \right. \right. \\ & \left. \left. - \left[\left\{ \vec{\mathbf{P}}^{(i)}(\vec{x}') \wedge \vec{\mathbf{P}}^{(i)}(\vec{x}'') \right\}_{\mathbf{3}} \cdot \vec{x}_0 \right] \vec{x}_0 (1 + \kappa |\vec{x}' - \vec{x}''|) \right\} \varphi dv' dv'' \right\} \end{aligned} \right\} \quad (62)$$

* It should be noted that the simplification occurs in the part of \vec{M}_{exch} which is not translation invariant (cf. the Appendix), *i. e.* directly results from the simpler form of \vec{J} on mixed theory.

with $\vec{x}_0 \equiv \frac{\vec{x}' - \vec{x}''}{|\vec{x}' - \vec{x}''|}$. In these formulae, $\vec{S}^{(i)}$ and $\vec{P}^{(i)}$ are the contributions of the i -th nucleon to the source densities \vec{S} and \vec{P} (cf. NF, p. 10).

According to NF (4, 24) we may write

$$\left. \begin{aligned} \vec{S}^{(i)}(\vec{x}) &= \frac{g_2}{\kappa} \mathbf{T}^{(i)} \varrho_3^{(i)} \vec{\sigma}^{(i)} D(\vec{x} - \vec{x}^{(i)}) \\ \vec{P}^{(i)}(\vec{x}) &= \frac{f_2}{\kappa} \mathbf{T}^{(i)} \vec{\sigma}^{(i)} D(\vec{x} - \vec{x}^{(i)}), \end{aligned} \right\} \quad (64)$$

the delta-function being replaced by a continuous distribution function $D(\vec{x} - \vec{x}^{(i)})$ (with $\int D(\vec{x}) dv = 1$). The formulae (62) and (63) then take the simpler form

$$\vec{M}_{\text{prop}} = -\mu_0 \mu \sum_i v_{\mathbf{3}}^{(i)} \vec{\sigma}^{(i)}, \quad (65)$$

where μ_0 denotes (as above) the nuclear magneton, while

$$\mu^{\text{vect}} = \frac{g_2^2}{\hbar c} \cdot \frac{M}{M_m} \cdot \frac{2}{\kappa} \int D(\vec{x}') D(\vec{x}'') (1 - \kappa \varrho) \varphi(\varrho) dv' dv'' \quad (66)$$

$$\mu^{\text{ps-scal}} = \frac{f_2^2}{\hbar c} \cdot \frac{M}{M_m} \cdot \frac{2}{3\kappa} \int D(\vec{x}') D(\vec{x}'') (1 - 2\kappa \varrho) \varphi(\varrho) dv' dv'', \quad (67)$$

with $\varrho \equiv |\vec{x}' - \vec{x}''|$; the last formula is valid on the assumption that the distribution function $D(\vec{x})$ is spherically symmetrical. The value of μ on the mixed theory is therefore (putting $f_2^2 = g_2^2$)

$$\left. \begin{aligned} \mu &= \mu^{\text{vect}} + \mu^{\text{ps-scal}} \\ &= \frac{g_2^2}{\hbar c} \cdot \frac{M}{M_m} \cdot \frac{2}{3\kappa} \int D(\vec{x}') D(\vec{x}'') (4 - 5\kappa \varrho) \varphi(\varrho) dv' dv''. \end{aligned} \right\} \quad (68)$$

According to (65), we should expect the total magnetic moments of proton (at rest) and neutron to be $\vec{\sigma} \mu_0 \mu_P$ and $\vec{\sigma} \mu_0 \mu_N$, respectively, with

$$\mu_P = 1 + \mu, \quad \mu_N = -\mu.$$

For a closer discussion of the expressions (66), (67), and (68) for μ , the reader is referred to the forthcoming paper by SERPE: it appears that the mixed theory will under reasonable assumptions for the distribution function D yield a value of μ of the right sign and the right order of magnitude. From the latest experimental determinations of the magnetic moments of proton [25] and neutron [26],

$$\left. \begin{aligned} \mu_P &= 2,7896 \pm 0,0008 \\ \mu_N &= -1,935 \pm 0,02, \end{aligned} \right\} \quad (70)$$

it would seem, however, that there is a slight difference between the "proper" magnetic moments of the two particles. Such a dissymmetry cannot be accounted for by the present theory, since it only arises in a higher approximation which should be discarded according to our general prescription. The manner in which the dissymmetry appears has been explained by FRÖHLICH, HEITLER and KEMMER [24]: in the expression of the self-energy of a nucleon in an external magnetic field, the second approximation terms consist of quotients of matrix elements independent of the magnetic field by differences of energy values for the initial and the intermediate states considered; now, such differences will contain a term proportional to the magnetic field which, when due account is taken of the magnetic energy $\pm \mu_0 H$ of the free proton states, turns out to be different for the alternative cases of a proton or a neutron. This effect clearly falls outside the scope of our theory* and can only be expected to present itself in a theory yielding a correct treatment of self-energy problems; we may provisionally allow for it by replacing the above expression (65) of \vec{M}_{prop} by

* It may be pointed out that an asymmetry of the kind here discussed could not be derived from the assumption of a small deviation from the symmetrical KEMMER combination of charged and neutral meson fields. In fact, the most general combination of such fields would be obtained by replacing in the expressions of the source densities each τ_i by a linear combination of the form $\sum_K a_{iK} \tau_K + b_i$ with c-number coefficients a_{iK}, b_i . Then, the same combination would also appear in the expressions of the static potentials and fields, and it is immediately apparent from formula (42) that this would not give rise to any term independent of τ_3 in the expression of the proper moment, such as would be required for the asymmetry effect in question.

$$\vec{M}_{\text{prop}}^{\text{exp}} = \mu_0 \sum_i \left\{ \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} (\mu_P - 1) + \frac{1 + \tau_{\mathbf{3}}^{(i)}}{2} \mu_N \right\} \vec{\sigma}^{(i)}, \quad (71)$$

where μ_P and μ_N have their empirical values (70).

§ 6. Electromagnetic properties of the deuteron.

A few remarks may still be added concerning the application of the preceding formulae to the simplest nuclear system, the deuteron. Comparing the treatment in meson theory with that in the previous theories, in which the proton was treated as a point charge in some arbitrarily assumed short range nuclear field, we see that, as regards the electric properties, there is no other difference than that arising from the influence of the latter field on the form of the wave functions. In fact, as stated in section 3 above, the exchange dipole moment gives only a negligible contribution, while the exchange quadrupole moment with respect to the centre of gravity of the deuteron vanishes according to formula (28). The exchange magnetic moment, on the other hand, does play an appreciable part which, in some cases, may even be quite considerable; examples of this are provided by PAIS' investigation of the photomagnetic effect of the deuteron and the inverse process of neutron capture by protons [7].

Since the exchange terms do not contribute anything to the expectation values of the corresponding quantities in stationary states of the system, no new element is introduced by our theory into the situation with regard to the magnetic moment of the deuteron in relation to its quadrupole moment. The existence of the latter, with a value estimated [27] to be

$$Q = 2,73 \cdot 10^{-27} \text{ cm}^2 \quad (72)$$

with an uncertainty of about 2 %, implies that the ground state of the deuteron must be a mixture of a 3S state with a small amount δ of 3D_1 state. The magnetic moment of the ground state can then easily be calculated in terms of δ . The exact expression of the z -component of the magnetic moment is

$$M_z = \mu_0 \sum_{i=1,2} \left\{ \left(\frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \mu_P + \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \mu_N \right) \sigma_z^{(i)} + \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} l_z^{(i)} - \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} \cdot (1 - \rho_3^{(i)}) (l_z^{(i)} + \sigma_z^{(i)}) \right\}, \quad (73)$$

$l_z^{(i)}$ denoting the z -component of the orbital angular momentum of the i -th nucleon. For its expectation value* in a stationary state of the nucleus, we get, since the mean value of $\tau_3^{(i)}$ is zero and that of $\varrho_3^{(i)}$ is approximately unity**,

$$\mathfrak{M}\{M_z\} \approx \mu_0 \left[(\mu_P + \mu_N) \mathfrak{M}\left\{\frac{1}{2}S_z\right\} + \frac{1}{2}\mathfrak{M}\{L_z\} \right],$$

where S_z , L_z represent the z -components of the total spin and total orbital momentum, respectively. Characterizing the stationary state in the usual way by the quantum numbers L , S , J , m , this may be written

$$\mathfrak{M}\{M_z\} \approx \mu_0 m \mu_D^{(LSJ)},$$

with the Landé factor

$$\begin{aligned} \mu_D^{(LSJ)} &= (\mu_P + \mu_N) \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \\ &+ \frac{1}{2} \frac{J(J+1) + L(L+1) - S(S+1)}{2J(J+1)}. \end{aligned}$$

For the ground state we thus get

$$\left. \begin{aligned} \mu_D &= (1 - \delta) (\mu_P + \mu_N) + \delta \left[-\frac{1}{2} (\mu_P + \mu_N) + \frac{3}{4} \right] \\ &= \left(1 - \frac{3}{2} \delta \right) (\mu_P + \mu_N) + \frac{3}{4} \delta. \end{aligned} \right\} (74)$$

The direct determination of μ_D [25],

$$\mu_D = 0,8565 \pm 0,0004, \quad (75)$$

lies very close to $\mu_N + \mu_P$, which, according to (70), is $0,85 \pm 0,02$; on the other hand, the amount δ of D -state necessary to account for the quadrupole moment (72), while depending on the form of meson theory adopted, is of the order of magnitude of a few percent***. So far, there is thus no certain discrepancy between

* The expectation value of A will be denoted by $\mathfrak{M}\{A\}$.

** The relativistic correction arising from the factors $(1 - \varrho_3^{(i)})$ was discussed by H. MARGENAU [28].

*** This is also the case, as a rough estimate shows, with the derivation of the quadrupole moment based on the non-static directional coupling of the

formula (74) and the empirical results (70), (75) and (72), though the margin of error appears rather strained. A decision on this point can only be arrived at by more precise measurements, especially of the neutron moment. It should also be remembered that a future theory, taking a correct account of the universal limiting length, might, analogously to the dissymmetry effect discussed in section 5, yield a small correction to the purely additive expression (73) of M_z arising from the nuclear field interaction between the neutron and the proton.

mixed theory, as developed in NF, Part III, if the nuclear source constants are chosen so as to give the right order of magnitude for the quadrupole moment. (If, however, the relations between these constants suggested by the five-dimensional formalism [4] are adopted, the quadrupole moment vanishes in that approximation).

Appendix.

Mean values for the zero-point meson field.

Let us take as usual progressive plane waves

$$u(\vec{k}; \vec{x}) = L^{-3/2} e^{i \vec{k} \cdot \vec{x}}, \quad (\text{A } 1)$$

satisfying a cyclic condition within a cube of side L :

$$\vec{k} \equiv \frac{2\pi}{L} (n_1, n_2, n_3); \quad (n\text{'s integers}) \quad (\text{A } 2)$$

the wave-vector \vec{k} is connected with the momentum \vec{p} and energy E of a meson by

$$\left. \begin{aligned} \vec{p} &= \hbar \vec{k} \\ E &= \hbar c \varepsilon_k \\ \text{with } \varepsilon_k &\equiv \sqrt{k^2 + \kappa^2}. \end{aligned} \right\} \quad (\text{A } 3)$$

We introduce quantized amplitudes $\mathbf{a}(\vec{k}; j)$ referring to mesons of a given wave-vector (momentum and energy); the index j will be given the value 0 for pseudoscalar mesons and the values 1, 2, 3 for vector mesons, 1 and 2 referring to two independent kinds of transversal (linear) polarization of the vector meson field, 3 to the longitudinal polarization. The directions of polarization will be characterized by three mutually orthogonal unit vectors $\vec{e}_j(\vec{k})$.

The meson fields may then be expressed as a superposition of such plane waves in the following way:

$$\left. \begin{aligned}
 \vec{U} &= \sqrt{\frac{\hbar c}{2}} \sum_{\vec{k}} \frac{1}{\sqrt{\varepsilon_k}} \sum_{j=1}^3 \vec{e}_j(\vec{k}) \cdot \left(\frac{\varepsilon_k}{\kappa}\right)^{j3} [\mathbf{a}(\vec{k}; j) u(\vec{k}) + \mathbf{a}^\dagger(\vec{k}; j) u^*(\vec{k})] \\
 \vec{F} &= i \sqrt{\frac{\hbar c}{2}} \sum_{\vec{k}} \sqrt{\varepsilon_k} \sum_{j=1}^3 \vec{e}_j(\vec{k}) \cdot \left(\frac{\kappa}{\varepsilon_k}\right)^{j3} [\mathbf{a}(\vec{k}; j) u(\vec{k}) - \mathbf{a}^\dagger(\vec{k}; j) u^*(\vec{k})], \\
 \Psi &= \sqrt{\frac{\hbar c}{2}} \sum_{\vec{k}} \frac{1}{\sqrt{\varepsilon_k}} [\mathbf{a}(\vec{k}; 0) u(\vec{k}) + \mathbf{a}^\dagger(\vec{k}; 0) u^*(\vec{k})] \\
 \Phi &= -i \sqrt{\frac{\hbar c}{2}} \sum_{\vec{k}} \sqrt{\varepsilon_k} [\mathbf{a}(\vec{k}; 0) u(\vec{k}) - \mathbf{a}^\dagger(\vec{k}; 0) u^*(\vec{k})].
 \end{aligned} \right\} \quad (\text{A 4})$$

The commutation rules for the amplitudes are

$$\left. \begin{aligned}
 [a_{\mathbf{n}}(\vec{k}; j), a_{\mathbf{n}'}^\dagger(\vec{k}'; j')]_- &= \delta_{\mathbf{n}\mathbf{n}'} \delta(\vec{k}, \vec{k}') \delta_{jj'}, \\
 \text{all other pairs commuting;}
 \end{aligned} \right\} \quad (\text{A 5})$$

the operators

$$N_{\mathbf{n}}(\vec{k}; j) = a_{\mathbf{n}}^\dagger(\vec{k}; j) a_{\mathbf{n}}(\vec{k}; j) \quad (\text{A 6})$$

represent the numbers of mesons of kind j , wave-vector \vec{k} and character \mathbf{n} with respect to the electric charge: $N_{\mathbf{3}}$ is the number of neutral mesons, $N_{\mathbf{1}} + N_{\mathbf{2}}$ the total number of positively and negatively charged mesons. We get the numbers of positive or negative mesons, separately, when replacing the amplitudes $a_{\mathbf{1}}$, $a_{\mathbf{2}}$ by

$$b_{\mathbf{1}} = \frac{1}{\sqrt{2}} (a_{\mathbf{1}} + i a_{\mathbf{2}}), \quad b_{\mathbf{2}} = \frac{1}{\sqrt{2}} (a_{\mathbf{1}} - i a_{\mathbf{2}}),$$

which obey similar commutation rules; then

$$N'_{\mathbf{1}} = b_{\mathbf{1}}^\dagger b_{\mathbf{1}}, \quad N'_{\mathbf{2}} = b_{\mathbf{2}}^\dagger b_{\mathbf{2}}$$

are the numbers of positive and negative mesons, respectively, the total charge of the meson field taking the form

$$\int \varrho_{\text{mes}} dv = e \sum_{\vec{k}} \sum_{j=0}^3 [N'_{\mathbf{1}}(\vec{k}; j) - N'_{\mathbf{2}}(\vec{k}; j)].$$

For the state corresponding to no meson, the above formulae yield, with

$$v(\vec{k}; \vec{x} - \vec{x}') \equiv u(\vec{k}; \vec{x}) u^*(\vec{k}; \vec{x}') = L^{-3} e^{i\vec{k}(\vec{x} - \vec{x}')},$$

the following mean values:

$$\begin{aligned} \mathfrak{M}_0 \{F_n^\mu(\vec{x}) F_n^\nu(\vec{x}')\} &= \frac{\hbar c}{2} \sum_{\vec{k}} \varepsilon_k \sum_j e_j^\mu e_j^\nu \left(\frac{\kappa}{\varepsilon_k}\right)^{2j} v(\vec{k}; \vec{x} - \vec{x}') \\ &= \frac{\hbar c}{2} \sum_{\vec{k}} \varepsilon_k \left[\delta^{\mu\nu} + \frac{k^\mu k^\nu}{k^2} \left(\frac{\kappa^2}{\varepsilon_k^2} - 1\right) \right] v(\vec{k}; \vec{x} - \vec{x}') \\ &= \frac{\hbar c}{2} \sum_{\vec{k}} \varepsilon_k \left(\delta^{\mu\nu} + \frac{k^\mu k^\nu}{\varepsilon_k^2} \right) v(\vec{k}; \vec{x} - \vec{x}'), \\ \mathfrak{M}_0 \{U_n^\mu(\vec{x}) U_n^\nu(\vec{x}')\} &= \frac{\hbar c}{2} \sum_{\vec{k}} \frac{1}{\varepsilon_k} \left[\delta^{\mu\nu} + \frac{k^\mu k^\nu}{k^2} \left(\frac{\varepsilon_k^2}{\kappa^2} - 1\right) \right] v(\vec{k}; \vec{x} - \vec{x}') \\ &= \frac{\hbar c}{2} \sum_{\vec{k}} \frac{1}{\varepsilon_k} \left(\delta^{\mu\nu} - \frac{k^\mu k^\nu}{\kappa^2} \right) v(\vec{k}; \vec{x} - \vec{x}'), \\ \mathfrak{M}_0 \{\Phi_n(\vec{x}) \Phi_n(\vec{x}')\} &= \frac{\hbar c}{2} \sum_{\vec{k}} \varepsilon_k v(\vec{k}; \vec{x} - \vec{x}'); \end{aligned} \quad (A7)$$

from the canonical commutation rules for \vec{U} and \vec{F} , it may further be deduced without calculation:

$$\mathfrak{M}_0 \{U_n^\mu(\vec{x}) F_n^\nu(\vec{x}')\} = -\mathfrak{M}_0 \{F_n^\mu(\vec{x}) U_n^\nu(\vec{x}')\} = \frac{\hbar c}{2i} \delta^{\mu\nu} \delta(\vec{x} - \vec{x}'); \quad (A8)$$

all other mean values of products of two field components vanish. The summations (or, in the limit $L \rightarrow \infty$, integrations) over the wave-vector components occurring in the preceding formulae may all be derived from the relation

$$K_0(zr) = 2\pi^2 \lim_{L \rightarrow \infty} \sum_{\vec{k}} \frac{1}{\varepsilon_k^3} v(\vec{k}; \vec{x} + \vec{x}'), \quad (r \equiv |\vec{x} - \vec{x}'|) \quad (A9)$$

where $K_0(z)$ is a Bessel function of the second kind with imaginary argument, connected with Hankel's functions by the general formula*

* For the definition and properties of $K_n(z)$, see G. WATSON, A treatise on the theory of Bessel functions (1922), 3'7 (p. 77), 3'71 (p. 79), 6'16 (p. 172). In E. WHITTAKER and G. WATSON, A course of modern analysis, 17'71 (4th Ed., p. 373), a function called $K_n(z)$, but differing from the one here used by a factor $(-1)^n$, is considered.

$$K_n(z) = \frac{\pi i}{2} e^{\frac{\pi i}{2} n} H_n^{(1)}(iz).$$

In fact, since

$$(\kappa^2 - \mathcal{A}) v = \epsilon_k^2 v,$$

we get

$$\left. \begin{aligned} \frac{\hbar c}{2} \sum_{\vec{k}} \frac{1}{\epsilon_k} v(\vec{k}; \vec{x} - \vec{x}') &= \frac{\hbar c}{4\pi^2} (\kappa^2 - \mathcal{A}) K_0(\kappa r) \\ &= \frac{\hbar c}{4\pi^2} \cdot \frac{\kappa}{r} K_1(\kappa r) \equiv \mathcal{S}(r), \\ \frac{\hbar c}{2} \sum_{\vec{k}} \epsilon_k v(\vec{k}; \vec{x} - \vec{x}') &= (\kappa^2 - \mathcal{A}) \mathcal{S}(r) \equiv \mathcal{D}(r); \end{aligned} \right\} \text{(A 10)}$$

the formulae (A 7) may then be written as follows:

$$\left. \begin{aligned} \mathfrak{M}_0 \{ F_{\mathbf{n}}^\mu(\vec{x}) F_{\mathbf{n}}^\nu(\vec{x}') \} &= \delta^{\mu\nu} \mathcal{D}(r) + \frac{\partial^2}{\partial x^\mu \partial x^\nu} \mathcal{S}(r) \\ \mathfrak{M}_0 \{ U_{\mathbf{n}}^\mu(\vec{x}) U_{\mathbf{n}}^\nu(\vec{x}') \} &= \delta^{\mu\nu} \mathcal{S}(r) - \frac{1}{\kappa^2} \frac{\partial^2}{\partial x^\mu \partial x^\nu} \mathcal{S}(r) \\ \mathfrak{M}_0 \{ \Phi_{\mathbf{n}}(\vec{x}) \Phi_{\mathbf{n}}(\vec{x}') \} &= \mathcal{D}(r). \end{aligned} \right\} \text{(A 11)}$$

We are interested in

$$\left. \begin{aligned} \mathfrak{M}_1 &= \int dv' \mathfrak{M}_0 \left\{ \left(\vec{f}^{(i)} \vec{U}_{\mathbf{1}} \right)_{\vec{x}} \left(\vec{f}^{(i)} \vec{U}_{\mathbf{1}} \right)_{\vec{x}'} \right\} \\ \mathfrak{M}_2 &= \int dv' \mathfrak{M}_0 \left\{ \left(\vec{f}^{(i)} \wedge \vec{F}_{\mathbf{1}} \right)_{\vec{x}} \left(\vec{f}^{(i)} \wedge \vec{F}_{\mathbf{1}} \right)_{\vec{x}'} \right\} \\ \mathfrak{M}_3 &= \int dv' \vec{f}^{(i)} \vec{f}^{(i)'} \mathfrak{M}_0 \{ \Phi_{\mathbf{n}} \Phi_{\mathbf{n}}' \}, \end{aligned} \right\} \text{(A 12)}$$

where

$$\vec{f}^{(i)}(\vec{x}) = \text{grad}^{(i)} \varphi (|\vec{x} - \vec{x}^{(i)}|) = -\text{grad} \varphi.$$

From (A 11) we get

Formula (A 9), after the summation over \vec{k} on the right has been replaced in the limit by $\frac{L^3}{8\pi^3} \int \dots dk_x dk_y dk_z$, or $\frac{L^3}{8\pi^3} \int \dots k^2 dk d\Omega$, and the angle integration performed, takes the form

$$K_0(\kappa r) = \frac{1}{r} \int_0^\infty \frac{\sin kr \cdot k dk}{(k^2 + \kappa^2)^{3/2}};$$

this last relation is in its turn readily derived from Basset's formula (see G. WATSON, loc. cit. 6'16 (p. 172))

$$K_0(\kappa r) = \int_0^\infty \frac{\cos k dk}{\sqrt{k^2 + (\kappa r)^2}} = \int_0^\infty \frac{\cos kr dk}{\sqrt{k^2 + \kappa^2}}$$

by a partial integration.

$$\left\{ \begin{aligned} \mathfrak{M}_1 &= \int dv' \vec{f}^{(i)} \vec{f}^{(i')} \mathcal{E} - \frac{1}{\kappa^2} \int dv' \sum_{\mu, \nu} f^{(i)\mu}(\vec{x}) f^{(i)\nu}(\vec{x}') \frac{\partial^2}{\partial x^\mu \partial x^\nu} \mathcal{E} \\ \mathfrak{M}_2 &= \int dv' \vec{f}^{(i)} \vec{f}^{(i')} (2 \mathcal{D} + \mathcal{A} \mathcal{E}) - \int dv' \sum_{\mu, \nu} f^{(i)\mu}(\vec{x}) f^{(i)\nu}(\vec{x}') \frac{\partial^2}{\partial x^\mu \partial x^\nu} \mathcal{E} \\ \mathfrak{M}_3 &= \int dv' \vec{f}^{(i)} \vec{f}^{(i')} \mathcal{D}. \end{aligned} \right.$$

By partial integrations, the last integral in \mathfrak{M}_1 and \mathfrak{M}_2 is easily transformed into

$$\int dv' \vec{f}^{(i)} \vec{f}^{(i')} \mathcal{A} \mathcal{E},$$

so that, taking account of (A 10), we get

$$\mathfrak{M}_1 = \kappa^{-2} \mathfrak{M}_3, \quad \mathfrak{M}_2 = 2 \mathfrak{M}_3. \quad (\text{A } 13)$$

Further, since $\varphi(|\vec{x} - \vec{x}^{(i)}|)$ satisfies the equation

$$(\kappa^2 - \mathcal{A}) \varphi = \delta(\vec{x} - \vec{x}^{(i)}),$$

we may write

$$\begin{aligned} \mathfrak{M}_3 &= \vec{f}^{(i)} \cdot \text{grad}^{(i)} \int dv' \varphi(|\vec{x}' - \vec{x}^{(i)}|) (\kappa^2 - \mathcal{A}') \mathcal{E} \\ &= \vec{f}^{(i)} \cdot \text{grad}^{(i)} \mathcal{E}(|\vec{x} - \vec{x}^{(i)}|) \\ &= \frac{d\varphi(r^{(i)})}{dr^{(i)}} \cdot \frac{d\mathcal{E}(r^{(i)})}{dr^{(i)}} \end{aligned}$$

with $r^{(i)} \equiv |\vec{x} - \vec{x}^{(i)}|$. From (A 10) and a recurrence formula for the K_n , we finally obtain

$$\left. \begin{aligned} \mathfrak{M}_3 &= -\frac{d\varphi(r^{(i)})}{dr^{(i)}} \cdot \frac{\hbar c}{4\pi^2} \cdot \frac{\kappa^2}{r^{(i)}} K_2(\kappa r^{(i)}) \\ &= \frac{e^{-\kappa r^{(i)}}}{4\pi r^{(i)}} \left(\kappa + \frac{1}{r^{(i)}} \right) \frac{\hbar c}{4\pi^2} \cdot \frac{\kappa^2}{r^{(i)}} K_2(\kappa r^{(i)}). \end{aligned} \right\} \quad (\text{A } 14)$$

The exchange and proper current density and related quantities.

We shall here be concerned with the calculation of

$$\int \vec{I}_{(0)} dv \quad \text{and} \quad \vec{M}_{(0)} = \frac{1}{2} \int \vec{x} \wedge \vec{I}_{(0)} dv,$$

$\vec{I}_{(0)}$ being given by (42). For the evaluation of the first integral, we need formula NF (89), which we re-write in the form

$$\psi(|\vec{x}-\vec{x}'|) \equiv \int \varphi(|\vec{x}-\vec{x}''|) \varphi(|\vec{x}'-\vec{x}''|) dv'' = \frac{e^{-\kappa|\vec{x}-\vec{x}'|}}{8\pi\kappa}; \quad (\text{A } 15)$$

from this, we deduce

$$\text{grad } \psi(|\vec{x}-\vec{x}'|) = -\frac{1}{2}(\vec{x}-\vec{x}') \varphi(|\vec{x}-\vec{x}'|) \quad (\text{A } 16)$$

and (since $\mathcal{A}\varphi - \kappa^2\varphi = -\delta(\vec{x}-\vec{x}')$)

$$\mathcal{A}\psi - \kappa^2\psi = -\varphi(|\vec{x}-\vec{x}'|). \quad (\text{A } 17)$$

On account of the symmetrization bars in formula (42), we may in all following calculations freely interchange the order of factors, even though they are not commutable; for simplicity, we shall omit the bars from now on. From NF (9, 14) we get, in the first place,

$$\vec{E}^\circ \wedge \vec{V}^\circ = \int \mathbf{N}' \wedge \mathbf{N}'' \text{grad}' \varphi(|\vec{x}-\vec{x}'|) \varphi(|\vec{x}-\vec{x}''|) dv' dv'', \quad (\text{A } 18)$$

whence, by (A 15),

$$\left. \begin{aligned} \int \vec{E}^\circ \wedge \vec{V}^\circ dv &= \int \mathbf{N}' \wedge \mathbf{N}'' \text{grad}' \psi(|\vec{x}'-\vec{x}''|) dv' dv'' \\ &= \frac{1}{2} \int \mathbf{N}' \wedge \mathbf{N} (\vec{x}-\vec{x}') \varphi(|\vec{x}-\vec{x}'|) dv dv', \end{aligned} \right\} (\text{A } 19)$$

by (A 16). The $\mathbf{3}$ -component of this quantity, multiplied by $-\frac{e}{\hbar c}$, coincides with the first term of the expression \vec{J} given by (31).

Similarly, we get from NF (14)

$$\begin{aligned} \vec{G}^\circ \wedge \vec{U}^\circ &= -\int \vec{G}^\circ \wedge (\vec{S}' \wedge \text{grad}' \varphi) dv' \\ &= -\int (\vec{S}' \wedge \vec{G}^\circ) \text{grad}' \varphi dv' + \int \vec{S}' \wedge (\vec{G}^\circ \text{grad}' \varphi) dv'; \end{aligned}$$

inserting in the first term the value NF (37) of \vec{G}° and in the second its value NF (10), we have

$$\left. \begin{aligned} \vec{G}^\circ \wedge \vec{U}^\circ &= -\kappa^2 \int (\vec{S}' \wedge \vec{S}'') \operatorname{grad}' \varphi (|\vec{x} - \vec{x}'|) \varphi (|\vec{x} - \vec{x}''|) dv' dv'' \\ &\quad - \int [(\vec{S}'' \operatorname{grad}'') \wedge (\vec{S}' \operatorname{grad}')] \\ &\quad \quad \quad \operatorname{grad}' \varphi (|\vec{x} - \vec{x}'|) \varphi (|\vec{x} - \vec{x}''|) dv' dv'' \\ &\quad + \int \vec{S}' \wedge (\vec{S} \operatorname{grad}' \varphi) dv' - \int \vec{S}' \wedge \operatorname{div} (\operatorname{grad}' \varphi \wedge \vec{U}^\circ) dv', \end{aligned} \right\} \text{(A 20)}$$

since $\operatorname{rot} \vec{U}^\circ \cdot \operatorname{grad}' \varphi = -\operatorname{div} (\operatorname{grad}' \varphi \wedge \vec{U}^\circ)$. Again using A (15) and A (16), we obtain immediately

$$\left. \begin{aligned} \int \vec{G}^\circ \wedge \vec{U}^\circ dv &= -\frac{1}{2} \int \kappa^2 (\vec{S}' \wedge \vec{S}') (\vec{x} - \vec{x}') \varphi (|\vec{x} - \vec{x}'|) dv dv' \\ &\quad + \int \{ [(\vec{S}' \operatorname{grad}') \wedge (\vec{S} \operatorname{grad})] \operatorname{grad} \psi \\ &\quad - \vec{S}' \wedge (\vec{S} \operatorname{grad} \varphi) \} dv dv', \end{aligned} \right\} \text{(A 21)}$$

the integral over the last term of (A 20) reducing to a vanishing surface integral. The first term of (A 21), by taking the $\mathbf{3}$ -component and multiplying by $\frac{e}{\hbar c}$, gives just the remaining term of the expression (31) of \vec{J} . The last term of (A 21) would subsist in the integral exchange current on a pure vector meson theory; in the mixed theory, however, it is compensated by the contribution from the pseudoscalar field.

In fact, from NF (30, 40) we have

$$\left. \begin{aligned} \vec{I}^\circ \wedge \Psi^\circ &= \int \vec{P} \wedge (\vec{P} \operatorname{grad}' \varphi) dv' \\ &\quad + \int [(\vec{P}' \operatorname{grad}'') \wedge (\vec{P} \operatorname{grad}')] \operatorname{grad}'' \varphi (|\vec{x} - \vec{x}'|) \varphi (|\vec{x} - \vec{x}''|) dv' dv'', \end{aligned} \right\} \text{(A 22)}$$

whence, by (A 15),

$$\left. \begin{aligned} \int \vec{I}^\circ \wedge \Psi^\circ dv &= \int \{ \vec{P} \wedge (\vec{P} \operatorname{grad} \varphi) \\ &\quad - [(\vec{P} \operatorname{grad}') \wedge (\vec{P} \operatorname{grad})] \operatorname{grad} \psi \} dv dv'. \end{aligned} \right\} \text{(A 23)}$$

In such an expression we may indeed, on the mixed theory, substitute \vec{S} for \vec{P} , apart from terms of higher order in the

nuclear velocities. Summing up, we have thus established formula (44).

We now turn to the calculation of $\vec{M}_{(0)}$, which runs entirely parallel to the preceding derivation of (44). Only, the role of the auxiliary relations (A 15), (A 16) is in this case played by

$$\left. \begin{aligned} \vec{\psi}(|\vec{x}-\vec{x}'|) &\equiv \int \varphi(|\vec{x}-\vec{x}''|) \varphi(|\vec{x}'-\vec{x}''|) \vec{x}'' dv'' \\ &= \frac{e^{-\kappa|\vec{x}-\vec{x}'|}}{8\pi\kappa} \cdot \frac{\vec{x}+\vec{x}'}{2}, \end{aligned} \right\} \text{(A 24)}$$

$$\left. \begin{aligned} \text{rot } \vec{\psi}(|\vec{x}-\vec{x}'|) &= -\frac{1}{2} \varphi(|\vec{x}-\vec{x}'|) \cdot (\vec{x}-\vec{x}') \wedge \frac{\vec{x}+\vec{x}'}{2} \\ &= -\frac{1}{2} \varphi(|\vec{x}-\vec{x}'|) \cdot \vec{x} \wedge \vec{x}'. \end{aligned} \right\} \text{(A 25)}$$

To begin with, we get from (A 18) and (A 24), (A 25),

$$\left. \begin{aligned} &\int \vec{x} \wedge (\vec{F}^{\circ} \wedge \vec{V}^{\circ}) dv = \\ &= -\int \mathbf{N}' \wedge \mathbf{N}'' \text{rot}' \{ \vec{x} \varphi(|\vec{x}-\vec{x}'|) \varphi(|\vec{x}-\vec{x}''|) \} dv' dv'' \\ &= -\int \mathbf{N}' \wedge \mathbf{N}'' \text{rot}' \vec{\psi}(|\vec{x}'-\vec{x}''|) dv' dv'' \\ &= \frac{1}{2} \int \mathbf{N}' \wedge \mathbf{N}'' \varphi(|\vec{x}'-\vec{x}''|) (\vec{x}' \wedge \vec{x}'') dv' dv''. \end{aligned} \right\} \text{(A 26)}$$

Further, from (A 20), we get by the same procedure

$$\begin{aligned} &\int \vec{x} \wedge (\vec{G}^{\circ} \wedge \vec{U}^{\circ}) dv = \\ &= -\frac{1}{2} \kappa^2 \int (\vec{S}' \wedge \vec{S}'') (\vec{x}' \wedge \vec{x}'') \varphi(|\vec{x}'-\vec{x}''|) dv' dv'' \\ &\quad -\frac{1}{2} \int [(\vec{S}'' \text{grad}'') \wedge (\vec{S}' \text{grad}')] (\vec{x}' \wedge \vec{x}'') \varphi(|\vec{x}'-\vec{x}''|) dv' dv'' \\ &\quad -\int (\vec{x}'' \wedge \vec{S}') \wedge (\vec{S}'' \text{grad}'' \varphi) dv' dv'' \\ &\quad -\int (\vec{x} \wedge \vec{S}') \wedge \text{div}(\text{grad}' \varphi \wedge \vec{U}^{\circ}) dv' dv''; \end{aligned}$$

the second and third terms together give

$$\frac{1}{2} \int (\vec{x}' - \vec{x}'') \wedge \vec{S}' \mathbf{\Lambda} (\vec{S}'' \text{grad}'' \varphi) \, d\vec{v}' \, d\vec{v}'' + \frac{1}{2} \int (\vec{x}' \wedge \vec{x}'') (\vec{S}'' \text{grad}'' \mathbf{\Lambda} (\vec{S}' \text{grad}' \varphi) \, d\vec{v}' \, d\vec{v}'',$$

while the last becomes

$$\begin{aligned} & - \int \vec{S}' \wedge (\text{grad}' \varphi \wedge \vec{T}^\circ) \, d\vec{v}' \, d\vec{v}'' = \int \vec{S}' \wedge \mathbf{\Lambda} [\text{grad}' \varphi (\vec{x} - \vec{x}') \wedge (\vec{S}'' \wedge \text{grad}'' \varphi (\vec{x} - \vec{x}''))] \, d\vec{v}' \, d\vec{v}'' \\ & = \int (\vec{S}' \wedge \vec{S}'') (\text{grad}' \varphi (\vec{x} - \vec{x}')) \cdot \text{grad}'' \varphi (\vec{x}' - \vec{x}'') \, d\vec{v}' \, d\vec{v}'' \\ & \quad - \int [\vec{S}' \wedge \text{grad}'' \varphi (\vec{x} - \vec{x}'')] \mathbf{\Lambda} [\vec{S}'' \cdot \text{grad}' \varphi (\vec{x} - \vec{x}')] \, d\vec{v}' \, d\vec{v}'' \\ & = - \int \vec{S}' \wedge \vec{S}'' \mathcal{L}' \psi \, d\vec{v}' \, d\vec{v}'' - \int (\vec{S}' \wedge \text{grad}' \mathbf{\Lambda} (\vec{S}'' \text{grad}'' \varphi)) \psi \, d\vec{v}' \, d\vec{v}'', \quad \text{by (A 15)} \\ & = - \int \vec{S}' \wedge \vec{S}'' (\kappa^2 \psi - \varphi) \, d\vec{v}' \, d\vec{v}'' - \frac{1}{2} \int (\vec{S}' \wedge \text{grad}' \mathbf{\Lambda} [\vec{S}'' \cdot (\vec{x}' - \vec{x}'')]) \varphi \, d\vec{v}' \, d\vec{v}'', \quad \text{by (A 17) and (A 16)} \\ & = - \int \vec{S}' \wedge \vec{S}'' (\kappa^2 \psi - \frac{\varphi}{2}) \, d\vec{v}' \, d\vec{v}'' - \frac{1}{2} \int (\vec{S}' \wedge \text{grad}' \varphi) \mathbf{\Lambda} [\vec{S}'' \cdot (\vec{x}' - \vec{x}'')] \, d\vec{v}' \, d\vec{v}'' \\ & = \int (\vec{S}' \wedge \vec{S}''') (\frac{\varphi}{2} - \kappa^2 \psi) \, d\vec{v}' \, d\vec{v}'' + \frac{1}{2} \int [\vec{S}' \wedge (\vec{x}' - \vec{x}'')] \mathbf{\Lambda} (\vec{S}'' \text{grad}'' \varphi) \, d\vec{v}' \, d\vec{v}'', \end{aligned}$$

since $\text{grad}' \varphi (\vec{x}' - \vec{x}'')$ is proportional to $\vec{x}' - \vec{x}''$; altogether, we thus get

$$\left. \begin{aligned} \int \vec{x} \wedge (\vec{G}^\circ \wedge \vec{T}^\circ) \, d\vec{v} &= \int (\vec{S}' \wedge \vec{S}''') (\frac{\varphi}{2} - \kappa^2 \psi) \, d\vec{v}' \, d\vec{v}'' - \frac{\kappa^2}{2} \int (\vec{S}' \wedge \vec{S}''') (\vec{x}' \wedge \vec{x}'') \varphi \, d\vec{v}' \, d\vec{v}'' \\ & \quad + \frac{1}{2} \int (\vec{x}' \wedge \vec{x}'') (\vec{S}'' \text{grad}'' \mathbf{\Lambda} (\vec{S}' \text{grad}' \varphi)) \, d\vec{v}' \, d\vec{v}''. \end{aligned} \right\} \text{(A 27)}$$

From (A 26), (A 27), and taking account of (A 15), we obtain for the contribution of the vector meson field alone to $\vec{M}_{(0)}$

$$\left. \begin{aligned} \vec{M}_{(0)}^{\text{vect}} &= \frac{e}{4\hbar c} \int \left\{ \vec{S}' \wedge \vec{S}'' \right\}_{\mathbf{3}} (1 - \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv'' \\ &- \frac{e}{4\hbar c} \int \left\{ \vec{N}' \wedge \vec{N}'' + \kappa^2 \vec{S}' \wedge \vec{S}'' \right\}_{\mathbf{3}} (\vec{x}' \wedge \vec{x}'') \varphi dv' dv'' \\ &+ \frac{e}{4\hbar c} \int (\vec{x}' \wedge \vec{x}'') \left\{ (\vec{S}'' \text{ grad}'') \wedge (\vec{S}' \text{ grad}') \right\}_{\mathbf{3}} \varphi dv' dv''; \end{aligned} \right\} \text{(A 28)}$$

owing to the factors $\vec{x}' \wedge \vec{x}''$, the last two terms are of the exchange type only, while the first also involves a proper magnetic moment given by (62).

Finally, we derive from (A 22), for the pseudoscalar field,

$$\begin{aligned} &\int \vec{x} \wedge (\mathbf{I}^{\circ} \wedge \Psi^{\circ}) dv \\ &= \int (\vec{x}'' \wedge \vec{P}') \wedge (\vec{P}' \text{ grad}' \varphi) dv' dv'' \\ &- \frac{1}{2} \int (\vec{P}' \text{ grad}'') \wedge (\vec{P}' \text{ grad}') (\vec{x}' \wedge \vec{x}'') \varphi dv' dv'' \\ &= - \int [(\vec{x}' - \vec{x}'') \wedge \vec{P}'] \wedge (\vec{P}' \text{ grad}' \varphi) dv' dv'' + \frac{1}{2} \int \vec{P}' \wedge \vec{P}' \varphi dv' dv'' \\ &- \frac{1}{2} \int (\vec{x}' \wedge \vec{x}'') (\vec{P}' \text{ grad}'') \wedge (\vec{P}' \text{ grad}') \varphi dv' dv''. \end{aligned}$$

With the notation $\vec{x}_0 \equiv \frac{\vec{x}' - \vec{x}''}{|\vec{x}' - \vec{x}''|}$, the first term may be written

$$\begin{aligned} &\frac{1}{2} \int [(\vec{x}_0 \wedge \vec{P}') \wedge (\vec{P}' \vec{x}_0) + (\vec{x}_0 \wedge \vec{P}') \wedge (\vec{P}'' \vec{x}_0)] (1 + \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv'' \\ &= \frac{1}{2} \int [(\vec{P}' \vec{x}_0) \wedge \vec{P}' - \vec{P}' \wedge (\vec{P}'' \vec{x}_0)] \wedge \vec{x}_0 (1 + \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv''; \end{aligned}$$

with the help of the vector relation $(\vec{a} \wedge \vec{b}) \wedge \vec{c} = (\vec{a} \vec{c}) \vec{b} - \vec{a} (\vec{b} \vec{c})$, which holds when \vec{b} and \vec{c} are commutable, this transforms into

$$\frac{1}{2} \int [(\vec{P}' \wedge \vec{P}') \wedge \vec{x}_0] \wedge \vec{x}_0 (1 + \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv'',$$

or, using again the same relation,

$$-\frac{1}{2} \int \left\{ \vec{\mathbf{P}}' \wedge \vec{\mathbf{P}}'' - \left[(\vec{\mathbf{P}}' \wedge \vec{\mathbf{P}}'') \vec{x}_0 \right] \vec{x}_0 \right\} (1 + \kappa |\vec{x}' - \vec{x}''|) \varphi dv' dv''.$$

Hence

$$\begin{aligned} \vec{M}_{(0)}^{\text{ps-scal}} = & -\frac{e}{4\hbar c} \int \left\{ \left\{ \vec{\mathbf{P}}' \wedge \vec{\mathbf{P}}'' \right\}_{\mathbf{3}} \kappa |\vec{x}' - \vec{x}''| \right. \\ & \left. - \left[\left\{ \vec{\mathbf{P}}' \wedge \vec{\mathbf{P}}'' \right\}_{\mathbf{3}} \vec{x}_0 \right] \vec{x}_0 (1 + \kappa |\vec{x}' - \vec{x}''|) \right\} \varphi dv' dv'' \\ & - \frac{e}{4\hbar c} \int (\vec{x}' \wedge \vec{x}'') \left\{ (\vec{\mathbf{P}}'' \text{ grad}'') \wedge (\vec{\mathbf{P}}' \text{ grad}') \right\}_{\mathbf{3}} \varphi dv' dv''; \end{aligned} \quad \left. \vphantom{\vec{M}_{(0)}^{\text{ps-scal}}} \right\} \quad (\text{A } 29)$$

here also the first integral only contributes to the proper magnetic moment, *viz.* the moment (63).

It now again appears that, to the first order in the nuclear velocities, the last term in (A 29) just compensates the corresponding term of (A 28), so that, using NF (3, 4, 24) and the notation (59), and taking into account that $f_2^2 = g_2^2$, we find for the exchange moment in the mixed theory the expression (60). It will be noticed that the exchange moment is not translation invariant; if the origin is displaced by \vec{a} , the exchange moment changes by $\frac{1}{2} \vec{a} \wedge \vec{J}$.

Expression of the proper magnetic moments of nucleons for special source distributions.

We shall here briefly indicate how the formulae (66), (67), (68) for the "Landé factor" μ of the proper magnetic moment may be evaluated under simple assumptions over the source density function $D(r)$. The integrals occurring in these formulae are of the form

$$\mathcal{F} = \int D(r') dv' \int D(r'') dv'' \frac{F(\kappa \varrho)}{4\pi\varrho},$$

with

$$\varrho = \sqrt{r'^2 + r''^2 - 2r'r''u},$$

u denoting the cosine of the angle between \vec{x}' and \vec{x}'' . Introducing suitable polar coordinates, we obtain

$$\mathcal{F} = \int \frac{D(r') dv'}{2\kappa r'} \int D(r'') r'' dr'' \int_{|\kappa|}^{\kappa(r'+r'')} F(y) dy. \quad (\text{A } 30)$$

For the contributions from the vector and pseudoscalar meson field, respectively, we have

$$F^{\text{vect}}(y) = (1 - y) e^{-y}, \quad F^{\text{ps-scal}}(y) = (1 - 2y) e^{-y}. \quad (\text{A } 31)$$

Consider a uniform source distribution on the surface of a sphere of radius r_0 ; the corresponding density function may be expressed by

$$D(r) = \frac{\delta(r - r_0)}{4 \pi r_0^2} \quad (\text{A } 32)$$

and (A 30) then reduces to

$$\mathcal{F}^{\text{sup}} = \frac{1}{8 \pi r_0^2 \kappa} \int_0^{2\kappa r_0} F(y) dy. \quad (\text{A } 33)$$

In this way, putting $\varepsilon \equiv \kappa r_0$, we obtain

$$\left. \begin{aligned} \mu^{\text{vect}} &= \frac{g_2^2}{4 \pi \hbar c} \cdot \frac{M}{M_m} \cdot \frac{2}{\varepsilon} e^{-2\varepsilon} \\ \mu^{\text{ps-scal}} &= \frac{f_2^2}{4 \pi \hbar c} \cdot \frac{M}{M_m} \cdot \frac{1}{3 \varepsilon^2} [e^{-2\varepsilon} (1 + 4\varepsilon) - 1] \\ \mu &= \frac{g_2^2}{4 \pi \hbar c} \cdot \frac{M}{M_m} \cdot \frac{1}{3 \varepsilon^2} [e^{-2\varepsilon} (1 + 10\varepsilon) - 1]. \end{aligned} \right\} \quad (\text{A } 34)$$

For small values of ε , the functions of ε occurring in (A 34) approximately become

$$\frac{2}{\varepsilon} - 4, \quad \frac{2}{3\varepsilon} - 2, \quad \frac{8}{3\varepsilon} - 6,$$

respectively.

In the case of a uniform volume distribution of the sources within a sphere of radius r_0 , we get from (A 30), (A 31), by a straightforward calculation,

$$\left. \begin{aligned} \mu^{\text{vect}} &= \frac{g_2^2}{4 \pi \hbar c} \frac{M}{M_m} \cdot \frac{18}{\varepsilon^6} \left[-\frac{\varepsilon^3}{3} + \varepsilon^2 - 2 + e^{-2\varepsilon} (\varepsilon + 1) (\varepsilon^2 + 2\varepsilon + 2) \right] \\ \mu^{\text{ps-scal}} &= \frac{f_2^2}{4 \pi \hbar c} \frac{M}{M_m} \cdot \frac{3}{\varepsilon^6} \left[-2\varepsilon^3 + 5\varepsilon^2 - 9 + e^{-2\varepsilon} (\varepsilon + 1) (4\varepsilon^2 + 9\varepsilon + 9) \right] \\ \mu &= \frac{g_2^2}{4 \pi \hbar c} \frac{M}{M_m} \cdot \frac{3}{\varepsilon^6} \left[-4\varepsilon^3 + 11\varepsilon^2 - 21 + e^{-2\varepsilon} (\varepsilon + 1) (10\varepsilon^2 + 21\varepsilon + 21) \right] \end{aligned} \right\} \quad (\text{A } 35)$$

For small values of ε , the approximate values of the functions of ε occurring in (A 35) are

$$\frac{36}{5\varepsilon}, \quad \frac{22}{5\varepsilon}, \quad \frac{58}{5\varepsilon},$$

respectively.

The formulae (A 34), (A 35) have been obtained by SERPE in another way; for a numerical discussion, the reader may be referred to SERPE's paper.

Correction:

p. 6, 29th line, 'it is out of question' read 'it is excluded'.

References.

1. C. MØLLER and L. ROSENFELD, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* **XVII**, 8, 1940.
2. A short re-statement of the theoretical arguments leading to the mixed theory, with a discussion of the relevant experimental evidence, is contained in L. ROSENFELD, *Proc. Amsterdam* **45**, 155, 1942. On the question of the life-time of the mesons see further J. JUILFS, *Naturw.* **30**, 584, 1942; G. COCCONI and V. TONGIORGI, *Naturw.* **31**, 108, 1943; J. JUILFS, *Naturw.* **31**, 109, 1943 and W. OPECHOWSKI, *Physica*, in the press.
3. S. ROZENTAL, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* **XVIII**, 7, 1941; *Phys. Rev.* **60**, 612, 1941.
See also T. CHANG, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* **XIX**, 10, 1942; S. SAKATA, *Proc. phys.-math. Soc. Japan* **23**, 291, 1941.
4. C. MØLLER, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.* **XVIII**, 6, 1941.
5. J. LUBAŃSKI and L. ROSENFELD, *Physica* **9**, 117, 1942.
For a summary of the whole discussion and further references, see L. ROSENFELD, *Proc. Amsterdam* **45**, 155, 1942.
See also N. SVARTHOLM, *Fysiogr. Sällsk. Lund Förhandl.* **12**, nr. 9, 1942.
6. A. PAIS, *Physica* **8**, 1137, 1941; **9**, 267, 1942; **9**, 407, 1942.
7. A. PAIS, *D. Kgl. Danske Vidensk. Selskab, Mat.-fys. Medd.*, **XX**, 17, 1943.
8. See Y. TANIKAWA and H. YUKAWA, *Proc. phys.-math. Soc. Japan* **23**, 445, 1941; H. YUKAWA, *ZS. f. Phys.* **119**, 201, 1942.
9. G. RACAH, *Rendic. Lincei* **25**, 223, 1937.
10. H. HELMHOLTZ, *Journ. f. Math.* **100**, 165, 1886 (*Werke* III, 236); A. MAYER, *Leipziger Ber.* **40**, 519, 1896.
11. W. PAULI, *Rev. Mod. Phys.* **13**, 203, 1941.
Cf. also F. BELINFANTE, *Physica* **7**, 449, 1940.
12. L. ROSENFELD, *Mém. Acad. roy. Belg.* **18**, no. 6, 1940.
Cf. also F. BELINFANTE, *Physica* **7**, 449, 1940.
13. The method here followed is analogous to that of W. HEISENBERG and W. PAULI, *ZS. f. Phys.* **59**, 168, 1930 and J. OPPENHEIMER, *Phys. Rev.* **35**, 461, 1930 (cf. L. ROSENFELD, *Ann. Inst. Poincaré* **2**, p. 63—65, 1932). For a similar treatment of Fermi's method, see F. BELINFANTE, *Leiden thesis* 1939.
14. E. MAJORANA, *Nuovo Cimento* **14**, 171, 1937.

15. W. PAULI and V. WEISSKOPF, *Helvet. phys. acta* **7**, 709, 1934.
 16. H. FRÖHLICH, W. HEITLER and B. KAHN, *Proc. Roy. Soc. A* **171**, 269, 1939. A correction in SERPE's paper.
 17. W. LAMB, *Phys. Rev.* **56**, 384, 1939.
H. FRÖHLICH, W. HEITLER and B. KAHN, *Phys. Rev.* **56**, 961, 1939.
W. LAMB, *Phys. Rev.* **57**, 458, 1940.
Cf. also B. KAHN, *Physica* **8**, 58, 1941; A. SOMMERFELD, *ZS. f. Phys.* **118**, 295, 1941.
 18. H. CASIMIR, *Teyler's Verhandelingen* no. 11, 1936.
 19. A. SIEGERT, *Phys. Rev.* **52**, 787, 1937.
Cf. W. LAMB and L. SCHIFF, *Phys. Rev.* **53**, 651, 1938.
 20. W. LAMB and L. SCHIFF, *Phys. Rev.* **53**, 651, 1938.
 21. A. PROCA, *Journ. de phys.* **7**, 347, 1936.
 22. N. KEMMER, *Proc. Roy. Soc. A* **173**, 91, 1939.
 23. Cf. H. YUKAWA, S. SAKATA, M. KOBAYASI and M. TAKETANI, *Proc. phys.-math. Soc. Japan* **20**, 720, 1938, § 9.
 24. Cf. H. YUKAWA, S. SAKATA and M. TAKETANI, *Proc. phys.-math. Soc. Japan* **20**, 319, 1938, § 5 (also earlier references).
H. FRÖHLICH, W. HEITLER and N. KEMMER, *Proc. Roy. Soc. A* **166**, 154, 1938, § 4; N. KEMMER, *Proc. Roy. Soc. A* **166**, 127, 1938; a more correct calculation in SERPE's paper.
 25. S. MILLMAN and P. KUSCH, *Phys. Rev.* **60**, 91, 1941.
 26. L. ALVAREZ and F. BLOCH, *Phys. Rev.* **57**, 111, 1940.
 27. J. KELLOGG, I. RABI, N. RAMSEY and J. ZACHARIAS, *Phys. Rev.* **57**, 677, 1940.
A. NORDSIECK, *Phys. Rev.* **58**, 310, 1940.
See also H. CASIMIR, *Physica* **7**, 169, 1940.
 28. H. MARGENAU, *Phys. Rev.* **57**, 383, 1940.
-

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 13

L'ACTION DE LA TRIÉTHYLAMINE
SUR L' α -BROMOPROPIONATE
D'ÉTHYLE

PAR

EINAR BIILMANN ET EVALD EGELUND PEDERSEN



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

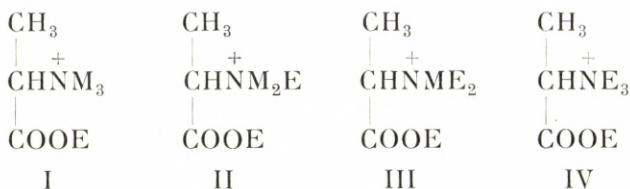
Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

Dans des mémoires précédents¹ il fut montré que les formes optiquement actives d'éthers-sels des propiobétaïnes comme aussi les éthers-sels analogues des phosphétines et des arsénétines, tous ayant un atome H en position α , subissent une racémisation par l'action d'une amine tertiaire ou de l'éthylate de sodium. De plus, il fut trouvé qu'une forme active de phényl-p-tolylacetate d'éthyle fut, elle aussi, racémisée par l'action de l'éthylate de sodium, tandis qu'elle ne l'est point par l'action de la triéthylamine, évidemment parce que l'amine a un effet moins basique que l'éthylate de sodium. Dans les recherches que nous allons décrire ci-dessous, nous avons étudié si une forme active d' α -bromopropionate d'éthyle, contenant comme les composés mentionnés ci-dessus un atome H dans la position α , est racémisée dans des conditions analogues. Quand on a affaire à de l' α -bromopropionate d'éthyle, on ne peut pas employer l'éthylate de sodium comme agent racémisant, parce qu'il donnerait naissance à de l' α -éthoxy-propionate d'éthyle. Par contre, il nous semblait possible que l'on pût appliquer de la triéthylamine, qui selon toute probabilité ne compliquerait pas la recherche par la formation d'un propiobétaïne triéthylé. Car on a trouvé que, bien que l'on puisse préparer sans difficulté les propiobétaïnes, qui contiennent soit trois groupements de méthyle (formule I), soit deux groupements de méthyle et un groupement d'éthyle (formule II), soit un groupement de méthyle et deux groupements d'éthyle (formule III), on n'a pas réussi à préparer un dérivé triéthylé qui corresponde à la formule IV². Les éthers-sels furent isolés à l'état de leurs ha-

¹ Bull. soc. chim. de France [5], 1, 1645, 1653 et 1661 (1934); *ibid.* [5], 3, 2295 et 2306 (1936).

² Bull. soc. chim. de France [5], 1, 1653 (1934).

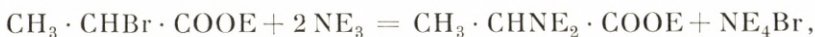
logenures, et les formules indiquent les cations correspondants. M représente le groupement méthyle, E le groupement éthyle. Voici les formules:



Certes, E. DUVILLIER¹ indique que, par l'action de la triéthylamine sur de l' α -bromopropionate d'éthyle, il a obtenu une très petite quantité d'un composé qui, après saponification, fut transformé en bromure de triéthyl-propionbétaine, et qui fut identifié par l'analyse de son chloroplatinate. Mais, dans la partie expérimentale (p. 10), nous montrerons qu'il est vraisemblable que la préparation analysée par DUVILLIER fut le chloroplatinate de l' α -diéthylaminopropionate d'éthyle. Les deux composés dont il s'agit sont isomères, de façon qu'une analyse n'est pas décisive. Nous citons aussi une recherche de J. W. BRÜHL², qui n'a pas réussi à obtenir un dérivé triéthylé de propionbétaine par une action très forte de triéthylamine sur de l' α -chloropropionate d'éthyle.

Mais, il s'est montré qu'il y a d'autres réactions, qui compliquent la recherche et qu'il faut envisager. Aussi avons-nous étudié à fond la réaction entre la triéthylamine et l' α -bromopropionate d'éthyle en chauffant un mélange de ces deux corps et à l'état pur et en solution toluénique. Nous avons trouvé que dans les deux cas il se forme un mélange liquide et un précipité cristallin, qui contient des bromures de tétraéthylammonium et de triéthylammonium.

La formation du bromure de tétraéthylammonium doit avoir lieu conformément au schéma:



ce qui exige une formation d' α -diéthylaminopropionate d'éthyle, et nous avons constaté la formation de ce composé. Par suite du schéma, le rapport moléculaire des deux produits de la réaction

¹ Bull. soc. chim. de France [3], 2, 141 (1889).

² Ber. deutsch. chem. Ges. 9, 35 (1876).

tion doit être 1,00, et nous avons trouvé 1,09. Vu les circonstances de la détermination des quantités des deux composés, l'accord peut être considéré comme satisfaisant.

La formation du bromure de triéthylammonium fait soupçonner un dédoublement de l' α -bromopropionate d'éthyle en acrylate d'éthyle et HBr conformément au schéma suivant:



Nous n'avons pas réussi à isoler de l'acrylate d'éthyle, ce qui est bien naturel, parce que ce composé se polymérise assez vite, et, son point d'ébullition (100°) étant plus élevé que le point d'ébullition de la triéthylamine (89°), nous n'avons pas pu isoler l'acrylate d'éthyle à mesure qu'il se forme. Nous n'avons pas non plus pu imaginer un mode opératoire pour le saisir sous forme de quelque dérivé. Mais en partant de l' α -bromopropionate de méthyle, nous avons réussi à isoler une très petite quantité d'acrylate de méthyle, dont le point d'ébullition n'est que 80° . Il fut transformé en l'anilide de l'acide β -anilinopropionique:



et ce corps fut identifié par son point de fusion comme aussi par le point de fusion d'un mélange de la préparation avec un échantillon de l'anilide de l'acide β -anilinopropionique obtenu par chauffage de l'acide β -iodopropionique avec de l'aniline.

Pourvu que nous admettions que les deux parties les moins volatiles formées par la réaction entre la triéthylamine et l' α -bromopropionate d'éthyle consistent en des composés formés par polymérisation de l'acrylate d'éthyle, nous pouvons tellement rendre compte des quantités de tous les composés employés ou formés par la réaction, que cela prouve que d'autres réactions n'ont pu avoir lieu qu'à un degré négligeable. En partant de 0,055 mol d' α -bromopropionate d'éthyle on a récupéré 0,009 mol de cet éther-sel et, en plus de cela, il s'est formé 0,012 mol d' α -diéthylaminopropionate d'éthyle et 0,032 mol d'acrylate d'éthyle, soit au total 0,053 mol au lieu des 0,055 mol dont on était parti. Quant au brome, nous retrouvons, naturellement, dans l' α -bromopropionate d'éthyle récupéré 0,009 mol, dans le bromure de tétraéthylammonium 0,011 mol et dans le bromure

de triéthylammonium 0,033 mol, soit de nouveau un total de 0,053 mol. Bien entendu, c'est un hasard que ces valeurs s'accordent si bien les unes avec les autres, mais après tout, les quantités trouvées rendent compte de 96,4 % du total.

Dans tous les essais que nous venons de décrire ci-dessus, nous avons employé la forme racémique d' α -bromopropionate d'éthyle. Cela a suffi pour nous faire connaître la plupart des réactions qui s'effectuent dans ces circonstances. Mais les essais n'élucident pas si une racémisation a eu lieu. En vue d'étudier cette question il faut que l'on emploie une forme active de l'éther-sel. Aussi avons-nous fait agir de la triéthylamine sur le (—)- α -bromopropionate d'éthyle. Afin de pouvoir déterminer le pouvoir rotatoire pendant l'essai, nous avons dissous les deux réactifs dans le toluène. Nous avons trouvé qu'à 50° la vitesse de la réaction est convenable pour l'exécution des mesurages, tandis que la réaction est trop vite à 90° et trop lente à 20°. Et nous nous sommes servis d'une pratique qui nous a permis de mesurer aux mêmes moments les pouvoirs rotatoires et les quantités de brome dans les mélanges des bromures de triéthylammonium et de tétraéthylammonium, qui se précipitent en état cristallin au cours du chauffage. Nous avons observé un décroissement du pouvoir rotatoire pendant le chauffage, voir les tableaux I—III de la partie expérimentale (p. 12).

Le décroissement des valeurs du pouvoir rotatoire fait voir qu'une racémisation de l'éther-sel a eu lieu, car les transformations de l'éther-sel en acrylate d'éthyle et en diéthylaminopropionate d'éthyle, dont la formation d'ion brome nous donne le pourcentage, n'embrassent que quelques rares pour-cent, quand la rotation est déjà près de zéro. Et nous avons observé le fait remarquable que l'on peut pousser le changement du pouvoir rotatoire jusqu'à un inversement du sens de la rotation. Les valeurs faiblement dextrogyres dans les tableaux II et III ont une signification réelle. C'est la formation du diéthylaminopropionate d'éthyle, qui est responsable de ce faible pouvoir dextrogyre. Dans la partie expérimentale nous montrerons qu'une préparation active d' α -diéthylaminopropionate d'éthyle n'est pas racémisée par l'amine tertiaire. Elle conserve toute son activité.

Les constantes de vitesse dans les tableaux I—IV sont calculées pour une réaction du premier ordre suivant l'équation:

$$K = 2,3026 \cdot \frac{1}{t} \cdot \log_{10} \frac{\alpha_0}{\alpha_t}$$

Les valeurs de K dans les tableaux I et II s'accordent assez bien, quand on considère que les deux séries de mesurages ont été exécutées avec deux préparations distinctes de l'éther-sel actif. Mais les valeurs de K dans le tableau III prouvent que ce n'est pas l'amine qui est le catalyseur de la racémisation, la concentration de l'amine dans le tableau III étant le double de celle des tableaux I et II.

Il fallait donc trouver une autre cause de la racémisation, et il semblait naturel d'assigner le rôle de catalyseur aux ions brome. Certes, les bromures de triéthylammonium et de tétraéthylammonium précipités au cours de la réaction sont très peu solubles dans le liquide; d'autre part, cette circonstance nous permet de supposer que la concentration d'ion brome dans la solution devienne constante peu de temps après le commencement des essais. Aussi avons-nous saturé une solution toluénique d' α -bromopropionate d'éthyle avec les deux bromures et mesuré le décroissement du pouvoir rotatoire. Les résultats de cet essai se trouvent dans le tableau IV. Nous avons déterminé les concentrations d'ion brome, $C_I - C_{IV}$, des solutions employées dans les essais enregistrés dans les tableaux I—IV. Voici les résultats:

$$C_I = C_{II} = 1,1 \cdot 10^{-3}, \quad C_{III} = 1,3 \cdot 10^{-3}, \quad C_{IV} = 3,7 \cdot 10^{-4},$$

et, par conséquent:

$$K_I : C_I = 43, \quad K_{II} : C_{II} = 47, \quad K_{III} : C_{III} = 47, \quad K_{IV} : C_{IV} = 32,$$

$K_I - K_{IV}$ étant les constantes de vitesse indiquées dans les tableaux I—IV. Les valeurs des trois premières fractions semblent révéler une proportionnalité entre constante de vitesse et concentration en ion brome. La dernière valeur (32) ne s'accorde pas si bien avec les trois autres; mais, d'abord, les déterminations des concentrations assez infimes d'ion brome deviennent peu exactes, et, de plus, dans l'essai IV la solution ne contient pas du tout de l'amine, tandis que dans les essais I et II la concentration en amine est de 0,84 molaire (8,5 %) et dans l'essai III de 1,68 molaire (17 %). Les solvants ne sont donc point identiques. Aussi sommes-nous d'avis que l'agent racémisant

n'est pas l'amine tertiaire même. Mais, par des réactions décrites plus haut, elle donne naissance à la formation d'ions brome, lesquels sont seuls responsables de la racémisation observée.

Partie expérimentale.

1. Action de la triéthylamine sur l' α -bromopropionate d'éthyle racémique.

Un mélange de 10 g de l'éther-sel (0,055 mol) et 11 g de triéthylamine (0,110 mol) a été chauffé à 107° en tube scellé, pendant 96 heures. Dès le commencement du chauffage il s'est déposé un corps blanc cristallin, dont la quantité augmentait peu à peu.

Le précipité a été séparé par filtration et lavé avec de l'éther. Rendement 8,4 g. Il s'est trouvé être un mélange des bromures de triéthylammonium et de tétraéthylammonium. Pour déterminer quantitativement ces constituants, on a ajouté une solution aqueuse de NaOH à un échantillon pesé des bromures, on a soumis ce mélange à une distillation et on a recueilli les vapeurs de l'amine tertiaire dans de l'acide chlorhydrique. Puis, par titrage, on a trouvé que l'amine tertiaire correspond à 71,6 % (6,0 g) du mélange des bromures. Si l'on admet que le reste (2,4 g) soit du bromure de tétraéthylammonium, le précipité cristallin doit contenir 28,4 % de ce composé, d'où on calcule qu'il contiendra au total 42,16 % de brome. Par titrage avec AgNO_3 nous avons trouvé 42,33 %.

Le liquide séparé des cristaux a été soumis à une distillation fractionnée, qui a conduit aux fractions suivantes:

a. Sous la pression ordinaire on a recueilli, à 87° — 90° , 3,0 g. Cette fraction est de la triéthylamine en état presque pur, le point d'ébullition de cette amine étant 89° . Nous avons trouvé, par titrage, que la valeur d'équivalent était de 105, calculée pour $(\text{C}_2\text{H}_5)_3\text{N} : 101$.

Dans le vide, sous 11 mm, on a obtenu les fractions suivantes:

b. Au-dessous de 53° : 4 g. Comme la précédente, cette fraction était de la triéthylamine. Équivalent trouvé comme ci-dessus: 105, calculé 101.

c. A 53° — 63° : 2,2 g. C'était un mélange d' α -bromopropionate d'éthyle et d' α -diéthylaminopropionate d'éthyle. Un dosage du

brome (Stepanow) a montré que la fraction contenait 68,6 % d' α -bromopropionate d'éthyle (1,55 g = 0,008 mol). Un titrage avec de l'acide a donné que le mélange contenait 29,6 % d' α -diéthylaminopropionate d'éthyle (0,65 g = 0,004 mol).

d. A 63° — 66° : 1,6 g, consistant en 13,2 % d' α -bromopropionate d'éthyle (0,2 g = 0,001 mol) et en 86,6 % d' α -diéthylaminopropionate d'éthyle (1,4 g = 0,008 mol).

e. A 80° — 154° : 0,8 g.

f. Résidu: 2,4 g.

2. α -Diéthylaminopropionate d'éthyle.

Un mélange de 60 g α -bromopropionate d'éthyle (forme racémique), 68 g triéthylamine et 100 cm³ toluène fut chauffé à 50° pendant 18 jours et nuits. Une masse cristalline s'était déposée. On a séparé la partie liquide par décantation, et on y a ajouté de l'acide chlorhydrique dilué jusqu'à réaction acide. Puis on a ajouté 200 cm³ d'éther. La solution dans le mélange d'éther et de toluène a été séparée de la couche aqueuse, on a épuisé le mélange avec 100 cm³ d'éther afin d'enlever les dernières traces de bromopropionate d'éthyle. On a ajouté K₂CO₃ (solide) à la solution aqueuse jusqu'à réaction alcaline, on a épuisé la solution deux fois avec 100 cm³ d'éther et on a séché la solution éthérée. Puis on a chassé l'éther au bain-marie et on a distillé le résidu dans le vide. A 73° et sous 11 mm on a recueilli 3,0 g de diéthylaminopropionate d'éthyle.

Un titrage avec de l'acide a donné pour l'équivalent la valeur 174, calculée 173.

La préparation a été transformée en son chloroplatinate. Analyse:

	Trouvé	calculé
Pt %	25,97	25,81

Point de fusion $114^{\circ},0$ — $114^{\circ},2$. NOVAK¹ indique 114° — 116° .

3. Formation d'acrylate de méthyle.

Un mélange de 10,1 g triéthylamine (0,1 mol) et 16,7 g α -bromopropionate de méthyle fut chauffé à 88° dans un ballon, muni d'un réfrigérant à air ascendant et d'un récipient, qui était réfrigéré avec de la glace et barré avec un coupe-air à mercure

¹ Ber. deutsch. chem. Ges. **45**, 842 (1912).

pour empêcher la diffusion des vapeurs de l'éther-sel cherché. Après 21 heures la quantité du distillat n'était que de 0,02 g, et elle ne s'est pas augmentée quand on a chauffé le ballon à 103° pendant deux jours et nuits. En élevant la température on a observé une faible distillation; quand la température du bain était de 155°, la température de la vapeur dans le ballon était de 72°. Après 30 minutes, pendant lesquelles la température s'est élevé à 190°, du bromure d'amine a commencé à sublimer. Alors on a terminé le chauffage. Dans le récipient se trouvait maintenant 0,48 g d'un liquide dont la réaction au papier de tournesol était faiblement alcaline. Pour l'identification on a transformé le corps dans l'anilide de l'acide β -anilino-propionique: Un mélange de 0,20 g du corps et de 0,54 g d'aniline fut chauffé dans un tube scellé à 180°—190° pendant 4 heures. Un corps brun sirupeux s'était formé. Il a été dissous dans 2 cm³ d'acide chlorhydrique de 20 % et refroidi avec de la glace. Une huile s'est séparée, mais l'hydrochlorure n'a pas cristallisé spontanément. Aussi a-t-on préparé un échantillon de l'hydrochlorure de β -anilinopropioanilide en partant d'acide β -iodopropionique pur, d'après W. AUTENRIETH et C. PRETZELL¹. A l'aide des cristaux de cette préparation on a amorcé la cristallisation de l'huile, et on a obtenu 0,19 g de l'hydrochlorure, que l'on a décomposé avec de Na₂CO₃. Ainsi on a pu obtenir 0,14 g de la base libre sous forme d'une poudre grisâtre. On l'a purifiée par cristallisation dans de la ligroïne chaude. Rendement 0,09 g. Point de fusion 89°,6—91°,8. AUTENRIETH et PRETZELL indiquent 91°—93°. La préparation obtenue par nous en partant d'acide β -iodopropionique fondait à 90°,2—92°,1 et un mélange en parties égales de nos deux préparations fondait à 90°,0—92°,2. Les deux corps sont donc identiques. De plus, nous avons analysé la préparation obtenue en partant de l'acrylate de méthyle. Trouvé dans deux analyses (micro-Kjeldahl): 11,53 % N et 11,53 % N, calculé 11,66 % N.

4. L'essai de DUVILLIER.

On a saponifié 1,0 g de l' α -diéthylaminopropionate d'éthyle par chauffage à ébullition pendant 5 heures avec 21,00 cm³ 0,2800-n Ba(OH)₂. Puis on y a ajouté la quantité calculée de

¹ Ber. deutsch. chem. Ges. **36**, 1264 (1903).

1-n H_2SO_4 . On a filtré et évaporé la solution filtrée au bain-marie jusqu'à cristallisation. Puis on a ajouté 1,5 g $\text{H}_2\text{PtCl}_6, 6 \text{H}_2\text{O}$ et 2 cm³ 3,36-n HCl. On a concentré la solution et on a ajouté 20 cm³ d'alcool. Alors on a chassé de nouveau le solvant au bain-marie et séché le résidu dans un dessiccateur sur du CaCl_2 . Puis on a amorcé la cristallisation du chloroplatinate de l' α -diéthylaminopropionate d'éthyle en ajoutant une trace de ce corps, on a additionné 10 cm³ d'alcool absolu. On a purifié le déposé par cristallisation dans 3 cm³ d'alcool absolu. Rendement 0,47 g. P. f. 114°,0—114°,3. Un mélange de cette préparation avec une quantité égale de chloroplatinate d' α -diéthylaminopropionate d'éthyle (p. f. 114°,0—114°,2) a montré un p. f. de 114°,0—114°,2.

Analyse: Pt, trouvé 25,84 %, calculé pour $\text{C}_{18}\text{H}_{40}\text{O}_4\text{N}_2\text{Cl}_6\text{Pt}$ 25,81 %.

Il nous semble donc naturel de croire que c'est ce composé que DUVILLIER a obtenu dans un essai et non pas son isomère, le chloroplatinate d' α -triéthylpropionbétaïne.

5. Action de la triéthylamine sur le (—)- α -bromopropionate d'éthyle en solution toluénique.

Les solutions des deux réactifs dans du toluène ont été chauffées à 50° dans des tubes scellés. Chaque tube contenait 5 cm³ de la solution étudiée. Les concentrations sont indiquées dans les tableaux I—III ci-dessous. A des moments appropriés on a plongé deux des tubes dans un mélange de glace et de sel, de façon à arrêter la réaction, et on les a laissés reposer dans le réfrigérant pendant une demi-heure. Ainsi on est arrivé à une solution tout à fait limpide, laquelle est aisément séparée de la masse cristalline qui s'est déposée, sans qu'une filtration soit nécessaire. L'un des deux tubes a servi à la détermination du pouvoir rotatoire. Le mélange dans l'autre tube fut employé pour le dosage de l'ion brome: On a ajouté 10 cm³ d'éther, 5 cm³ d'acide nitrique dilué et 5 cm³ d'eau; on a bien agité, afin d'obtenir une solution des bromures dans l'eau. On a séparé la couche aqueuse et on a déterminé l'ion brome par microtitrage.

Dans les tableaux I—III, M_E indique la concentration moléculaire en éther-sel et M_A la concentration moléculaire en triéthylamine de la solution toluénique, dont on était parti. Les temps sont indiqués en heures. Le premier mesurage fut exécuté après

deux heures ($t = 0$). α indique les valeurs du pouvoir rotatoire, mesurées à 20° dans un tube de 2 dm de longueur. Dans la dernière colonne des tableaux on trouve les quantités d'ion brome, en pour cent du brome total. K représente la constante de vitesse du décroissement de la rotation, calculée pour une réaction du premier ordre (cfr. p. 7).

Tableau I.

 $M_E = 0,84, M_A = 0,84.$

Solvant: toluène.

t	α	$K \cdot 10^2$	Br ⁻
0	-4,18	—	0,62
2	-3,83	4,4	0,89
4	-3,42	5,0	1,13
6	-3,11	4,9	1,30
8	-2,90	4,6	1,47
12	-2,36	4,8	1,80
20	-1,59	—	2,39
34	-0,53	—	3,28
49	-0,14	—	4,20

Moyenne: $K_I = 4,7 \cdot 10^{-2}$.

Tableau II.

 $M_E = 0,84, M_A = 0,84.$

Solvant: toluène.

t	α	$K \cdot 10^2$	Br ⁻
0	-7,21	—	0,71
7	-5,04	5,1	1,40
16,8	-3,02	5,2	2,12
22	-2,30	5,2	2,45
24,5	-1,95	5,3	—
30	-1,28	—	3,12
43	-0,34	—	—
72	+0,02	—	5,21
96	+0,08	—	6,07
120	+0,06	—	6,82
168	+0,07	—	—

Moyenne: $K_{II} = 5,2 \cdot 10^{-2}$.Moyenne des deux constantes: $K = 5,0 \cdot 10^{-2}$.

Tableau III.

 $M_E = 0,84, M_A = 1,68.$

Solvant: toluène.

t	α	$K \cdot 10^2$	Br ⁻
0	-4,65	—	1,23
2	-4,09	6,4	1,61
4	-3,61	6,3	1,90
8	-2,91	5,9	2,77
20	-1,48	5,7	5,07
34	-0,41	—	7,35
45	-0,15	—	8,90
96	+0,12	—	—
144	+0,09	—	—
336	+0,11	—	—

Moyenne: $K_{III} = 6,1 \cdot 10^{-2}$.

Comme nous l'avons dit, les valeurs faiblement dextrogyres dans les tableaux II et III ont une signification réelle. C'est la formation du diéthylaminopropionate d'éthyle qui est responsable du faible pouvoir dextrogyre. Afin de le prouver, nous sommes partis des deux tubes dont la valeur d' α dans 2 dm tube était de $-0^{\circ},34$ (tableau II). Au liquide de ces deux tubes on a ajouté 10 cm^3 d'acide chlorhydrique dilué de façon à épuiser le diéthylaminopropionate d'éthyle, et on a établi une séparation complète des deux couches à l'aide du centrifugeur. Puis on a mesuré les pouvoirs rotatoires. Pour la solution en toluène la rotation s'était augmentée de $-0^{\circ},34$ à $-0^{\circ},39$ malgré le fait que la solution dans l'acide chlorhydrique était, elle aussi, lévogyre, sa valeur d' α étant de $-0^{\circ},02$. L'explication très simple de ce phénomène est que le diéthylaminopropionate d'éthyle est dextrogyre, tandis que sa solution dans l'acide chlorhydrique est lévogyre.

Puis, en vue de confirmer la justesse de cette explication, nous avons employé une préparation dextrogyre de diéthylaminopropionate d'éthyle. La forme racémique de ce composé fut obtenue par BILMANN et BERG¹ par l'action de diéthylamine sur l' α -bromopropionate d'éthyle à 50° – 60° . Pour l'obtention d'une forme active nous sommes partis d'une préparation lévogyre d' α -bromopropionate d'éthyle ($\alpha = -5^{\circ},48$ dans un 1 dm tube). On a laissé reposer à la température ordinaire pendant 10 jours et nuits un mélange de 18 g de l'éther-sel et 25 g de diéthylamine. Du bromure de diéthylammonium s'est déposé au cours de l'essai. On a ajouté au mélange 50 cm^3 d'éther et on a filtré à la trompe. On a acidulé la solution par addition d'acide chlorhydrique dilué, on a séparé la couche aqueuse et on l'a épuisée de nouveau avec 50 cm^3 d'éther. Afin d'avoir le diéthylaminopropionate d'éthyle à l'état libre, on a ajouté à la solution aqueuse une quantité convenable de K_2CO_3 solide, puis on a épuisé la solution 3 fois par 100 cm^3 d'éther. Les extraits réunis, on les a séchés avec K_2CO_3 . Par distillation dans le vide (12 mm) on a recueilli entre 72° et 73° 14,1 g de diéthylaminopropionate d'éthyle, c'est-à-dire 82 % de la théorie. Équivalent par titrage avec de l'acide: Trouvé 173, calculé 173. $\alpha = +8^{\circ},57$ ($l = 1\text{ dm}$). Le pouvoir rotatoire d'une solution de 1 g de l'éther-

¹ Bull. soc. chim. de France [5], 1, 1657, (1934).

sel dans 10 cm³ de toluène était de + 1°,83 (l = 2 dm), tandis que la valeur d' α pour la même quantité d'éther-sel, dissoute dans 10 cm³ d'acide chlorhydrique dilué, était de -0°,25.

6. Racémisation de l' α -bromopropionate d'éthyle par l'action de l'ion brome.

A une série de tubes on a donné un mélange de NE₃HBr et de NE₄Br en état solide. Puis on a ajouté à chaque tube 5 cm³ d'une solution 0,84 molaire de (-)- α -bromopropionate d'éthyle en toluène, et on a scellé les tubes. Puis on les a agités dans un thermostat à 50°. A des instants convenables on a plongé les tubes les uns après les autres dans un mélange réfrigérant à -10° environ, et après déposition complète des bromures, qui étaient présents en grand excès, on a mesuré la rotation de la solution tout à fait claire. Les observations se trouvent dans le tableau IV, dans lequel le temps t = 0 correspond à 18 heures après l'introduction des tubes dans le thermostat, de façon que l'on s'est assuré une saturation complète de la solution toluénique avec les bromures d'amine dès le commencement des mesurages.

Tableau IV.
M_E = 0,84; Br⁻ = 3,7 · 10⁻⁴-n.
Solvant: toluène.

t	α	K · 10 ²
0	- 7,20	-
24	- 5,46	1,2
48	- 4,21	1,1
96	- 2,27	1,2
106,5	- 1,96	1,2
120	- 1,76	1,2
144	- 1,22	1,2

Moyenne: K_{IV} = 1,2 · 10⁻².

7. Détermination des concentrations d'ion brome dans des solutions correspondant aux tableaux I-IV.

On a préparé des solutions comme pour les essais résumés dans les tableaux I-IV, et on les a saturées à 50° avec un mélange de NE₃HBr et de NE₄Br (solides). Puis on a épuisé les

liquides claires 3 fois avec du H_2SO_4 dilué, on a séparé la couche aqueuse, et on a déterminé la concentration en ion brome par microtitrage avec du AgNO_3 et du NH_4SCN . Les titrages ont été exécutés avec environ $0,05 \text{ cm}^3 \text{ AgNO}_3$ 0,1-n en titrant jusqu'à un excès de NH_4SCN . Ces déterminations donnent naissance à une erreur systématique. Voilà pourquoi on a déterminé la grandeur de cette erreur par titrage d'une solution de KBr , dont la concentration en ion brome avait une grandeur connue et du même ordre que celle des solutions étudiées. On a trouvé que l'erreur correspond à $0,006 \text{ cm}^3$, qu'il faut ajouter à la quantité de AgNO_3 employée. Avec cette correction on a obtenu les concentrations en ion brome indiquées ci-dessous:

$$C_I = C_{II} = 1,1 \cdot 10^{-3}; \quad C_{III} = 1,3 \cdot 10^{-3}; \quad C_{IV} = 3,7 \cdot 10^{-4}.$$

8. 1- α -Diéthylaminopropionate d'éthyle et triéthylamine.

Dans la partie générale nous avons dit que l' α -diéthylaminopropionate d'éthyle n'est pas racémisé par l'action de NE_3 . Nous l'avons constaté de la manière suivante: Une solution 1,0 molaire tant de l' α -diéthylaminopropionate d'éthyle que de NE_3 dans du toluène a été tenue en thermostat à 50° pendant 12 jours et nuits. La rotation du mélange au commencement de l'essai était $\alpha = -8^\circ,02$ dans 2 dm tube et n'a pas changé pendant l'essai. Nous ajoutons que le pouvoir rotatoire de l'éther-sel pur que nous avons employé était $\alpha = -20^\circ,08$ pour 1 dm tube.

9. Dédoublément de l'acide α -bromopropionique en ses composants actifs.

Nous avons exécuté le dédoublément de l'acide α -bromopropionique par cristallisation de son sel normal cinchoninique dans de l'eau comme indiqué par L. RAMBERG¹, et, en opérant suivant ses indications, nous avons fait des observations, qui nous semblent d'un certain intérêt et que nous allons décrire:

On a dissous 255 g $\text{CH}_3\text{CHBrCOOH}$ (1,67 mol) et 247 g cinchonine (0,84 mol) dans 5000 cm^3 d'eau à 35° . On a laissé la solution à la température ordinaire. Au cours de 12 heures 37 % du sel cinchoninique s'est déposé. On a récupéré l'acide

¹ Liebig's Ann. d. Chemie, **349**, 324 (1906).

et constaté qu'il était lévogyre. Par concentration fractionnée de l'eau mère dans le vide à 45°, on a obtenu d'autres précipités du sel cinchoninique, qui donnent, eux aussi, des préparations lévogyres de l'acide, tandis qu'une préparation dextrogyre de l'acide s'obtient de l'eau mère de ces précipitations quand on l'acidule avec de l'acide sulfurique et qu'on épuise le mélange avec de l'éther. Suivant RAMBERG, on doit chasser par évaporation une quantité d'eau, qui se monte à 3500 cm³. Par des essais de dédoublement réitérés nous avons appris qu'il vaut mieux terminer l'évaporation, dès que 2500 cm³ d'eau sont enlevés. En continuant l'évaporation, on arrive rapidement à une préparation dextrogyre de l'acide. Néanmoins nous n'avons pas essayé d'obtenir l'acide dextrogyre par son sel cinchoninique, mais nous avons séparé une préparation dextrogyre de l'acide en acidulant l'eau mère et en épuisant avec de l'éther comme indiqué par RAMBERG. Nos observations au cours de ces essais sont enregistrées dans le tableau ci-dessous.

	H ₂ O évaporée g	sel cincho- ninique g	acide g	α	acide lévogyre p. ct. du total
		Essai no 1			
	0	182	93	— 10°,87	16,1
	2700	131	67	— 4°,15	4,4
Total...	2700	313	160	— 8°,04	20,5
		Essai no 2			
	0	192	98	— 10°,81	16,8
	2460	105	53	— 5°,85	5,0
Total...	2460	297	151	— 9°,04	21,8
		Essai no 3			
	2330*	266	136	— 9°,40	20,3
	1170	59	30	+ 0°,97	— 0,5
Total...	3500	325	166	— 7°,52	19,8

* Sans filtration préalable.

Dans la première colonne se trouvent les quantités d'eau évaporées. La deuxième colonne contient les quantités précipitées du sel cinchoninique en grammes, et dans la troisième

colonne se trouvent les quantités d'acide des sels cinchoniques. Les pouvoirs rotatoires de ces acides sont indiqués dans la colonne suivante. A l'aide de ces valeurs on a calculé les quantités d'acide lévogyre qui sont présentes à côté du mélange racémique, et ces quantités se trouvent dans la dernière colonne, exprimées en pourcents de la quantité globale d'acide lévogyre dans l'acide racémique (255 g), dont on était parti. Pour ce calcul nous avons employé la valeur de $-49^{\circ},3$ comme pouvoir rotatoire de l'acide lévogyre pur (RAMBERG¹).

Des préparations actives obtenues comme indiqué ci-dessus furent purifiées par cristallisation comme indiqué par RAMBERG². Ainsi nous avons obtenu une préparation lévogyre dont la valeur d' α dans 1 dm tube était de $-32^{\circ},82$. Nous sommes déjà arrivés à cette purification en réfrigérant à $+10^{\circ}$. Comme cette activité a suffi tout à fait à nos buts, nous n'avons pas poussé la purification plus loin.

¹ Liebig's Ann. d. Chemie, **370**, 236 (1909).

² Ibid. **349**, 327 (1906).

*Universitetets kemiske Laboratorium,
Copenhagen.*

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 14

LES RECHERCHES SUR
LE MAGNÉTISME
TERRESTRE

PAR
DAN LA COUR



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A S

Le contenu de la présente étude a été communiqué à l'Académie par Dan la Cour dans sa séance du 27 février 1942. Le 10 mars suivant La Cour remit un manuscrit rédigé en danois, qui, après la mort de l'auteur, a été jugé bon à paraître en français dans les »Matematisk-fysiske Meddelelser« de l'Académie.

MARTIN KNUDSEN
Secrétaire de l'Académie.

Il y a cent ans, ce mois-ci [Février 1942], que furent commencées dans notre pays, sous les auspices de l'Académie Royale des Sciences et des Lettres, des observations régulières du magnétisme terrestre; s'il est vrai qu'on avait pris des observations magnétiques avant cette date et que d'autre part elles n'ont pas été faites sans interruptions depuis lors, il est juste toutefois de commémorer ce centenaire en rappelant brièvement l'effort que l'Académie a fait dans le temps pour entreprendre ce travail, et de donner quelques détails sur les résultats de cette initiative.

C'est, on le sait, l'Académie qui a donné le branle tant aux recherches sur le magnétisme terrestre qu'aux observations météorologiques, deux activités qui de nos jours ont été rattachées à l'Institut météorologique, et déjà bien des années avant 1842 elle s'était intéressée de différentes façons aux mesures du magnétisme terrestre. C'est ce que prouvent de nombreuses communications sur la déclinaison qu'on trouve dans les Travaux de l'Académie d'autrefois. Avec les instruments et les idées de ce temps, les recherches étaient d'abord, naturellement, peu sûres et tâtonnantes, bien qu'on ait dès la seconde moitié du XVIII^e siècle d'assez longues séries complètes d'observations de la déclinaison. On ne saurait pourtant utiliser ces matériaux qu'avec une extrême prudence. C'est ainsi qu'une série de mesures météorologiques et magnétiques relevées pendant plusieurs années à la Tour Ronde à Copenhague et publiées in extenso avec des observations provenant d'autres pays dans les éphémérides de la Societas meteorologica Palatina à Mannheim et par extraits dans les Travaux de l'Académie, constituent une documentation très sujette à caution. Pour les observations météorologiques, par exemple, on se servait en même temps d'un baromètre suspendu dans une pièce de la tour et d'un thermomètre à l'air

libre, et probablement parce qu'on constatait que cela allait très bien avec le baromètre dans la pièce, on y mit aussi pendant quelque temps le thermomètre, procédé évidemment peu heureux. Les mesures magnétiques étaient faites, elles aussi, pour la plupart dans ce local, mais on ne tenait pas compte des ancras et des autres pièces de fer qui se trouvaient dans les murs et les installations de la tour, et les éphémérides montrent que les mesures conservées de la déclinaison présentaient souvent pendant plusieurs journées de suite exactement la même valeur à la même minute, ce qui, d'après nos connaissances actuelles, n'a pas pu être juste.

Une amélioration sensible fut apportée par Hansteen grâce à son importante concentration d'informations sur les observations magnétiques de différentes régions de la terre, et à ses cartes magnétiques, qui firent date. C'est sans doute à cette occasion et sur l'initiative d'Ørsted que l'Académie mit au concours en 1811 le sujet: »Peut-on expliquer tous les phénomènes magnétiques de la terre par un seul axe magnétique ou faut-il en supposer plusieurs?« On sait que Hansteen prouva qu'un seul axe était insuffisant pour rendre compte des phénomènes, et que son mémoire fut couronné par l'Académie — mais comme on traversait une période difficile, Hansteen dut lui-même »rappeler à plusieurs reprises la médaille qu'on avait oubliée« avant de l'obtenir, et l'Académie n'avait pas les moyens nécessaires à l'impression de ses importantes »Untersuchungen über den Magnetismus der Erde«.

Pour ce qui est antérieur à 1842, je mentionnerai encore qu'Ørsted, après son voyage à Göttingue, où il était allé voir l'observatoire de Gauss en 1834, obtint l'adhésion du Danemark à l'association magnétique de A. von Humboldt, qui organisait dans plusieurs pays, suivant un plan prémédité, des mesures simultanées du magnétisme terrestre, en faisant prendre des observations toutes les cinq minutes pendant 6 jours de l'année. En Danemark ces observations furent effectuées à la nouvelle Haute École Polytechnique. On n'y observait que les variations de la direction de l'aiguille aimantée, et non la valeur absolue de la déclinaison. Or, lorsque toutes les observations des différents lieux furent réunies, les matériaux présentaient plusieurs phénomènes intéressants, entre autres celui d'un trouble magné-

tique — une tempête — qui s'était manifesté le même jour et de manière à peu près identique à Copenhague et à Milan. Détail intéressant: le volume de quelques-uns des aimants de Gauss était environ 30.000 fois plus grand que celui des aimants dont on se sert actuellement. L'aimant de Hansteen à Christiania, comme on disait alors, était encore plus grand que celui de Gauss.

Or, ce qui arriva il y a cent ans, c'est que les mesures du magnétisme terrestre eurent leur centre propre et qu'elles cessèrent d'être faites à la Tour Ronde, à l'École Polytechnique ou dans d'autres endroits où il fallait craindre des facteurs troublants. Le 1^{er} février 1842 les observations commencèrent dans un observatoire magnétique spécial, que l'Académie avait installé au bastion dit Gyldenløve près de la porte de Vesterport. La déclinaison y fut relevée, au début, quotidiennement toutes les trois heures de 5 heures du matin à 11 heures du soir. Quelques années après on y ajouta des observations de l'intensité horizontale. A ce nouvel observatoire c'était toujours les variations de la direction de l'aiguille aimantée qu'Ørsted étudiait surtout, fait qui ressort nettement de ce qu'il donnait dès le début tous les mois, dans les Travaux de l'Académie, une formule déduite de la marche diurne normale de la déclinaison. Il l'obtenait en faisant abstraction, de l'ensemble des matériaux, des observations prises les jours qui lui paraissaient anormaux, soit près de la moitié. C'est là un procédé assez risqué; en tout cas une expédition, qui a eu lieu bien plus tard aux régions arctiques — où, il est vrai, les troubles sont sensiblement plus grands qu'en Danemark — a trouvé, en écartant les jours où la marche diurne ne ressemblait pas à celle observée dans d'autres localités, une marche qui était à peu près l'inverse de ce qu'elle est en réalité au nord de la zone aurorale. On verra par la figure 1 ci-après comment Ørsted se tirait de ces difficultés. On y voit la marche diurne trouvée par le grand physicien pour les mois de février, mars et avril 1842 et la marche diurne pendant les mêmes mois de l'année 1932 — deux années qui offrent les mêmes rapports à la période des taches solaires, et qu'on doit par conséquent supposer sensiblement identiques. Il faut se rappeler, en comparant les deux courbes, qu'Ørsted ne disposait pour ses calculs que d'observations de toutes les

trois heures pendant une partie de la journée, tandis que la courbe de 1932 est basée sur des enregistrements continus — et en outre que la déclinaison était, en 1842, d'un peu plus de 17° W et en 1932 de 6° W à peine, d'où il s'ensuit que les extrêmes devaient être atteints un peu plus tôt à l'époque d'Ørsted.

Après la mort d'Ørsted, c'est le professeur d'Arrest qui devint l'animateur des travaux de l'Académie sur le magnétisme terrestre. Ils furent continués sans interruption et les résultats

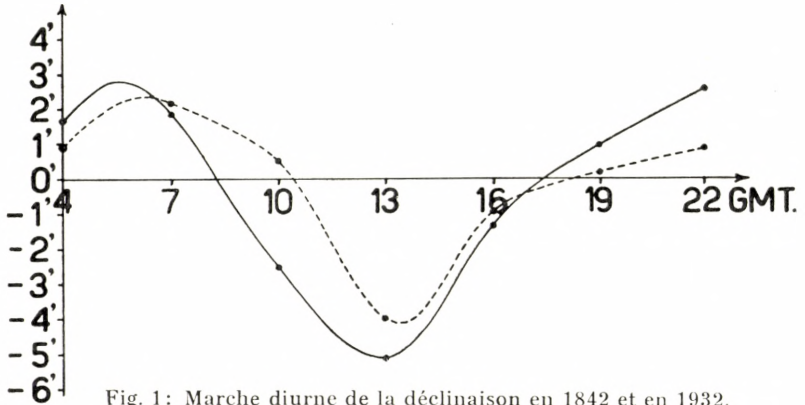


Fig. 1: Marche diurne de la déclinaison en 1842 et en 1932.

—•— Moyenne pour février, mars et avril 1842.

---•--- » » » » » 1932.

en étaient communiqués sommairement tous les ans pour le mois de décembre dans l'aperçu des travaux de l'Académie. Dans l'aperçu de 1859, d'Arrest, en faisant l'historique des recherches, a utilisé les matériaux pour expliquer — autant que le permettaient les observations prises — le phénomène connu sous le nom de variation séculaire. Depuis le début du siècle on avait cherché à déterminer, sans arriver au même résultat, le moment où l'aiguille aimantée qui, autant qu'on le savait, tournait vers l'ouest depuis plus de cent ans, allait revenir vers l'est. D'Arrest y allait avec beaucoup de méthode et calcula, après avoir soumis les matériaux à la critique et à la correction congrues, une formule d'interpolation qui, disait-il, devait non seulement remplacer l'énoncé de la loi vraie, mais dont on devait pouvoir encore se servir pour en tirer toutes les conclusions sur le passé et sur l'avenir qu'on tirerait de la formule même, si elle était connue. Si l'on tient compte de tous les matériaux disponibles,

la formule d'Arrest donne pour résultat que la déclinaison sera zéro à Copenhague en 1954. Il s'en faudra bien encore à cette date de 1 ou 2 degrés, mais c'est tout de même un très beau pronostic séculaire fait sur une base qu'on ose bien qualifier de quelque peu incertaine. La déclinaison zéro est d'ailleurs actuellement si proche de nous que ce moment historique est déjà atteint dans certaines localités de Bornholm. Il faut d'autant plus admirer la formule de d'Arrest qu'on considère sur

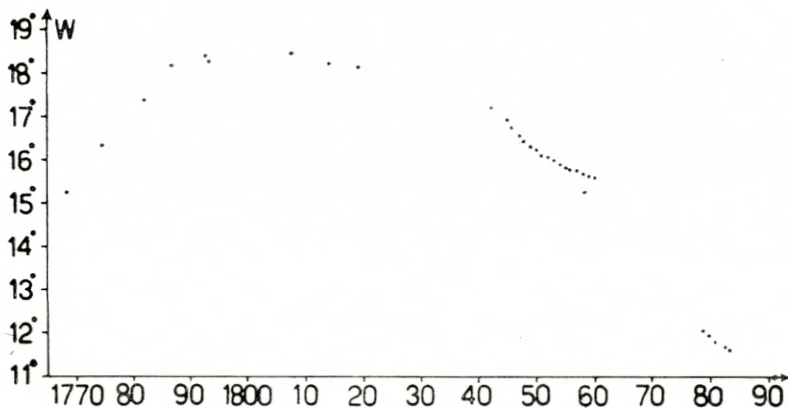


Fig. 2: La déclinaison à Copenhague.

quoi elle est basée. La figure 2 montre les valeurs de 1770; d'abord éparses, elles deviennent plus continues à partir de l'installation de l'observatoire du bastion Gyldenløve. On paraît y avoir eu quelque difficulté à maintenir les constantes des instruments; on voit par la figure comment les valeurs se suivent systématiquement, et en 1858 une partie en sont au-dessus d'une observation faite cette même année par le professeur Lamont, qui visitait la Scandinavie avec ses instruments de précision. La valeur de Lamont, qui est bien plus près d'une ligne régulière rejoignant les valeurs trouvées vers 1880, est probablement plus exacte. Car d'après ce que l'on sait, la variation séculaire suit un cours assez régulier, ce qui ressort par exemple du graphique ci-après (fig. 3) de la déclinaison à Copenhague pendant ces dernières années.

En 1862 l'observatoire magnétique permanent de l'Académie fut transféré pour 1800 rixdales au bastion de Rosenborg, mais

en même temps l'activité devint, si j'ose dire, plus dispersée. Pour des raisons d'ordre pratique — surtout pour répondre aux besoins de la navigation — les mesures magnétiques furent faites dans plusieurs localités du pays même et du Groenland. Les Danois firent un pas décisif pendant la première année dite

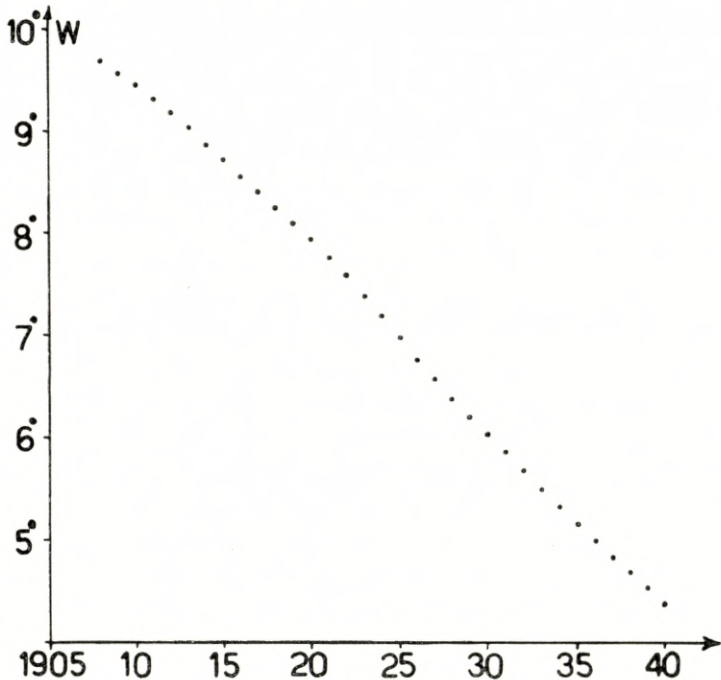


Fig. 3: La déclinaison à Rude Skov.

polaire internationale par l'établissement d'un observatoire magnétique à Godthaab au Groenland sous la direction d'Adam Paulsen. A son retour de Groenland, le terrain était déjà préparé à l'entreprise de travaux plus systématiques. Entre autres choses le Service hydrographique de la marine avait attiré l'attention du ministère de la Marine sur l'insuffisance des observations magnétiques dans notre pays, et la Faculté des Sciences et le Conseil de l'Université appuyèrent un projet en vue de créer un observatoire magnétique rattaché à l'Institut météorologique; les instruments qui avaient servi au Groenland réduiraient les frais de l'installation. Nommé chef, comme on disait alors, de l'Institut météorologique en 1884, Adam Paulsen proposa d'établir un

observatoire sur les remparts de la citadelle, mais ce fut toutefois le Jardin botanique qu'on finit par préférer lorsque le Parlement accorda les fonds nécessaires au budget de 1888—89.

Le fait le plus important pour les travaux, c'est sans doute la création à l'Institut météorologique d'un poste de sous-directeur pour les travaux sur le magnétisme terrestre, assurant ainsi un personnel suffisant pour la tâche dont Adam Paulsen se préoccupait si vivement et à laquelle il consacra tant d'énergie jusqu'à sa mort. L'observatoire commença ses travaux sans tarder — d'abord par des observations à des heures fixes comme précédemment — toutefois sans l'observation de 5 hs du matin — mais on commença aussi de se servir d'appareils enregistreurs pour l'enregistrement continu des variations, et — comme dans d'autres pays — on constata par là coup sur coup de nouveaux phénomènes. Si les travaux étaient pendant les premières années quelque peu entravés par les difficultés du début, cela allait bientôt mieux, et l'Institut météorologique entreprit la publication d'un Annuaire magnétique rédigé en français, selon le renseignement du livre des adresses, mais c'était surtout des chiffres, seules quelques phrases isolées et les rubriques étant en français; d'ailleurs le livre n'eut guère de lecteurs en Danemark. A côté de ces travaux Adam Paulsen s'occupa avec ardeur du dressage de cartes magnétiques du pays, et il fit chaque année, en personne, des tournées de mesurages; malgré beaucoup d'énergie les travaux n'avancèrent que lentement en raison des instruments dont on se servait — et c'était avant les automobiles. On recueillit toutefois une documentation précieuse pour la détermination des variations ultérieures.

L'installation au Jardin botanique ne fut pas de longue durée, car les tramways de Copenhague étant électrifiés au renouvellement du siècle, la situation devint désastreuse pour l'observatoire. En abritant convenablement les instruments, on obtenait bien des courbes relativement nettes, mais il était impossible d'effectuer les observations indispensables pour l'interprétation de ces courbes, et il fallut cesser la publication de l'Annuaire. Par la suite on a toutefois pu utiliser les enregistrements, à certaines fins tout au moins. Selon la concession de la Compagnie des Tramways, celle-ci devait réparer le dommage qu'elle causerait, et, surtout grâce à l'intervention du professeur K. Prytz, la Com-

pagnie paya la construction d'un nouvel observatoire à la forêt de Rude Skov et accepta de verser une subvention annuelle de 3000 couronnes pour l'augmentation des frais d'exploitation occasionnée par le déménagement — subvention que l'Institut météorologique reçoit encore, mais de la Ville.

Le nouvel observatoire fut construit en 1906, mais les difficultés du début durèrent encore cette fois-ci quelque temps; ces difficultés surmontées, la publication de l'Annuaire fut reprise, et les travaux ont continué depuis lors pratiquement sans interruption. Il manque seulement les enregistrements d'un petit nombre d'heures par suite d'une tempête magnétique si violente que les courbes dépassèrent les bords de la bande ou parce que la foudre est tombée sur l'installation, faisant sauter les plombs. On a pourtant réussi peu à peu à assurer le fonctionnement régulier dans la mesure du possible: on a installé des instruments susceptibles de faire des enregistrements par des tempêtes magnétiques même extraordinairement violentes, on dispose de lignes électriques tant aériennes que souterraines, et on a établi dans le pays un réseau de points d'appui, où l'on fait des mesures régulières, de sorte que, même si l'observatoire était détruit par une bombe, le fil qui relie le passé et l'avenir ne serait plus coupé. Mais ces mesures de sécurité sont aussi indispensables, car il faut envisager par exemple le cas où l'électrification successive des moyens de transport rendra inutilisable également l'observatoire de Rude Skov.

Le levé magnétique du pays a lui aussi été continué. Il a subi un développement curieux au cours des années. Du temps d'Ørsted c'était surtout de la déclinaison qu'on s'occupait, les autres éléments magnétiques — l'intensité horizontale et surtout l'intensité verticale — étant bien plus difficiles à mesurer et d'une manière générale bien plus incertains; c'était aussi la déclinaison qui seule importait pour les besoins de la vie pratique, notamment pour la navigation. De nos jours, ce sont les mesures de la déclinaison qui font le plus de difficulté, parce qu'elles demandent le plus de temps. Certaines activités de la vie pratique s'intéressent de nos jours particulièrement à la mesure de l'intensité verticale pour la prospection du sous-sol, et grâce aux instruments modernes ces mesures sont aujourd'hui non seulement les plus rapides et les plus faciles à faire, mais don-

nent encore les résultats relativement les plus exacts. Dans notre cas il faut ajouter l'examen électrique du sol. L'Institut météorologique est entré en collaboration étroite avec les organisations internationales pour le magnétisme terrestre et l'Annuaire magnétique qu'il publie est fait conformément aux directives adoptées par ces organisations relativement au contrôle et au contenu. Une publication parallèle, les »Communications magnétiques«, comprend d'autres travaux sur le magnétisme terrestre, jusqu'à présent surtout des monographies sur des instruments de mesure, dont l'Institut a fait pendant ces dernières années une exportation assez importante.

En 1926 on a installé un observatoire magnétique permanent très bien équipé à Godhavn au Groenland. Il relève comme tous les autres établissements groenlandais de la Direction du Groenland, laquelle a confié la conduite des travaux scientifiques et la publication de l'Annuaire magnétique du Groenland à l'Institut météorologique. A Godhavn comme à Rude Skov les mesures et les enregistrements ont été continués sans interruption depuis la fondation, et en plus du haut intérêt scientifique des matériaux recueillis, l'observatoire constituera la base nécessaire de la vaste exploration géophysique projetée du Groenland, qui fut confiée, à la veille de la guerre, à l'Institut géodésique et à l'Institut météorologique; il pourra être aussi une aide pour la mesure de la déclinaison et la connaissance de ses variations dans les eaux groenlandaises. Après le 9 avril 1940, qui coupa la liaison avec le Groenland, le Département du magnétisme terrestre de l'Institut Carnegie a bien voulu, à la demande de l'Institut météorologique, se charger de fournir l'observatoire de papier photographique, de piles, etc., de sorte que le service de l'observatoire paraît pleinement assuré même dans les conditions actuelles. Grâce à un don de l'Institut Carnegie et à une subvention de la fondation Rask-Ørsted, l'observatoire a été mis en état de faire — jusqu'à nouvel ordre comme le seul observatoire de la zone de l'aurore boréale — des enregistrements continus de l'intensité des rayonnements cosmiques. Les enregistrements sont étudiés en Amérique avec des enregistrements analogues d'autres latitudes.

Enfin il faut mentionner qu'au cours de la deuxième année polaire internationale 1932/33 — 50 ans après la première —

l'Institut météorologique avait installé, avec des subventions de l'État, de la fondation Carlsberg et de la fondation Rask-Ørsted, des observatoires magnétiques complets à Julianehaab et à Thulé au Groenland; et que, notamment grâce à des subventions de la fondation Rockefeller, l'Institut a aidé plusieurs autres pays à s'équiper d'instruments pour les mesures du magnétisme terrestre et leur a donné des instructions.

J'espère qu'on se sera rendu compte par ce que j'ai dit de l'étendue et de la consolidation des travaux qui ont été les fruits de l'initiative que l'Académie a prise dans le temps en donnant le branle aux recherches sur le magnétisme terrestre.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 15

UNTERSUCHUNGEN ÜBER
CALCIUMCHLORID UND DESSEN
MOONHYDRAT

VON

AKSEL TOVBORG JENSEN UND AXEL LANNUNG



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

INHALTSÜBERSICHT

	Seite
Einleitung	3
I. Bestimmung der Teilchengröße von durch Dehydratation von Monohydrat bei niedriger Temperatur hergestellten energiereichen Calciumchloridpräparaten	11
II. Gitterdynamische Klarlegung des Energieüberschusses des feinteiligen Calciumchlorids	13
III. Nachweis von zwei neuen Calciumchloridmodifikationen	17
IV. Betrachtungen über den Mechanismus des Hydratations-Dehydratationsprozesses mit besonderem Hinblick auf die Reversibilität, durch einige Beispiele erläutert. Dampfdruckmessungen am System Natriumbromid - Natriumbromiddihydrat	24

EINLEITUNG

In einer früheren Arbeit¹ wies der eine von uns nach, dass der Gleichgewichtsdampfdruck über dem System $\text{CaCl}_2 - \text{CaCl}_2, \text{H}_2\text{O}$ je nach der Darstellungsweise des Systems variiert, und zwar in einem Grade, der den üblichen verbreiteten Vorstellungen, dass die Salzhydratisierung ein fast vollständig reversibler Prozess ist, und dass die innere Energie eines — wasserhaltigen oder wasserfreien — Salzes von der Vorgeschichte des Salzes unabhängig ist, nicht entsprach. So bildeten sich bei teilweiser Entwässerung von Calciumchloridmonohydrat Anhydrid-Monohydratsysteme, deren Gleichgewichtsdampfdruck umso geringer war, je niedriger die angewandte Entwässerungstemperatur gewesen war. Bei Entwässerung bei 18° entstand ein System (1), dessen Gleichgewichtsdampfdruck bei 18° 0.002 mm betrug. Entwässerung bei 180° mit nachfolgender Sinterung bei 400° ergab ein System (10), dessen Gleichgewichtsdampfdruck bei 18° etwa 25 Mal grösser als der des 18° -Systems war. Ausgehend von einem geschmolzenen Calciumchloridpräparat, dessen besondere strukturelle Verhältnisse später näher besprochen werden, wurden Systeme (12) dargestellt, deren Gleichgewichtsdampfdrucke bei 18° bis zu 250 Mal grösser waren als der des erstlich erwähnten 18° -Systems. Die Hydratationswärmen der drei Systeme wurden zu 13.7, 2.5 und etwa 3 kcal pro Mol H_2O berechnet. Ferner wurde gezeigt, dass dieser Energieunterschied in der CaCl_2 -Phase des Systems zu suchen ist, da eine Hydratisierung immer zu Systemen $\text{CaCl}_2, \text{H}_2\text{O} - \text{CaCl}_2, 2\text{H}_2\text{O}$ führte, deren Dampfdruckverhältnisse so gut wie unabhängig davon waren, aus welchem $\text{CaCl}_2 - \text{CaCl}_2, \text{H}_2\text{O}$ -System sie entstanden waren.

¹ AXEL LANNUNG, D. Kgl. Danske Vidensk. Selskab, Math.-fys Medd. **XVII**, 10, 1940.

Da wir vermuten, dass Erscheinungen wie die erwähnten recht verbreitet sind, selbst wenn sie in der Regel quantitativ unbedeutender sind als hier, haben wir das System einer eingehenden röntgenkristallographischen Untersuchung unterworfen; unter anderem haben wir die gradweise Entwässerung von $\text{CaCl}_2 \cdot \text{H}_2\text{O}$ bis zu $\text{CaCl}_2 \cdot 0.14 \text{H}_2\text{O}$ bei 24.2° mit ständigen Röntgenaufnahmen verfolgt, um die thermodynamischen Daten mit einigen Aufschlüssen über die Feinstruktur der verschiedenartig behandelten CaCl_2 -Präparate zu ergänzen.

Röntgentechnik.

Wie in der früheren Arbeit erwähnt wurde, hat Dr. L. MISCHE einige der in dieser Arbeit behandelten Präparate röntgenographisch untersucht und nachgewiesen, dass es sich immer um dieselbe Monohydratform handelte, und soweit die angewandte Untersuchungsmethode eine Beurteilung zuließ, dass es auch immer die gleiche CaCl_2 -Modifikation war, die sich in allen untersuchten Präparaten fand. Debye-Linien anderer kristalliner Modifikationen liessen sich nicht auffinden. Da Dr. MISCHE's Diagramme keine sicheren Aufschlüsse über feinere Einzelheiten geben konnten, haben wir versucht die angewandte Technik an verschiedenen Punkten zu verbessern. Viele Ca-Salze geben eine sehr störende Untergrundschwärzung auf dem Debye-Diagramm. Diese haben wir vermindert, indem wir den Film mit 20μ Al-Folie bedeckten. Um die Linienbreite besser messen zu können, haben wir statt der von Dr. MISCHE benutzten Rundkamera mit der üblichen Peripherie von 180 mm eine solche mit 600 mm Peripherie benutzt. Das Präparat machte eine Umdrehung in der Minute. Durch die Anwendung von $\text{Cu K}\alpha$ -Strahlung statt der von Dr. MISCHE benutzten $\text{Fe K}\alpha$ -Strahlung haben wir allerdings etwas von der gleichzeitig erreichten höheren Dispersion wieder eingebüsst, aber dafür wenigstens eine dreifache Verminderung der Belichtungszeit erreicht. Die hygroskopischen Eigenschaften der Präparate erforderten nämlich auch während der Aufnahme der Filme ihre Anbringung in einer Glaskapillare (Hanff & Buest, Berlin), und deren Glaswand absorbiert weit mehr von der Fe- als von der Cu-Strahlung.

Frühere Arbeiten über die Kristallstruktur von wasserfreiem Calciumchlorid.

Durch ganz langsame Abkühlung der Schmelze stellten VAN BEVER und NIEUWENKAMP¹ Einkristalle von Calciumchlorid her, deren Struktur sie auf Grund von Drehkristallaufnahmen als ein rhombisches Pseudo-Rutilgitter mit 2 Molekülen in der Zelle und den Dimensionen: $a = 6.24$, $b = 6.43$, $c = 4.20 \text{ \AA}$ bestimmten. Sie gaben weiterhin an, dass sich diese Kristalle bei mechanischer Bearbeitung zu einer anderen Modifikation umwandeln, da sie undurchsichtig wurden, und da das Pulverdiagramm der zerkleinerten Einkristalle sich nicht mit den genannten Gitterkonstanten indizieren liess. VAN BEVER und NIEUWENKAMP schreiben, dass sich von der Pseudo-Rutilmodifikation keine Pulverdiagramme herstellen lassen. Später haben jedoch DÖLL und KLEMM² durch Entwässerung von Calciumchloridhydrat im Hochvakuum und Sinterung des Entwässerungsproduktes bei nicht angegebener Temperatur ein Präparat hergestellt, von dessen Pulverdiagramm sie für 13 Linien $\sin^2\theta$ Werte angeben, die sich mit einer kleinen Änderung, die indessen innerhalb der Versuchsungenauigkeit liegt, mit den von VAN BEVER und NIEUWENKAMP angegebenen Dimensionen indizieren lassen.

Bekanntlich kommt Titandioxyd in drei Modifikationen vor: Rutil, Anatas und Brookit. Das Verhältnis zwischen den Ionenradien von Titan und Sauerstoff ist 0.49 und das entsprechende Verhältnis für Calcium und Chlor 0.55, also noch immer innerhalb des Radiusquotientgebietes 0.41 bis 0.73, wo man erwarten kann, dass eine Verbindung AX_2 in einem Koordinationsgitter mit 6:3 Koordination³ kristallisieren wird. Daher muss die Existenz von anderen Calciumchloridmodifikationen der Typen Anatas und Brookit — eventuell Deformationsstrukturen von diesen — von vornherein als wahrscheinlich angesehen werden. Dagegen lässt sich nicht erwarten, dass sich Calciumchlorid des Rutiltyps bei mechanischer Behandlung in eine weniger stabile Modifikation umlagern lässt. VAN BEVER und NIEUWENKAMP's Beobachtung lässt sich jedoch durch die Annahme erklären, dass in ihrem Präparat zwischen den grossen Kristallen des Pseudo-Rutiltyps

¹ A. K. VAN BEVER, W. NIEUWENKAMP, Z. f. Kristallographie, **90** (1935), 374.

² W. DÖLL, W. KLEMM, Z. anorg. allg. Chem., **241** (1939), 239.

³ O. HASSEL, Crystal Chemistry, London (Heinemann) (1935), 36.

kleine Mengen der beiden — oder einer der beiden — instabilen Calciumchloridmodifikationen, die wir gefunden haben, und die im folgenden näher besprochen werden, eingelagert waren. Kleine, für die Drehaufnahme ausgewählte Einkristalle enthielten keine instabile Modifikation, während sie in einer etwas grösseren Menge des Präparats für die Pulverdiagramme zu finden war.

Untersuchung von bei 35° und 24.2° entwässertem Calciumchlorid.

Es lag uns sehr daran, bei der röntgenographischen Untersuchung des Entwässerungsproduktes von $\text{CaCl}_2, \text{H}_2\text{O}$ die Pulverisierung unter Paraffinöl, die DR. MISCH vorgenommen hatte, um die Präparate für die Pulveraufnahme in Glasröhren anzubringen, zu vermeiden. Teils kann ein stark hygroskopischer Stoff wie Calciumchlorid bei einer derartigen Behandlung leicht etwas Wasser aufnehmen, teils und namentlich risikiert man leicht, dass eventuelle instabile feste Phasen bei der Pulverisierung in die stabile Form übergehen. Vgl. die Verhältnisse bei der Pulverisierung der instabilen Zinksulfidmodifikation Wurtzit, die sich leicht in die stabile Zinkblende umwandelt¹.

Das Calciumchloridmonohydrat wurde wie in der zitierten früheren Arbeit beschrieben, durch teilweise Entwässerung — im Vakuumexsikkator über konzentrierter Schwefelsäure — eines analysierten Merck-Präparates, das im wesentlichen aus dem Dihydrat bestand, hergestellt. Mit Hilfe einer Art von Pipette, die aus einer zu einer passenden feinen Spitze ausgezogenen Glasröhre hergestellt und mit einem kleinen Gummiball versehen war, werden zwei Glaskapillaren mit Monohydrat gefüllt und in einem Halter in einer Glasröhre angebracht, die mit Hilfe eines Normalschliffs in direkter Verbindung mit einer Quecksilberdampfpumpe aus Stahl steht. Die Glasröhre mit den Präparatgläsern wird in einem Thermostaten bei 35° angebracht und 130 Stunden lang ununterbrochen ausgepumpt, worauf durch die Pumpe trockene Luft in die Glasröhre hineingelassen wird. Das eine Präparatglas wurde zugeschmolzen, und der Inhalt röntgenographisch untersucht; es erwies sich, dass er aus einer Mischung

¹ A. SCHLEEDE, H. GANTZCKOW, Z. f. Physik, **15** (1923), 184.

von Pseudo-Rutil- CaCl_2 und weniger $\text{CaCl}_2, \text{H}_2\text{O}$ bestand. Die Entwässerung des Inhalts des zweiten Glases wurde fortgesetzt bis zu 530 Stunden im ganzen, aber das Präparat verunglückte vor der röntgenographischen Untersuchung.

Bei der Entwässerung des Monohydrats bei 24.2° wurde das Vorgehen derart verändert, dass es möglich war, den langwierigen Entwässerungsprozess durch Wägung des Calciumchlorids in den Glaskapillaren zu verfolgen; von diesen werden zu dem Versuch 11 benutzt, die in zwei mit im ganzen 11 nummerierten Ausbohrungen versehenen Aluminiumblöcken angebracht werden, die während der Entwässerung in eine durch Normalschliff mit der Stahl-Quecksilberdampfpumpe verbundene Glasröhre gestellt werden. Wenn nach der Entwässerung noch Platz für das Zuschmelzen und danach für die Vermischung des Inhalts sein soll, können in den Präparatkapillaren nur etwa 10 mg Monohydrat angebracht werden. Die Gewichtsverluste, die bestimmt werden sollen, sind daher von der Grössenordnung 1—2 mg. Darum wird eine aperiodische Kuhlman Mikro-Ultra Wage benutzt. Die Kapillarröhren werden einzeln in einem speziellen Aluminiumwägerohr gewogen, das derart ausgebohrt ist, dass sie bequem und sicher darin angebracht werden können, und dass der innere Hohlraum so klein wie möglich ist. Das Wägerohr, das nur 2.7 g wiegt, wird mit einem darauf geschraubten Deckel luftdicht verschlossen; es hat blankpolierte Oberflächen, wodurch der Einfluss von hygroskopisch darauf gebundenem Wasser praktisch vollständig eliminiert wird.

Die Kapillarröhren werden mit Calciumchloridmonohydrat gefüllt, das darauf im Vakuumexsikkator über konzentrierter Schwefelsäure einige Tage lang nachgetrocknet wird, wodurch es, wie bereits früher aufgezeigt und hier röntgenographisch verifiziert, genau die Zusammensetzung $\text{CaCl}_2, \text{H}_2\text{O}$ erhält. Nach der Wägung wird das Monohydrat wie oben erwähnt mit der Stahl-Quecksilberdampfpumpe bei 24.2° entwässert, wobei nach bestimmten Zeiträumen eins der Präparate entnommen und gewogen wird. Nach der abschliessenden Wägung im Aluminiumrohr, wobei die Präparate keine Gelegenheit hatten, merkbare Mengen Wasserdampf aus der Luft aufzunehmen, wird die Kapillarröhre schnell mit einem Mikrobrenner zugeschmolzen, so dass eine Wasseraufnahme wenig wahrscheinlich war.

Bei der Entwässerung, die sich, was das zuletzt entnommene Präparat anbelangt, über 58 Tage erstreckte, wurde kein Trockenmittel angewandt; es hätte zwar die Rede davon sein können, Phosphorpentoxyd zu benutzen, aber wir befürchteten, dass dieses im Hochvakuum im Laufe der langen Versuchszeit in gewissem Ausmass zu den Präparaten hinüber sublimieren würde. Die Anwendung einer Quecksilberfalle erübrigte sich, da die elektrische Heizung der Pumpe in passender Weise eingestellt und für effektive Wasserkühlung gesorgt wurde.

Bei der Entwässerung zeigte es sich, dass man scheinbar nicht mit der gleichen Entwässerungsgeschwindigkeit für alle Präparate rechnen konnte, obgleich das Ausgangsmaterial ja identisch war. Röntgenographisch liess sich in zwei Fällen nachweisen, dass die Entwässerung des Monohydrats in den Kapillarröhren nicht gleichartig in der ganzen Masse vor sich geht, sondern von oben her und allmählich nach unten fortschreitet. Zur näheren Beleuchtung dieser Beobachtung haben wir die Menge Wasserdampf, die nach den Mikrowägungen aus einer mit Monohydrat gefüllten Kapillarröhre durchschnittlich in der Sekunde ausgepumpt wird, mit der Menge verglichen, die unter gewissen, im Folgenden näher angegebenen Bedingungen nach unserer Berechnung ausgepumpt werden muss. Unter der Voraussetzung, dass in unserem Monohydratpräparat während des Auspumpens ständig ein Wasserdampfdruck von 0.5 Dyn/cm^2 aufrechterhalten wird, entsprechend etwa $1/10$ des Gleichgewichtsdampfdrucks — einer Grösse, für die wir von der früheren Arbeit her gewisse Anhaltspunkte haben — wollen wir zunächst berechnen, wieviel Wasserdampf sich in einer Sekunde aus einer Höhlung im Präparat durch eine zylindrische Pore mit dem Durchmesser $0,01 \text{ cm}$ und der Länge 1 cm auspumpen lässt. Diese Pore reicht von der Oberfläche bis zur Mitte der Präparatsäule, und wir setzen weiterhin voraus, dass der Wasserdampf ausschliesslich aus der Höhlung und nicht von den Wänden der Pore herrührt.

Nach M. KNUDSEN¹ lässt sich die Menge Q_t eines Gases, das bei der Temperatur $t^\circ\text{C}$ und dem Druckunterschied $p_1 - p_2$ in 1 Sekunde durch einen zylindrischen Kanal mit dem Durchmesser $2R$ strömt, wenn $2R$ verschwindend klein ist im Verhältnis zur Länge L des Kanals und zu der durchschnittlichen freien Weg-

¹ M. KNUDSEN, Ann. d. Physik, 28 (1909), 75.

länge der Moleküle des betreffenden Gases, durch folgende Formel berechnen:

$$Q_t = \frac{4}{3} \cdot \sqrt{2\pi} \cdot \frac{R^3}{L} \cdot \sqrt{\frac{1 + \alpha t}{\varrho_0}} (p_1 - p_2)$$

Bei Anwendung von absoluten Einheiten bedeutet Q_t die Menge des Gases ausgedrückt durch das Produkt von dem Volumen in cm^3 und dem Druck in Dyn/cm^2 . α ist der Ausdehnungskoeffizient des Gases, ϱ_0 dessen spezifisches Gewicht bei 0° und 1 Dyn/cm^2 Druck.

Im vorliegenden Falle finden wir, wenn wir $p_1 - p_2$ zu 0.5 Dyn/cm^2 ansetzen, $Q_t = 0.0077 \text{ (cm}^3 \cdot \text{Dyn/cm}^2)$. Da nach den Resultaten der Mikrowägungen im Laufe von 2 Monaten aus einem Präparat $1 \text{ mg H}_2\text{O}$ oder durchschnittlich $0.26 \text{ (cm}^3 \cdot \text{Dyn/cm}^2)$ in der Sekunde ausgepumpt wird, so bedeutet das, dass wir in unserem Präparat gleichzeitig 34 solcher Poren in Tätigkeit haben müssen. Da der Durchmesser der Kapillarröhren 0.1 cm betrug, so sieht man, dass die Grössenordnung mit diesem Wert übereinstimmt. Scheinbare Verschiedenheiten in der Entwässerungsgeschwindigkeit von der einen Kapillarröhre zu der anderen lassen sich daher durch zufällige grössere oder geringere Packungsdichte des Monohydrats erklären. Eine nach der Tiefe zunehmende Entwässerungsverzögerung ist nach diesen Ausführungen gleichfalls leicht erklärlich.

Die beobachteten Entwässerungsgeschwindigkeiten waren jedoch trotzdem von der gleichen Grössenordnung wie der durch Interpolation von Daten der früheren Arbeit gefundene Wert. Es muss hier angeführt werden, dass die benutzte Stahl-Quecksilberdampfmaschine schneller arbeitet als die früher benutzte Pumpe, und dass die Pumpleitung abgesehen von den Kapillarröhren bei unseren Versuchen weit günstigere Dimensionen hat.

Alle Linien, die von den 24.2° - und 35° -Präparaten stammen, gehörten der Pseudo-Rutil- CaCl_2 -Modifikation an, und ihre Intensität stieg in den 24.2° -Präparaten gleichmässig mit fallendem Wassergehalt. Wir ziehen hieraus den Schluss, dass keine andere Modifikation von CaCl_2 während der Entwässerung zu-

gegen war, und dass der abnorm hohe Energieinhalt des wasserfreien CaCl_2 ausschliesslich dessen Zustand zuzuschreiben ist. Hierüber geben nun die Röntgendiagramme gewisse Aufschlüsse. Das Diagramm des 35° -Präparates enthält 28 messbare Linien hinauf bis $\theta = 41.4^\circ$. Diese gehören mit Ausnahme einer einzigen alle der Pseudo-Rutilform für CaCl_2 an. Diese eine Linie war die stärkste der Linien des Monohydrats; jedoch konnten auch zwei andere der stärkeren Monohydrat-Linien bei sehr genauer Durchsicht des Films nachgewiesen werden. Die θ -Werte dieser 27 Linien waren natürlich nicht ganz genau gleich den θ -Werten für geschmolzenes oder gesintertes CaCl_2 , aber die meisten Differenzen sind nur 0.1—0.2 mm auf dem Film, und es finden sich ausserdem beinahe ebenso viele positive wie negative Abweichungen. Es lässt sich also kein Unterschied zwischen den Gitterkonstanten der geschmolzenen, der gesinterten und der durch Entwässerung im Hochvakuum hergestellten Calciumchloridpräparate feststellen. Da die Hochvakuumpräparate keine höheren Linien als bis zu 41.4° und nur verwaschene Linien ergeben, so zeigt eine Überschlagsrechnung, dass die Gitterkonstanten für das 35° -Präparat mit einer Unsicherheit von $\pm 1/2\%$ definiert sind.

Die im Hochvakuum hergestellten Präparate geben, wie eben erwähnt, verwaschene Linien, und diese können sich bei höheren θ -Werten als 41° gegen die starke Untergrundschwärzung nicht geltend machen. Es ist wohl nicht überflüssig darauf hinzuweisen, dass diese Unschärfe der Linien in den Diagrammen von Präparaten, die bei niedriger Temperatur hergestellt worden sind, im vorliegenden Falle keine anderen Ursachen haben kann als die sehr geringe Grösse der Elementarteilchen. Die Linien eines Pulverdiagrammes können abnorm breit werden, wenn die Teilchen zwar gross sind, aber etwas verschiedene Gitterkonstanten haben. Dies ist gewöhnlich die Ursache für die Linienverbreiterung bei Diagrammen von isomorphen Mischkristallen, was jedoch hier nicht zu erwarten ist, da wir es mit einem sehr reinen Stoff zu tun haben. Es ist zwar ausser reinem Calciumchlorid auch Wasser vorhanden, aber man kann sich nur schwer vorstellen, dass bei der Entwässerung noch etwas Wasser in den grösseren Teilchen des Stoffes zurückbleiben könnte, so dass die Teilchen in der Hauptsache Pseudo-Rutilstruktur haben

sollten, die als typisches Koordinationsgitter durch ein sehr kompaktes Arrangement der Atome charakterisiert ist, denn in diesem könnten die Wassermoleküle schwerlich abwechselnd an die Stelle von Ca^{++} - und Cl^- -Ionen treten.

Eine dritte Form von Gitterunvollkommenheit hatten wir von vornherein erwartet, nämlich reine Aufbaufehler wie Leerstellen, »Verhakungen« und andere Fehlordnungen mit den sich daraus ergebenden Ausbeulungen und Schiefheiten in den Atomebenen, kurzum, die Art von Gitterstörungen, die formell als »eingefrorene Wärmeschwingungen« angesehen werden können. Derartige Fehler führen zu keiner Verbreiterung der Linien, sondern nur dazu, dass ihre Intensität systematisch geschwächt wird. Daher können sie also nicht für die abnorme Linienbreite in unseren Diagrammen verantwortlich gemacht werden. Ob derartige Gitterstörungen ebenfalls vorhanden sind, können wir nicht sagen. Relative photographische Messungen der Linienintensität reichen nicht aus, um das zu entscheiden¹.

Bestimmung der Teilchengröße von bei niedriger Temperatur hergestelltem CaCl_2 .

Es wurden photometrische Messungen an den Filmen der folgenden drei Präparate vorgenommen: ein weiter unten eingehender behandeltes Präparat, 2 $\frac{1}{2}$ Stunden bei 400° gesintert, das 35°-Präparat mit unbekanntem aber kleinem Wassergehalt, sowie ein 24.2°-Präparat vollständig ohne Monohydratlinien, obgleich es nach Mikrowägungen noch 0.14 Mol H_2O enthielt. Das Präparat bildete jedoch einen so festen Pfropfen im Glas, dass es nur noch im untersten Teil desselben Monohydratphase enthielt.

Diese Filme wurden mit einem Zeiss Registrierphotometer photometriert, und die Halbwertsbreite einer Anzahl von Linien mit dem Messmikroskop ausgemessen. Als Halbwertsbreite wurde einfach die Breite der Photometerberge in halber Höhe angenommen, und irgendwelche Versuche, eine Korrektur für den krummen Verlauf der Schwärzungskurve des Films anzubringen, wurden nicht unternommen.

¹ Literatur bei R. FRICKE, Z. f. Elektrochem., **46** (1940), 491.

Bei der Berechnung der Partikeldimensionen benutzen wir die Formel

$$b = \frac{k \cdot \lambda}{t \cos \theta} \quad (1)$$

wobei b die Halbwertsbreite der Linie in reinen Zahlen bedeutet, θ den Glanzwinkel, λ die Wellenlänge und t die Dicke der Teilchen in denselben Einheiten. k ist eine Konstante = 0.90. Die Formel zeigt, welche Abhängigkeit zwischen der Linienbreite und den Partikeldimensionen in dem idealen Fall, dass die Primärstrahlen parallel verlaufen und das Präparat unendlich dünn ist, besteht. b lässt sich also nicht direkt messen, aber wie an anderer Stelle¹ aufgezeigt wurde, lässt es sich aus der Breite einer mit gewöhnlicher Kamera aufgenommenen Linie hkl annähernd berechnen, falls man diese sowohl für ein Präparat mit »grossen« Teilchen (die Halbwertsbreite sei hier B) und für ein in derselben Aufstellung aufgenommenes Präparat eines feinen Pulvers mit den gleichen Dimensionen und gleicher Dichte kennt (die Breite sei hier B'). b für das feine Präparat ist dann gleich $\sqrt{B'^2 - B^2}$.

In Tabelle 1 ist die auf den Photometerplatten gemessene doppelte Linienbreite von drei klaren und isolierten Linien auf den drei Diagrammen aufgeführt.

Tabelle 1.
Linienbreite in mm auf Photometerplatten von drei Präparaten von CaCl_2 .

Linienindices	Linienbreite der Präparate			b_{35}	$b_{24.2}$
	B	B'_{35}	$B'_{24.2}$		
122	1.6	2.5	3.2	1.92	2.77
140	1.8	2.5	3.0	1.73	2.40
222	1.65	2.5	3.0	1.90	2.53
				1.85	2.57

Die drei Linien liegen sehr nahe beisammen, und für sie alle gilt $\frac{1}{\cos \theta} \approx 1.14$.

¹ A. TOVBORG JENSEN, D. Kgl. Danske Vidensk. Selskab, Mat.-fys. Medd. **XX**, 8, 1943.

Aus den Durchschnittswerten für b erhält man

$$t_{35^\circ} = 163 \text{ \AA} \quad \text{und} \quad t_{24.2^\circ} = 117 \text{ \AA}.$$

Die nach Formel (1) berechnete Zahl t_{hkl} ist die in \AA angegebene Dicke der Teilchen in einem monodispersen System, das dieselbe Linienverbreiterung wie das vorliegende liefert, wobei t rechtwinklig zu der Fläche hkl gemessen wird.

Die berechneten t -Werte sind keine exakt definierten Grössen, ergeben aber trotzdem vermutlich einen zutreffenden Ausdruck für die Teilchengrösse der Präparate. Über die Form der CaCl_2 -Teilchen können wir nichts mit Sicherheit wissen, ausser dass von extremer Ausbildung in bestimmten Richtungen wie bei ARNFELT'S¹ Schichtengittern nicht die Rede ist, denn die Linienverbreiterung weist keine systematischen Variationen mit den Linienindices auf.

117 und 163 \AA sind recht grosse Dimensionen für kolloidale Teilchen, weit grösser als sie sich bei den feinen Metallkolloiden finden. Die beiden Präparate enthalten ca. 19450 beziehungsweise 52600 Moleküle CaCl_2 pr. Teilchen. Die Teilchen sind jedoch so klein, dass ein grosser Teil der Atome in der Oberfläche sitzt. Würfelförmige CaCl_2 -Kristalle mit der Seitenlänge 117 bzw. 163 \AA würden somit 26 % respektive 19.2 % aller Elementarzellen in der Oberfläche enthalten. 13 % beziehungsweise 9.6 % der Ca^{++} - und Cl^- -Ionen würden »an der Aussenseite sitzen«.

Gitterdynamische Erwägungen.

Aus den Dampfdruckdaten der früheren Arbeit wurde die molare Hydratationswärme ($\text{CaCl}_2 + \text{H}_2\text{O}_{fl.} \rightarrow \text{CaCl}_2, \text{H}_2\text{O}$) für ein im Hochvakuum bei 35° entwässertes CaCl_2 -Präparat zu 10.6 kcal berechnet; durch Interpolation erhält man für ein entsprechendes bei 24.2° hergestelltes Präparat 12.6 kcal, während der Wert für ein gut kristallisiertes, mikrokristallines durch Entwässerung bei 180° und nachfolgender $2^{1/2}$ -stündiger Sinterung bei 400° hergestelltes Präparat 2.5 kcal betrug. Es handelt sich also um ganz bedeutende Energiemengen, die auf diese Weise in den CaCl_2 -Präparaten akkumuliert werden, die, wie die Röntgenuntersuchungen ergaben, bei einjähriger Aufbewahrung

¹ H. ARNFELT, Arkiv för Matematik, Astronomi och Fysik, **23** B (1932) Nr. 2.

in zugeschmolzenen Glasröhren bei gewöhnlicher Temperatur keine merkbare Veränderung erleiden.

Wir wollen nunmehr die gitterdynamische Bedeutung dieser Überschüsse von innerer Energie im Betrage von 8.1 bzw. 10.1 kcal/Mol bei den CaCl_2 -Präparaten von geringer Teilchengrösse betrachten. Wir setzen wie früher voraus, dass die Teilchen würfelförmig und ohne innere Gitterfehler sind. Dann kann man sich die Überführung eines aus kleinen Teilchen bestehenden Präparates in ein solches mit grossen Teilchen durch blosses Aufeinanderstapeln der kleinen Kristalle vorstellen. Dadurch kann die ganze überschüssige Energie gewonnen werden. Diese ist beinahe gleich der Arbeit, die die alleroberste Ionenschicht in jeder Kristallfläche an der allerobersten Ionenschicht der Kristallfläche, die an sie herangelegt wird, ausführt. Wir wollen von der Arbeit, die von Kräften aus tiefer gelegenen Schichten ausgeführt wird, absehen, da die Kraftwirkung, die von einer Fläche mit gleichmässig verteilten positiven und negativen Ionen auf ein einzelnes Ion ausgeübt wird, ungeheuer schnell mit zunehmendem Abstand abnimmt, wie z. B. von LENNARD-JONES und DENT¹ hervorgehoben worden ist. Wir wollen ausserdem davon absehen, dass die Ionen in den Kanten und Ecken stärker angezogen werden als die Ionen etwas innerhalb einer Fläche. Von diesen Überlegungen ausgehend wollen wir die Betrachtungs- oder Ausdrucksweise einführen, dass die Ionen im Innern des Kristalls die volle Gitterenergie ausgenützt haben, während die Ionen in der Oberfläche (in unseren Präparaten 9.6 bzw. 13% der gesamten Anzahl) noch einen unausgenutzten Bruchteil x der Gitterenergie übrig haben. Diese unausgenutzte Gitterenergie ist gleich dem Energieüberschuss. Setzen wir mit VAN ARKEL und DE BOER² die Gitterenergie für CaCl_2 zu 525 kcal/Mol an, lässt sich x aus folgenden Gleichungen berechnen:

$$525 - 90.4/100 \cdot 525 - 9.6/100 (1 - x) 525 = 8.1 \quad (\text{I})$$

$$525 - 87/100 \cdot 525 - 13/100 (1 - x) 525 = 10.1 \quad (\text{II})$$

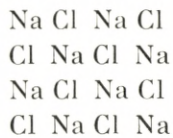
(I) ergibt $x = 0.161$; (II) $x = 0.148$, also im Durchschnitt zwischen $1/6$ und $1/7$. Diese Zahl beruht ausschliesslich auf experimentellen Daten.

¹ J. E. LENNARD JONES, BERYL M. DENT, *Transact. Faraday Soc.*, **24** (1928), 92.

² A. E. VAN ARKEL, J. H. DE BOER: *Chemische Bindung als elektrostatische Erscheinung*, Lpz. (Hirzel) 1931, 66.

Wir wollen im folgenden versuchen, die zu erwartende Grösse von x durch rein gittertheoretische Betrachtungen entsprechender Verhältnisse bei NaCl, die verhältnismässig leicht zu übersehen sind, zu finden.

In einem NaCl-Kristall ist die Fläche 100 in folgender Weise mit Na^+ und Cl^- besetzt:



Jedes Ion in der Oberfläche hat also fünf Nachbarionen mit entgegengesetzter Ladung, vier in der Oberfläche und eins darunter, während ein Ion im Inneren des Kristalls sechs hat. In erster Annäherung möchte man daher annehmen, dass x für diese Fläche gleich $1/6$ ist. Die Zahl $1/6$ ist jedoch zu gross. Für den Coulombschen Teil der Gitterenergie gestattet uns eine Berechnung nach von KOSSEL¹ und HØJENDAHL² angegebenen Linien gewisse Schlüsse zu ziehen. Das Potential eines Ions, das in richtiger Weise ausserhalb einer monomolekularen NaCl-Schicht wie der oben beschriebenen angebracht ist, lässt sich danach zu 0.0674 berechnen, wo das Potential eines Ions im Inneren des Kristalls den bekannten Madelungschen Wert 1.75 in denselben relativen Einheiten hat. Legt man zwei NaCl-100-Flächen zusammen, gewinnt man gerade die potentielle Energie eines einzelnen Ions für jedes Molekül, das während des Prozesses von der Oberfläche verschwindet. x wird unter den angeführten Voraussetzungen gleich $0.0674/1.75 = 0.0385$ oder 3.85% der gesamten Gitterenergie, also erheblich weniger als $1/6$, ein Wert, der im ersten Augenblick plausibel erscheint.

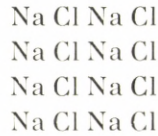
Man begeht keinen grossen Fehler, wenn man die Kräfte aus tiefer liegenden Schichten vernachlässigt. Das Potential eines Ions, in richtiger Weise über der Fläche 100 in einem unendlich dicken NaCl-Kristal angebracht, ist 0.0663, noch nicht 2% verschieden von dessen Potential über einer monomolekularen Schicht des Kristalls. Bei diesen Betrachtungen ist nicht berücksichtigt

¹ W. KOSSEL, Naturwissensch., 18 (1930), 906.

² K. HØJENDAHL, D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd. XVI,2 1938, 144.

worden, dass die Anziehung an den Kanten und Ecken grösser ist als innerhalb der Fläche, ein Umstand, der bei abnehmender Kristallgrösse steigende Bedeutung erhält.

Für die meisten anderen Flächen als die Fläche 100 lässt sich x nicht ohne stark willkürliche Annahmen über den Charakter der betreffenden Fläche berechnen. So besteht 110 aus Schichten nach dem Schema



111 sogar aus wechselnden Schichten von



Schichten wie diese können keine Begrenzung eines Kristalls über ein grösseres Areal hin bilden.

Wenn die Flächen 110 und 111 ausgebildet sind, sind die sie begrenzenden Ionenschichten nach KOSSEL¹ wahrscheinlich nicht eben, wie bei Fläche 100, sondern bei 110 mit Rillen und bei 111 mit Buckeln versehen. x wird für derartige Flächen zweifellos weit grösser als 3.85⁰/₁₀₀ sein.

Die Berechnung von x im Anschluss an KOSSEL und HØJENDAHL kann am Rutilgitter nicht vorgenommen werden. Für die Fläche 100 im Rutil² bestehen ähnliche Schwierigkeiten, wie wir sie bei 110 und 111 im NaCl sahen. 100 und 010 im CaCl₂ werden aus diesem Grunde einen grossen x -Wert haben.

001 im Rutil ist am leichtesten übersehbar, und es ist denkbar, dass sie eine ausgedehnte, völlig ebene Kristallfläche bildet. Begrenzt diese Fläche einen Kristall, so hat ein positives Ion in der Oberfläche vier negative Nachbarn, während ein Ion im Kristallinnern sechs hat. Ein negatives Ion in der Oberfläche hat zwei positive Nachbarn an Stelle von drei im Kristallinneren. Man kann also

¹ W. KOSSEL, loc. cit.

² Siehe das ausgezeichnete Bild in M. v. LAUE und R. v. MISES: Stereoskopbilder von Kristallgittern II, Berlin (Springer) 1936.

erwarten, dass x für diese Fläche weit grösser wird als für 100 im NaCl, wo die entsprechende Anzahl von Nachbarionen fünf bzw. sechs war.

Diese Betrachtungen scheinen uns zu zeigen, dass sich für x für CaCl_2 in Übereinstimmung mit den experimentell gefundenen ca. 15% ein Wert von der Grössenordnung 5—15% erwarten lässt.

Indessen haben wir bei unseren Berechnungen von x unter anderem vorausgesetzt, dass die Teilchen würfelförmig und innerhalb jedes Präparates von vollständig gleicher Grösse sind. In Wirklichkeit ist dies kaum der Fall, und es verhält sich daher so mit unseren Untersuchungen, dass die thermodynamischen Daten sich hauptsächlich auf die kleinsten Teilchen beziehen, während die röntgenographischen Daten eine Funktion aller vorhandenen Teilchen sind. Aus diesem Grunde lässt sich einsehen, dass die berechneten x -Werte prinzipiell zu gross ausfallen müssen, wenn die Präparate von ungleichmässiger Teilchengrösse sind.

Untersuchung von gesintertem und geschmolzenem Calciumchlorid.

Nachweis von zwei neuen Modifikationen.

Ein Präparat von CaCl_2 , das durch Entwässerung des Monohydrats bei 180° hergestellt und danach $2\frac{1}{2}$ Stunde bei 400° gesintert wurde, und das daher nach der Darstellungsweise dem Anhydrid in dem in der Einleitung erwähnten System (10) der früheren Arbeit entsprechen sollte, ergab schöne Diagramme mit 67 messbaren α_1 -Linien. Das Diagramm war wie zu erwarten ein Pseudo-Rutildigramm, enthielt aber ausserdem drei schwache Linien, die nicht zu diesem gehörten, sondern von der später entdeckten CaCl_2 -Modifikation 2, die weiter unten besprochen werden wird, herrührten. Das Diagramm liess sich im übrigen mit folgenden Gitterkonstanten indizieren: $a = 6.21_7$, $b = 6.40_6$, $c = 4.13_5 \text{ \AA}$, die nur ganz wenig von den von VAN BEVER und NIEUWENKAMP und von DÖLL und KLEMM gefundenen Werten abweichen. Die Indizierung ist bis zu $\sin^2 \theta = 0.51$ (37 Linien) durchgeführt, und die Übereinstimmung zwischen berechneten

und gemessenen $\sin^2 \theta$ -Werten ist zufriedenstellend. Die relativen Intensitäten, die von einem Kristall mit der von VAN BEVER und NIEUWENKAMP angegebenen Pseudo-Rutilstruktur zu erwarten waren, wurden berechnet. Es bestehen bis zu $\sin^2 \theta = 0.5145$ 86 Reflektionsmöglichkeiten, und die Übereinstimmung zwischen berechneten und beobachteten Intensitäten ist so gut, dass sie kaum durch kleine Änderungen der variablen Parameter der Struktur wesentlich verbessert werden könnte. Tabelle 2.

Bei Abkühlung von reinem, wasserfreiem Calciumchlorid, das in einem Platintiegel unter Zuleitung von wasserfreiem Chlorwasserstoff geschmolzen war, um die Bildung von basischem Salz zu verhindern, kristallisierte ein Präparat, das in der früheren Arbeit durch Dampfdruckmessungen bei allmählicher Hydratisierung untersucht und in dieser als System (12) bezeichnet wurde. Die Diagramme dieses Präparates zeigen überraschende Übereinstimmung mit dem Diagramm von gesintertem CaCl_2 , enthalten jedoch ausser den Linien, die der Pseudo-Rutilstruktur entsprechen, noch eine kleinere Anzahl von Linien, von denen einige mit niedrigem θ von mittlerer Stärke sind. Es handelt sich nicht um das Diagramm eines Umwandlungsproduktes, das durch Pulverisierung von einmal gebildeten Einkristallen entstanden ist, denn alle Linien, auch die neu hinzugekommenen, lösen sich in wenige Flecken von einzelnen grösseren Kristallen auf, wenn man die Drehung des Präparates im Apparat zum Stillstand bringt. Die Ausmessung der zur Pseudo-Rutilstruktur gehörigen Linien zeigt vollständige Übereinstimmung mit den Diagrammen sowohl der bei 24.2° und 35° hergestellten Präparate wie des gesinterten Calciumchlorids. Bis ganz hinauf zu den Linien, deren $\alpha_1 - \alpha_2$ -Aufspaltung 3–4 mm beträgt, weichen die Diagramme von gesintertem und geschmolzenem Calciumchlorid also nur innerhalb der Messgenauigkeit — ein paar Zehntel mm — von einander ab, so dass also die Gitterkonstanten für die beiden Präparate innerhalb von 1:10000 miteinander übereinstimmen (vergl. BRADLEY¹).

Mit dieser Genauigkeit können wir allerdings die Gitterkonstanten nicht angeben, da wir die Indizierung dieser »hohen Ordnungen« nicht durchgeführt haben.

¹ A. J. BRADLEY, A. H. JAY, Proc. Phys. Soc., **44** (1932), 563 und **45** (1933), 507.

Tabelle 2.

Beobachtete und berechnete Intensitäten im Pulverdiagramm von
gesintertem CaCl_2 , Pseudo-Rutilstruktur.

hkl	beob. Int.	ber. Int.	hkl	beob. Int.	ber. Int.
110	m	82	232	0	0
011	0	0	322	v	10
101	w	17	421	0	7
020	0	1	113	w	25
200	w	22	340	v	12
111	s	180	042	0	1
120	m	64	430	0	0
210	v?	8	150	v	5
121	m—s	53	402 } 142 }	m—s	{ 25 25
211	m—s, s	107	123	0	10
220	m—s	66	213	w	25
002	m—s	52	051	0	1
130	v	5	510	0	9
310 } 221 }	v	{ 4 5	412	0	7
031	m—s	84	341	0	4
112	w	24	332	m	24
301	m	33	431 } 151 }	m	{ 44 13
131	0	1	501	0	1
311	m	41	250	0	1
230	0	0	223	0	2
320	w	16	242	w	16
022	0	0	511	0	10
202	0	6	520 } 422 }	w	{ 4 0
122	m	37	033 }		{ 11
212	0	4	303	v	26
231	0	10	133	w	33
321	0	3	251	0	0
040	0	0	313	v?	14
400 } 140 }	w—m	{ 21 21	521	w	19
222	m—s	50	440	0	1
410	0	6	350	0	0
330	w—m	19	233	0	4
141	w	12	323	0	1
132	0	4	342	v	18
312 } 240 }	w	{ 3 12	441	0	7
411	0	1	530	0	0
331	0	1	432	0	0
420	0	0	152	0	0
013	0	0	060	0	0
241 } 103 }	v	{ 3 4			

Gewisse Linien im Pseudo-Rutildiagramm des geschmolzenen CaCl_2 weisen Intensitäten auf, die sich von den Linien im Diagramm des gesinterten Präparates unterscheiden. Hierbei ist jedoch nur von einem Orientierungseffekt die Rede. Das aus der Schmelze kristallisierte Präparat zerspringt nämlich bei der Zerkrümmung in Stäbchen, in denen ein grosser Teil der Einkristalle ausgeprägt orientiert ist. Wir haben das Diagramm eines derartigen Stäbchens ($5 \times 0.8 \text{ mm}$) mit der gewöhnlichen Pulverkamera aufgenommen. Die Linien $hk0$ traten allzu stark hervor und waren nur in der Mitte geschwärzt (Fiberdiagramm). Die Stäbchen sind also zu einem grossen Teil aus Kristallen, deren c -Achse parallel mit der Achse der Stäbchen verläuft, aufgebaut.

Die fremden Linien, die sich im Diagramm des geschmolzenen Calciumchlorids finden, und die nicht in die quadratische Form von Pseudo-Rutil- CaCl_2 hineinpassen, müssen von anderen Kristallen herrühren, wobei es sich nur um kristalline Modifikationen von CaCl_2 handeln kann, da die chemische Analyse zeigte, dass das Präparat vollständig neutral war und weder basisches Salz noch Wasser enthielt. Die Möglichkeit, dass sich die Linien auf eines der niedrigeren Hydrate des Calciumchlorids zurückführen lassen, wurde selbstverständlich untersucht. Sie stimmen indessen nicht mit den stärksten Linien des Mono- oder Dihydrats überein. Nach der Anzahl und Stärke der Linien zu urteilen, — zwei von ihnen sind recht kräftig, und da das Diagramm schon im voraus sehr dicht mit Linien von normalem CaCl_2 besetzt ist, vermuten wir, dass mehrere Linien der neuen Modifikationen mit anderen Linien zusammenfallen —, kann man schätzen, dass das Präparat mindestens 10% von den neuen Modifikationen enthält.

Wie schon früher erwähnt, vermuten wir, dass VAN BEVER und NIEUWENKAMP in dem Pulver, dessen Diagramm sie nicht indizieren konnten, wo aber die Abweichung vom Rutilgitter »immerhin noch gering« ist, eine ähnliche Mischung vor sich hatten, was wir allerdings nicht beweisen können, da sie in ihrer kurzen Abhandlung nicht die dafür ausreichenden Zahlen mitteilen.

Wird die Mischung einer Wärmebehandlung unterworfen, so verschwinden die fremden Linien nach und nach aus den Röntgendiagrammen. Mehrstündiges Erwärmen auf 50° oder 100° war nicht ausreichend, aber zehntägiges Erwärmen auf 200° brachte

sie beinahe zum Verschwinden. Die Pseudo-Rutilmodifikation ist also bei dieser Temperatur die stabile Form. Nach Versuchen, die im folgenden beschrieben werden, vermuten wir, dass sie von Zimmertemperatur an bis zum Schmelzpunkt die stabile Modifikation ist, dass also die Verhältnisse analog den Verhältnissen für die TiO_2 -Modifikationen sind, wo Rutil unter allen bekannten Bedingungen die stabile Form ist.

Um die Eigenschaften des instabilen CaCl_2 näher zu bestimmen, wurden einige Versuche unternommen, wobei das geschmolzene Salz mit verschiedener Geschwindigkeit abgekühlt wurde, um womöglich ein reines Präparat der instabilen Modifikation zu erhalten. Das misslang jedoch.

Wie schon erwähnt, erhält man durch Schmelzen eines bei ca. 150° vollständig entwässerten CaCl_2 in einem Platintiegel bei gleichzeitiger Zuleitung von wasserfreiem Chlorwasserstoff eine neutrale Schmelzmasse. Mit Hilfe eines auf die Erstarrungstemperaturen von Natriumchlorid (800°) und Kaliumchlorid (770°) geeichten Heraeus Platin-Platinrhodiumthermoelementes in Verbindung mit einem Millivoltmeter für die Temperaturmessungen wurden die Abkühlungskurven für geschmolzenes Calciumchlorid unter verschiedenen Bedingungen bestimmt. Die Versuche ergaben, dass chlorwasserstoffreies und chlorwasserstoffhaltiges Calciumchlorid bei der erwähnten recht groben Versuchsanordnung den gleichen Erstarrungspunkt hatten. Es wurde festgestellt, dass der Chlorwasserstoff bei der Kristallisation ausgetrieben wird. Wiederholtes kurz andauerndes Schmelzen ohne Zuleitung von Chlorwasserstoff bewirkte nur unbedeutende Alkalisierung des Salzes. Die Versuche ergaben, dass der Erstarrungspunkt des reinen Calciumchlorids zwischen 771 und 775° lag, die am häufigsten gefundene Temperatur war 774° ; das stimmt mit den Angaben in der Literatur überein, die, wenn wir von den ältesten, sehr abweichenden Angaben absehen, zwischen 770 und 782° ¹ liegen.

¹ 770° F. C. A. H. LANTSBERY, R. A. PAGE, Journ. Soc. Chem. Industr., **39** (1920), 37.

772° B. KARANDÉEFF, Z. anorg. Chem., **68** (1910), 188.

774° K. ARNDT, Z. f. Elektrochem., **12** (1906), 337.

W. PLATO, Z. physik. Chem., **58** (1907), 350.

G. GRUBE, W. RÜDEL, Z. anorg. allg. Chem., **133** (1924), 375.

R. LORENZ, W. HERZ, Z. anorg. allg. Chem., **145** (1925), 88.

777° O. MENGE, Z. anorg. Chem., **72** (1911), 162.

782° E. KORRENG, N. Jahrb. Min. Bgbd., **37** (1914), 51.

In vielen Fällen liess sich eine Unterkühlung von 20—30° beobachten, aber die Kristallisation ging bei gut definierter Temperatur vor sich. In einem einzigen Fall war der Verlauf ein anderer. Nach erneuter Justierung der thermometrischen Aufstellung, wobei sich wiederum im Platintiegel geschmolzenes Natriumchlorid und Kaliumchlorid befunden hatten, wurde das Calciumchlorid wie vorher geschmolzen und einige Minuten lang auf einer Temperatur von 880° gehalten. Bei der Abkühlung fiel die Temperatur gleichmässig bis auf 692°, wo eine Kristallisation stattfand. Darauf stieg jedoch die Temperatur recht plötzlich auf 756° und schwankte danach unregelmässig zwischen den erwähnten Temperaturen hin und her. Bruchstücke der Schmelzmasse wurden unter Paraffinöl pulverisiert und röntgenographisch untersucht. Die Untersuchung ergab das gewöhnliche Pseudo-Rutildigramm mit Ausnahme einer einzelnen schwachen Linie, die nicht dazu gehörte. Wir haben hier also eine neue Modifikation vor uns gehabt. Jedoch trat leider sofort eine beinahe vollständige Umwandlung in die gewöhnliche stabile Modifikation ein. Zahlreiche spätere Untersuchungen des Erstarrungspunktes der Schmelze ergaben nur die gewöhnlichen Temperaturen um 774°. Es scheint somit schwierig zu sein, auf diese Weise instabiles CaCl_2 darzustellen, was vielleicht auch nicht so merkwürdig ist, wenn man bedenkt, dass das Ausgangsmaterial immer recht gut kristallisiertes, gewöhnliches CaCl_2 ist, das bald alle Geräte mit Kristallkeimen infiziert haben wird.

Alles deutet darauf hin, dass die Umwandlung von instabilem CaCl_2 in die gewöhnliche stabile Modifikation monotrop ist und bei höheren Temperaturen sehr schnell verläuft. Verschiedene Versuche, die Schmelze abzuschrecken, wovon nur ein einzelner näher besprochen werden soll, ergaben immer Präparate, deren Pulverdiagramme einige fremde Linien enthielten. Das Präparat, dessen Diagramm die meisten fremden Linien enthielt, wurde durch Eintauchen einer kalten, 8 mm dicken Nickelstange in geschmolzenes CaCl_2 und schnelles Herausziehen hergestellt. Das Diagramm des erstarrten Calciumchlorids enthielt ausser den 23 kräftigsten der gewöhnlichen Linien (die jedoch alle etwas verwaschen waren, die Kristalle waren also recht schlecht ausgebildet) 15 fremde Linien, davon 11 sichere und 4 zweifelhafte.

Eine Prüfung aller fremden Linien auf einer Reihe von CaCl_2 -Diagrammen zeigte, dass es ausser der gewöhnlichen Form nicht nur eine sondern zwei Modifikationen von CaCl_2 geben muss. Die gemessenen Linien der beiden Modifikationen sind in Tabelle 3 aufgeführt.

Tabelle 3.

Linien, die den instabilen Modifikationen von CaCl_2 angehören.
CuK α -Strahlung.

Modifikation 2		Modifikation 3	
θ	Int.	θ	Int.
9.59	m	15.23	w—m
14.31	s	19.79	w—m
16.00	w—m	23.44	v
16.55	w	25.88	v?
17.38	w	28.63	v
18.51	v	31.59	v?
19.92	m	36.69	v
21.45	v		
23.05	v		
26.31	v?		
26.67	v?		
27.03	v?		
27.82	v?		
33.08	v		

Wir nennen die beiden Modifikationen 2 und 3. Der Beweis dafür, dass es ausser der gewöhnlichen zwei und nur zwei Modifikationen gibt, besteht darin, dass die Intensitäten der Linien, die wir einer der Modifikationen zuschreiben, auf einem willkürlichen Diagramm immer im gleichen gegenseitigen Verhältnis stehen.

Die Anordnung der Linien in zwei getrennten Gruppen wird sehr dadurch erleichtert, dass das auf der Nickelstange abgekühlte Präparat alle gefundenen Linien der Modifikation 2 und nur eine der Modifikation 3 (nämlich die stärkste, die hier nur mit der Intensität v registriert worden war) enthielt. Das auf der Nickelstange abgeschreckte CaCl_2 bestand also grösstenteils aus ordinärem CaCl_2 , einer geringeren Menge von CaCl_2 Modifikation 2 und nur ganz wenig Modifikation 3.

Das früher erwähnte Präparat (12) enthielt in seinem Diagramm 7 Linien der Modifikation 3 und 3 Linien der Modifikation 2, nämlich die drei stärksten. Die Zusammensetzung dieses Präparates war also folgendermassen: zum grössten Teil gewöhnliches, ausserordentlich gut kristallisiertes CaCl_2 , eine kleine Menge der Modifikation 3 und noch weniger von der Modifikation 2.

Die gemessene Linienanzahl scheint leider nicht ausreichend zu sein, um zu entscheiden, ob es sich bei den instabilen Modifikationen um wirkliche Brookit- und Anatas-Strukturen, oder ob es sich um Deformationsstrukturen handelt.

Einige Betrachtungen über den Dehydratationsprozess.

Seit PARTINGTON¹ und LANGMUIR² ihre Auffassung von der Kinetik beim Dehydratationsprozess und dazu analogen Prozessen skizzierten, sind innerhalb dieses Gebietes zahlreiche experimentelle Arbeiten erschienen; von diesen haben viele die Geschwindigkeit, womit Dehydratationsprozesse verlaufen, behandelt, und hierbei sind mehrere interessante Beobachtungen gemacht worden. Über den kinetischen Verlauf des Dehydrationsprozesses selbst sind verschiedene Theorien³ aufgestellt worden, in denen sich in gewissem Grade von einander abweichende Auffassungen geltend machen. Auf Grund dieser Theorien haben wir uns ein Bild von diesen Verhältnissen gemacht, das wir im folgenden darzustellen suchen.

In einem System, das aus einem Hydrat und Wasserdampf besteht, werden sich die Wassermoleküle im gesamten Existenzbereich des Hydrates infolge thermischer Beeinflussung bei konstantem Dampfdruck zwischen den Kristallflächen des Hydrates und der Dampfphase mit gleicher Geschwindigkeit in beiden Richtungen hin und her bewegen. In diesem Bereich wird es bei einer gegebenen Temperatur einen bestimmten Wasserdampf-

¹ J. R. PARTINGTON, J. Chem. Soc., **99** (1911), 466.

² I. LANGMUIR, J. Am. Chem. Soc., **38** (1916), 2263.

³ B. TOPLEY, M. L. SMITH, J. Chem. Soc., (1935), 321.

M. VOLMER, G. SEYDEL, Z. physik. Chem. A., **179** (1937), 153.

J. D. BERNAL, Transact. Faraday Soc., **34** (1938), 981.

J. COLVIN, J. HUME, Transact. Faraday Soc., **34** (1938), 969.

J. COLVIN, - - - - - 983.

W. E. GARNER, - - - - - 981.

J. Y. MACDONALD, - - - - - 983.

druck geben, bei dem eine oder unter gewissen Voraussetzungen alle äusseren Hydratkristallflächen die der Formel des Hydrats entsprechende Zusammensetzung haben. Wird nun eine geringe Menge Wasserdampf aus dem System entfernt, so geben die äusseren Kristallflächen Wassermoleküle ab, und es entsteht eine Oberflächenschicht mit relativem Unterschuss an Wassermolekülen oder mit einem relativen Überschuss an den Ionen, aus denen der Hydratkristall zusammen mit den Wassermolekülen aufgebaut ist; hierdurch stellt sich ein neues Dampfdruckgleichgewicht ein, da jeder Zusammensetzung der Oberflächenschicht ein bestimmter Dampfdruck entspricht. Man muss sich vorstellen, dass die Oberflächenschicht einen beschützenden Einfluss auf die zunächst darunterliegende Kristallfläche ausübt, so dass diese die Wassermoleküle nicht mit der gewöhnlichen Geschwindigkeit aussendet — was sonst eine nach und nach in der Tiefe des Kristalls abnehmende Umbildung zur Folge haben müsste, — sondern nur mit einer Geschwindigkeit, die dem Gleichgewichtsdampfdruck der Oberflächenschicht entspricht.

Fügt man zu dem ursprünglichen System eine geringe Menge Wasserdampf hinzu, so entsteht eine Oberflächenschicht mit relativem Überschuss an Wassermolekülen im Gleichgewicht mit höherem Wasserdampfdruck. Prinzipiell ähnlich wie es hier für ein bestimmtes Hydrat geschildert worden ist, müssen sich die niedrigeren Hydrate des Salzes und zu einem gewissen Grad auch das Anhydrid verhalten.

Die Annahme eines bestimmten Aggregatzustandes in klassischer Bedeutung in einer Oberflächenschicht erscheint uns nicht erforderlich und auch nicht zutreffend zu sein, aber da die Teilchen in der Schicht eine gewisse Beweglichkeit längs den äusseren Kristallflächen, an welche sie elektrostatisch gebunden sind, haben müssen, kann man eine gewisse Ähnlichkeit mit einer Flüssigkeit annehmen.

Entfernt man eine grössere Menge Wasserdampf aus einem Hydrat, so entstehen instabile Oberflächenschichten, in denen die relative Konzentration der Ionen gross ist, und aus denen oft nach grösserer oder geringerer Übersättigung Kristallkeime des niedrigeren Hydrats oder Anhydrids entstehen. Diese Kristallkeime, d. h. Kristallgitter von ganz geringer Ausdehnung, werden, wenn die physikalischen Bedingungen eine ständige Berührung

mit der Oberflächenschicht, die bei der Entwässerung ständig regeneriert wird, gestatten, fortgesetzt Ionen daraus aufnehmen und daher wachsen und derartige Grössen erreichen, dass das System, wenn man es sich selbst überlässt, einen definierten Gleichgewichtsdampfdruck haben wird. Wenn das Gleichgewicht in einem solchen System erreicht worden ist, müssen die zusammenstossenden Oberflächenschichten der beiden festen Phasen energetisch identisch und dampfdruckbestimmend sein.

Der Dehydratations-Rehydratationsprozess wird reversibel verlaufen, wenn das niedrigere Hydrat oder das Anhydrid ständig die Verbindung mit der Oberflächenschicht des höheren Hydrats in ausreichendem Umfang bewahrt, da man nicht annehmen kann, dass die Ionen sich ausserhalb derselben bewegen können — jedenfalls nicht in einem Ausmass, das irgendwelche Bedeutung haben kann. Hört während einer Dehydratation der Kontakt zwischen den Kristallen des höheren und des niedrigeren Hydrats (Anhydrids) auf, oder ist die Geschwindigkeit, womit die letzteren imstande sind zu wachsen d.h. Ionen von der Oberflächenschicht aufzunehmen, ausserordentlich gering, so ist kein Abfluss für die Ionen da. Dadurch steigert sich der Übersättigungsgrad für die Ionen in Beziehung auf das niedrigere Hydrat oder das Anhydrid, und der Wasserdampfdruck nimmt gleichzeitig ab, wodurch die Entwässerungsgeschwindigkeit vermindert wird. Aus der stark übersättigten Oberflächenschicht können so zahlreiche neue Kristallkeime entstehen, die vielleicht nicht eine solche Grösse erreichen, dass sie imstande sind ein Röntgendiagramm zu geben; man erhält so ein amorphes Entwässerungsprodukt mit einem entsprechend grossen Inhalt an freier Energie und somit Systeme mit abnormer Hydratationswärme, deren Gleichgewichtsdampfdrucke infolgedessen bei niedrigerer Temperatur abnorm niedrig sind.

Der Dehydratationsprozess $\text{MnC}_2\text{O}_4 \cdot 2\text{H}_2\text{O} \rightarrow \text{MnC}_2\text{O}_4 + 2\text{H}_2\text{O}$ ist von TOPLEY und SMITH¹ und später noch von VOLMER und SEYDEL² untersucht worden. Topley und Smith bestimmten bei 76°, wo der Gleichgewichtsdampfdruck 135 mm beträgt, die Geschwindigkeit des Prozesses bei wechselndem Wasserdampfdruck und beobachteten ein Geschwindigkeitsminimum bei

¹ B. TOPLEY, M. L. SMITH, *Nature*, **128** (1931), 302. *J. Chem. Soc.*, (1935), 321.

² M. VOLMER, G. SEYDEL, *Z. physik. Chem. A*, **179** (1937), 153.

0.1 mm und ein Maximum bei ungefähr 1 mm Wasserdampfdruck. VOLMER und SEYDEL wiederholen TOPLEY und SMITH'S Versuche und bestätigen ihre Beobachtungen, kritisieren jedoch ihre Auslegung und zeigen auf, dass das im Hochvakuum entwässerte Manganooxalat röntgenographisch amorph wird, und dass die Hydratationswärme dieses Hochvakuumproduktes um ca. 2000 cal pr. Mol grösser ist als die Werte, die sie für kristallisierte Entwässerungsprodukte, die unter Anwesenheit von Wasserdampf bei höherem Druck entstanden sind, finden.

KOHLSCHÜTTER und NITSCHMANN¹ erhalten bei Entwässerung von $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ im Hochvakuum bei 20 und 40° als gewichtskonstantes Produkt ein röntgenographisch völlig amorphes Monohydrat, während die Entwässerung von $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ im Trockenschrank bei 100° ein gut kristallisiertes Monohydrat ergibt.

Wenn die Kristallkeime wie hier in einigen Fällen extrem langsam wachsen, könnte das darauf beruhen, dass sich zu wenige Ionen an einer solchen Stelle befinden, wo sie einen Platz in der angefangenen Gitterschicht einnehmen könnten; setzt man aber eine wirkliche Übersättigung voraus, so ist die Ursache eher darin zu suchen, dass die Ionen bei der relativ tiefen Temperatur und vielleicht auch bei den speziellen Verhältnissen des Mediums seltener so grosse thermische Energie erlangen, dass sie ausreicht um die Potentialbarriere, die die Kristallkeime umgibt, zu forcieren. Es ist früher die Wahrscheinlichkeit dafür erbracht worden², dass diese Barriere desto höher ist, je härter der Kristall ist, eine Eigenschaft, die im allgemeinen in naher Beziehung zu relativ hohem Schmelzpunkt und grosser Gitterenergie steht.

Es ist naheliegend zu folgern, dass es mindestens drei Faktoren von besonderer Bedeutung für die Reversibilität des Dehydratationsprozesses gibt:

- 1) Aufrechterhaltung eines nicht zu niedrigen Wasserdampfdruckes während der Entwässerung, wodurch grosse Übersättigung vermieden wird.
- 2) Die Härte des entstandenen niedrigeren Hydrats oder Anhydrids darf nicht zu gross sein.

¹ V. KOHLSCHÜTTER, H. NITSCHMANN, Z. physik. Chem. A, Bodenst.-Festbd. (1931), 494.

² A. TOVBORG JENSEN, Z. physik. Chem. A, 180 (1937), 93.

- 3) Das Verhältnis zwischen der Dehydratationstemperatur und dem Schmelzpunkt oder der Umwandlungstemperatur des Dehydratationsproduktes in der absoluten Temperaturskala darf nicht zu klein sein.

Demnach sollte man a priori erwarten, dass die letzten Entwässerungsstufen, die in der Bildung der Anhydride resultieren, weniger reversibel als die vorhergegangenen Stufen sein sollten, wo höhere Hydrate bloss zu niedrigeren entwässert werden. Die wirklichen Verhältnisse zeigen jedoch, dass eine derartige Verallgemeinerung zu weit geht.

Das System $\text{NaBr}, 2\text{H}_2\text{O} + \text{NaBr}$ wurde während der vorliegenden Arbeit näher untersucht. Die Entwässerung von $\text{NaBr}, 2\text{H}_2\text{O}$ geschieht im Vakuum bei Zimmertemperatur mit grosser Schnelligkeit unter Bildung von NaBr , das sich bei röntgenographischer Untersuchung als gut kristallisiert erwies, indem die α_1 — α_2 -Aufspaltung ebenso vollkommen wie bei einem Präparat von sorgfältig wärmebehandeltem NaCl -Pulver war. Der Gleichgewichtsdampfdruck stellt sich relativ schnell ein, und es ergaben sich bei 0, 18, 25 und 30° folgende Dampfdrucke: 1.20, 5.16, 8.72 und 12.49 mm. Die Entwässerung dieses Präparats war bei 18° vorgenommen worden. Die Messungen wurden mit Präparaten von verschiedenem Wassergehalt, die einige Zeit gestanden hatten, wiederholt. Die grössten Abweichungen innerhalb dieser Versuche betragen weniger als 1 % der gemessenen Werte. Nach 24 stündigem Sintern des Entwässerungsproduktes bei 300° wurden für die entsprechenden Temperaturen folgende Werte gefunden: 1.20, 5.14, 8.70 und 12.43 mm. In einer früheren Arbeit¹ wurden bei 18° 5.15 ± 0.05 mm gefunden.

Die Entwässerung von $\text{NaBr}, 2\text{H}_2\text{O}$ verläuft so schnell, dass der Druck trotz Anwesenheit von wasserentziehenden Mitteln so hoch bleibt, dass extreme Übersättigung vermieden wird. Die Härte von Natriumbromid ist relativ klein, nach REIS und ZIMMERMANN² von der Ritzhärte 2.0 oder von der gleichen Härte wie rhombischer Schwefel, Kaliumchlorid, Silberchlorid und

¹ AXEL LANNUNG, Z. physik. Chem. A, **170** (1934), 134.

J. N. BRONSTED und A. PETERSEN, D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd. I,3 (1917), fanden bei 20° einen Dampfdruck, der 50,2 mm Bromnaphthalinsäule oder umgerechnet 5.47 mm Hg entsprach.

H. LESCOEUR, Ann. chim. phys., **19** (1890), 533, fand bei 10, 20 und 30° 3.9, 6.7—7.6 und 15.5 mm.

² A. REIS, L. ZIMMERMANN, Z. physik. Chem., **102** (1922), 298.

Merkurichlorid, während zum Vergleich Natriumchlorid die Ritzhärte 3.5, $\text{CuSO}_4, 5 \text{H}_2\text{O}$ 7.8—8.1 und Kalkspat 7.5—15.2 haben.

Das Verhältnis zwischen Entwässerungstemperatur und Schmelzpunkt in der absoluten Temperaturskala ist ungefähr 0.3, also nach TAMMANN¹ nicht günstig für eine Rekrystallisation. Die geringe Härte von NaBr scheint demnach von überragendem Einfluss für den reversiblen Verlauf des Prozesses zu sein.

Es fehlen Daten, um die Verhältnisse bei der Dehydratation von $\text{MnC}_2\text{O}_4, 2 \text{H}_2\text{O}$ und $\text{CuSO}_4, 5 \text{H}_2\text{O}$ von den oben angeführten Gesichtspunkten aus völlig zu beleuchten. Über die Härte von MnC_2O_4 haben wir keine Angaben in der Literatur gefunden, jedoch sind die Härten von anderen untersuchten Oxalaten recht gross. $\text{CuSO}_4, 5 \text{H}_2\text{O}$ hat für ein höheres Hydrat eine beträchtliche Härte, und man muss erwarten, dass die Härte des Monohydrats so gross ist, dass sie von überwiegendem Einfluss auf die Reversibilität des Prozesses sein wird.

Wir wollen nunmehr die Verhältnisse bei der Dehydratation von $\text{CaCl}_2, \text{H}_2\text{O}$ betrachten. Die in der früheren Arbeit gefundenen und hier in der Einleitung kurz erwähnten Eigentümlichkeiten des Dampfdrucks von $\text{CaCl}_2\text{—CaCl}_2, \text{H}_2\text{O}$ -Systemen, die durch teilweise Entwässerung des Monohydrats bei Temperaturen zwischen 180 und 18° hergestellt worden sind, lassen sich, wie hier röntgenographisch nachgewiesen wurde, im wesentlichen auf Unterschiede in der Teilchengrösse des Anhydrids zurückführen. Geht man von höherer zu tieferer Temperatur, so erhält man einen gleichmässigen Übergang von relativ gut kristallisiertem CaCl_2 zu Entwässerungsprodukten mit immer kleineren Partikelgrössen und stark steigenden Hydratationswärmen. Der Wasserdampfdruck während der Entwässerung variierte von 760 mm bei 180° bis zu $< 10^{-4}$ mm bei 18°. Die Entwässerungsgeschwindigkeiten nahmen im gleichen Gebiet mit einem Faktor von mindestens 10^6 (extrapoliert) ab.

Wir haben die Härte von geschmolzenem CaCl_2 untersucht und gefunden, dass sie geringer als die des Kalkspats, aber grösser als die von $\text{CuSO}_4, 5 \text{H}_2\text{O}$ ist. Der Schmelzpunkt ist wie erwähnt 774°. Das Verhältnis zwischen Entwässerungstemperatur und Schmelzpunkt variiert also in diesem Gebiet von 0.43—0.28. Man hat hier offenbar ein Gebiet vor sich, in dem

¹ G. TAMMANN, Z. anorg. allg. Chem., **157** (1926), 321.

sich die Faktoren gradweise derartig verändern, dass sie sich in steigendem Masse zu Ungunsten der Reversibilität dieses Entwässerungsprozesses auswirken.

Dass die gemessenen Dampfdrucke jedes einzelnen Präparates trotz alledem ziemlich reproduzierbar sind, deutet darauf hin, dass diese instabilen Entwässerungsprodukte innerhalb jedes einzelnen Präparates von gleichartiger Teilchengrösse sind. Übrigens sind diese energiereichen, instabilen Formen von CaCl_2 jahrelang haltbar.

Betrachten wir schliesslich die Hydratation von geschmolzenem CaCl_2 , so sind die Verhältnisse bei dem früher untersuchten Präparat (12) etwas kompliziert. Dies Präparat enthielt ausser gewöhnlichem und besonders gut kristallisiertem CaCl_2 geringere Mengen der beiden anderen CaCl_2 -Modifikationen. Die Dampfdruckverhältnisse eines derartigen Systems sollten ganz instabil sein. Bei den Messungen, die seinerzeit vorgenommen wurden, zeigte es sich denn auch, dass der Dampfdruck des Systems nach der anfänglich zugesetzten geringen Menge Wasserdampf, die weniger als 0.01 Mol H_2O entsprach, schwer reproduzierbar war, was damals durch Anwesenheit von zu wenig Monohydrat erklärt wurde. Aus diesem Grunde konnte auch die Hydratationswärme nicht mit grösserer Genauigkeit angegeben werden. Nach Zuleitung von mehr Wasserdampf entstanden Systeme mit reproduzierbaren Gleichgewichtsdampfdrucken, die sich jedoch sehr langsam einstellten. Die aus diesen Messungen berechneten Hydratationswärmen waren um etwa 1 kcal grösser als für gesintertes CaCl_2 .

Wie die Hydratation eines solchen, aus groben, zusammengesinterten Kristallmassen gebildeten Systems verläuft, ist schwer zu entscheiden, aber es ist wenig wahrscheinlich, dass die relativ geringen Mengen der instabilen Modifikationen bis zu einem Wassergehalt von 0.83 Mol H_2O /Mol CaCl_2 hinauf dampfdruckbestimmend sein könnten. Von gewisser Bedeutung kann es sein, dass die Kristalle des gewöhnlichen CaCl_2 etwas einseitig nach der c-Achse entwickelt waren. Besonders interessant war es, dass sich dies System nach Aufnahme der 0.83 Mol H_2O in nicht geringem Umfang reversibel bei 50° entwässern liess. Erst nachdem ungefähr 0.4 Mol Wasser entfernt worden waren, entstand ein energiereicheres Entwässerungsprodukt mit der berechneten Hydratationswärme ($\text{CaCl}_2 + \text{H}_2\text{O}_{\text{fl.}} \rightarrow \text{CaCl}_2 \cdot \text{H}_2\text{O}$) 8.7 kcal.

In Übereinstimmung mit unserem Bild lässt sich das dadurch erklären, dass die Ionen bei der Dehydratation des Monohydrats anfänglich auf den Überresten der ursprünglich grossen Calciumchloridkristalle Platz finden, so dass der Kontakt zwischen diesen und den Monohydratkristallen durch die Oberflächenschicht bewahrt bleibt. Später wird er unterbrochen, zahlreiche neue Keime entstehen, aber erreichen nur geringe Grösse.

Der besonders hohe Dampfdruck über Systemen, die aus geschmolzenem CaCl_2 hergestellt worden sind, ist in der vorliegenden Arbeit durch Untersuchung von geschmolzenem CaCl_2 , dessen Erstarrungspunkt zu 774° bestimmt wurde, verifiziert worden. Für ein System von der Zusammensetzung $\text{CaCl}_2, 0.77\text{H}_2\text{O}$ waren die Dampfdrucke bei $18, 35$ und 50° $0.36, 1.38$ und 3.69 mm, woraus sich die Hydratationswärme ($\text{CaCl}_2 + \text{H}_2\text{O}_{fl.} \rightarrow \text{CaCl}_2, \text{H}_2\text{O}$) zu 3.1 kcal berechnen lässt. Diese Dampfdrucke sind etwa 8 Mal so gross wie diejenigen, die über Systemen, die aus mikrokristallinem gesintertem CaCl_2 hergestellt worden waren, gemessen wurden. Von unserem Bild des Hydratationsprozesses ausgehend erklärt sich dieser grosse Unterschied daraus, dass die Ionen in den räumlich am weitesten ausgedehnten Gittern mit der grössten Kraft festgehalten werden, wodurch die Oberflächenschicht relativ wasserreicher wird und folglich der Gleichgewichtsdampfdruck entsprechend grösser. Vgl. die Löslichkeit von grossen und kleinen Kristallen.

Zusammenfassung.

In einer früheren Arbeit über Systeme von $\text{CaCl}_2 - \text{CaCl}_2, \text{H}_2\text{O}$ wurde aufgezeigt, dass der Gleichgewichtsdampfdruck über derartigen Systemen bei 18° je nach der Vorgeschichte der Gleichgewichtssysteme im Verhältnis 1 zu 250 variieren konnte.

1) Um diesen Umstand näher zu beleuchten, wurde in der vorliegenden Arbeit eine röntgenographische Untersuchung einer Reihe von unter ganz verschiedenen Umständen dargestellten Präparaten von CaCl_2 und $\text{CaCl}_2, \text{H}_2\text{O}$ und Mischungen von beiden vorgenommen. Diese Untersuchung zugleich mit den Resultaten der früheren Arbeit zeigt, dass die Hydratphase in allen Präparaten immer die gleiche ist, und dass der Unterschied zwischen den untersuchten Systemen im Anhydrid zu suchen ist, das ganz

verschiedene Zustände und deshalb auch verschiedene Energieinhalte und Affinitäten zu Wasser haben kann.

2) Das stabile Gitter für CaCl_2 scheint im ganzen untersuchten Temperaturgebiet ein rhombisches Pseudo-Rutilgitter zu sein.

3) Stellt man CaCl_2 durch Entwässerung von $\text{CaCl}_2 \cdot \text{H}_2\text{O}$ bei niedriger Temperatur her, so erhält man sehr kleine Partikel, und je niedriger die Entwässerungstemperatur ist, desto grössere Affinität erhält der Stoff zum Wasser. Die Entwässerung im Hochvakuum nahe der Zimmertemperatur geht ungeheuer langsam vor sich und führt bei 35° zu Teilchen, deren Dicke wir zu 163 \AA bestimmt haben. Bei 24.2° erhält man Teilchen von der Dicke 117 \AA .

4) Die molare Hydratationswärme ($\text{CaCl}_2 + \text{H}_2\text{O}_{\text{fl.}} \rightarrow \text{CaCl}_2 \cdot \text{H}_2\text{O}$) wurde für ein bei 35° hergestelltes Präparat zu 10.6 kcal berechnet; durch Interpolation findet man für ein bei 24.2° hergestelltes Präparat 12.6 kcal , während der Wert für ein mikrokristallines bei 400° gesintertes Präparat 2.5 kcal ist. Diese Energieüberschüsse der feinteiligen Präparate im Betrage von 8.1 bzw. 10.1 kcal/Mol haben wir mit der unausgenutzten Gitterenergie der Ionen in der Oberfläche der Kristalle identifiziert. In einigen gitterdynamischen Überlegungen haben wir durch vom NaCl -Gitter aus gezogene Analogieschlüsse versucht, eine Klarlegung des Wertes dieses Energieüberschusses zu geben.

5) Röntgenuntersuchungen zeigen, dass geschmolzene, aber nicht abgeschreckte Präparate von CaCl_2 immer hauptsächlich aus relativ grossen (Dicke $> 2000 \text{ \AA}$) Partikeln des Pseudo-Rutilgitters, oft mit geringeren Beimischungen von einer oder zwei neuen instabilen CaCl_2 -Modifikationen bestehen; von diesen haben wir durch Untersuchung von abgeschreckten Präparaten von CaCl_2 eine Reihe von Debye-Linien nachgewiesen, worin diese neuen Modifikationen bisweilen in recht grossen Konzentrationen auftreten.

6) In einigen allgemeinen Betrachtungen über den Mechanismus von Hydratations- und Dehydratationsprozessen wird auf drei Faktoren von besonderer Bedeutung für die Reversibilität hingewiesen:

a) Aufrechterhaltung eines nicht zu niedrigen Dampfdruckes bei der Entwässerung.

- b) Dass die Härte des entstandenen niedrigeren Hydrats oder Anhydrids nicht zu gross ist.
- c) Dass das Verhältnis von Dehydratationstemperatur zu dem Schmelzpunkt (Umwandlungspunkt) des Dehydratationsproduktes in der absoluten Temperaturskala nicht zu klein ist.
- 7) Der Gleichgewichtsdampfdruck über dem System $\text{NaBr} + \text{NaBr} \cdot 2 \text{H}_2\text{O}$ ist bei verschiedenen Temperaturen gemessen worden.
- 8) Die Vorgänge bei der Dehydratation von $\text{NaBr} \cdot 2 \text{H}_2\text{O}$, $\text{MnC}_2\text{O}_4 \cdot 2 \text{H}_2\text{O}$, $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ und $\text{CaCl}_2 \cdot \text{H}_2\text{O}$ sind von den erwähnten Gesichtspunkten aus diskutiert worden.
-

Es ist uns eine Freude, dem Direktor der Kgl. Tierärztlichen und Landwirtschaftlichen Hochschule Herrn Prof. Dr. phil. N. BJERRUM unsern besten Dank für seine Anregung und sein beständiges Interesse für diese Zusammenarbeit darzubringen.

Ferner sind wir Herrn Prof. Dr. phil. EBBE RASMUSSEN für seine liebenswürdige Unterstützung bei der Aufnahme der Registrierkurven zu Dank verpflichtet.

*Kopenhagen. Apotheke und Chemisches Laboratorium
der Kgl. Tierärztlichen und Landwirtschaftlichen Hochschule.*

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 16

ON THE CONTINUOUS ABSORPTION
OF THE NEGATIVE HYDROGEN ION
IN THE ULTRAVIOLET

BY

MOGENS RUDKJØBING



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

Through recent work on the theory of stellar atmospheres^{1, 2, 3, 4} it has become clear that the continuous absorption in atmospheres of the solar type is mainly due to the negative hydrogen ion, H^- .

Making use of HYLLERAAS'⁵ calculation of the ionisation potential of the negative hydrogen ion, and of the estimates of its continuous absorption by JEN⁶, and by MASSEY and SMITH⁷, WILDT showed that H^- probably contributed most of the continuous absorption in the case of the solar atmosphere. After the astrophysical importance of H^- had been revealed by WILDT's discovery, MASSEY and BATES⁸ carried out a much more accurate calculation of the continuous absorption coefficient of H^- as a function of the wave length. The results of this calculation were used by WILDT², and by B. STRÖMGREN³ in work on model solar atmospheres according to the revised concept of the mechanism of continuous absorption.

The continuous absorption coefficient of the negative hydrogen ion has hitherto been calculated only for transitions to states in which the neutral atom left by the electron is in the ground-state. The fact that the integrated f -sum for these transitions is considerably smaller than the theoretical value 2 indicates, however, that we have to consider other transition processes also, namely such which leave the atom in an excited or even ionised state. Transitions of this kind, having their absorption edges in the ultraviolet, have little influence on the value of the opacity

¹ R. WILDT, *Ap. J.* **89**, 295, 1939.

² R. WILDT, *Ap. J.* **90**, 611, 1939.

³ B. STRÖMGREN, *Festschrift für Elis Strömgren*, p. 218, 1940.

⁴ P. TEN BRUGGENCATE, *V. J. S.* **75**, 203, 1940.

⁵ E. HYLLERAAS, *Zs. f. Phys.* **60**, 624, 1930.

⁶ C. K. JEN, *Phys. Rev.* **43**, 540, 1933.

⁷ H. S. W. MASSEY and R. A. SMITH, *Proc. Roy. Soc. A* **155**, 472, 1936.

⁸ H. S. W. MASSEY and D. R. BATES, *Ap. J.* **91**, 202, 1940.

of the matter in the solar atmosphere, provided that there are no important deviations from the black body curve of emission in the far ultraviolet. But in atmospheres where the temperature is higher and the electron pressure is great enough to balance the ionisation of the negative ions due to the higher temperature, this absorption which is partly placed in the spectral "window" on the long wave length side of the Lyman continuum may have some influence.

The aim of the present work has been the calculation of the transition probabilities mentioned, to see if any of them were large enough to be responsible for the theoretically expected remainder of the total absorption.

The formulae used for the transition probabilities and the equations determining the eigenfunctions (Fock's equations) have been taken from D. R. HARTREE and W. HARTREE'S work on calcium¹. The cases are analogous, as in both cases we have to deal with two valence electrons.

In the final state of the transition considered we have to deal with one s - and one p -electron (one of them bound and the other free). We therefore get the following two simultaneous Fock-equations which correspond to HARTREE'S equations for calcium neglecting the interaction of the valence electrons with the inner, closed shells (with n between 1 and 3). In HARTREE'S notation—with indices s and p only instead of $4s$ and $4p$ —the two equations are as follows:

$$\left. \begin{aligned} \left[\frac{d^2}{dr^2} + \frac{2 - 2Y_0(p,p)}{r} - \epsilon_s \right] \cdot P(s) - \frac{2}{3} \frac{Y_1(s,p)}{r} \cdot P(p) &= 0; \\ \left[\frac{d^2}{dr^2} + \frac{2 - 2Y_0(s,s)}{r} - \epsilon_p - \frac{2}{r^2} \right] \cdot P(p) - \frac{2}{3} \frac{Y_1(s,p)}{r} \cdot P(s) &= 0. \end{aligned} \right\} (1)$$

The negative sign of the last term in the equations is due to the fact that we are dealing with singlet states only. ϵ is the energy of the electron with its sign reversed. It is measured in units of the Rydberg-constant equal to the ionisation energy of

¹ D. R. HARTREE and W. HARTREE, Proc. Roy. Soc. A **164**, 167, 1938.

neutral hydrogen. (In the present work we have made no distinction between the Rydberg-constant for hydrogen and for infinite mass, the difference being wholly negligible in our calculations).

$$Y_k(a, b) = \int_{r_1=0}^{r_1=r} P(a) \cdot P(b) \cdot \left(\frac{r_1}{r}\right)^k dr_1 + \int_{r_1=r}^{r_1=\infty} P(a) \cdot P(b) \cdot \left(\frac{r}{r_1}\right)^{k+1} dr_1. \quad (2)$$

In the case that one of the two electrons is free the above equation (1) for the other electron is reduced to the simple SCHRÖDINGER-equation for a neutral hydrogen atom. The equation for the free electron itself undergoes no similar reduction, as it is homogenous in the unknown eigenfunction so that all its terms are reduced in the same way when the state of the electron approaches the free state¹. Instead of two simultaneous equations we get only one in which beside the required eigenfunction for the free electron we have the well-known hydrogen eigenfunction for the considered stationary state of the atom.

Between the two equivalent *s*-electrons in the negative ion there are no exchange forces in this approximation used by HARTREE. Its two identical *s*-functions may therefore be determined by the ordinary self consistent field-method. These eigenfunctions are used throughout the present work because HARTREE's formulae for the transition probabilities can be utilised directly for such eigenfunctions.

For the transition probability we use HARTREE's formula for the transition between the two lowest states of neutral calcium, this transition being exactly analogous to the transitions of the hydrogen ion here considered.

The formula for the transition probability *A* is²

$$A = 2.66 \cdot 10^8 \cdot \frac{1}{3} \cdot \left(\frac{\tilde{\nu}}{R}\right)^3 \cdot S, \quad \left. \begin{array}{l} \text{where} \\ S = 2 \cdot \left[\int_0^\infty P_A(s) \cdot P_B(s) dr \right]^2 \cdot \left[\int_0^\infty r \cdot P_A(s) \cdot P_B(p) dr \right]^2 \end{array} \right\} \quad (3)$$

¹ H. BETHE: Handbuch der Physik, Kap. 3, 350, 1933.

² D. R. HARTREE and W. HARTREE, Proc. Roy. Soc. A **164**, 167, 1938.

r is—as in the equations—measured in atomic units. The indices A and B refer to the ion-state and the state with one s - and one p -electron respectively. $\tilde{\nu}$ is the wave-number of the absorbed radiation and R the Rydberg-constant. The number 3 in the denominator in the formula for A is the statistical weight of the upper state, this being a singlet P -state.

Between A and the oscillator-strength of absorption f we have LADENBURG'S relation:

$$A = \frac{g}{g'} \cdot \frac{8\pi^2 e^2}{3mc^2} \cdot \nu^2 \cdot f = \frac{g}{g'} \cdot 3 \cdot \left(\frac{\tilde{\nu}}{R}\right)^2 \cdot f \cdot 2.66 \cdot 10^8 \text{ sec}^{-1}. \quad (4)$$

Here g' and g denote the statistical weights of the upper and the lower state respectively.

In our case, when we have to deal with continuous upper states, we are interested in $\frac{df}{dE}$, the oscillator strength per interval of energy. The eigenfunctions for the free electrons must then be normalised in such a way that their limiting amplitudes are $\pi^{-\frac{1}{2}} \cdot E^{-\frac{1}{4}}$, where E is the energy of the electron¹. The formula for $\frac{df}{dE}$ is then the same as for f and is written:

$$\frac{df}{dE} = \frac{1}{3} \cdot \frac{\tilde{\nu}}{R} \cdot 2 \cdot \left[\int_0^\infty P_A(s) \cdot P_B(s) dr \right]^2 \cdot \left[\int_0^\infty r \cdot P_A(s) \cdot P_B(p) dr \right]^2, \quad (5)$$

where the eigenfunction of the free electron (being an s - or an p -electron if the atom is in a P - or a S -state respectively) is normalised in the way described above.

The energy of the absorbed radiation is the sum of the ionisation energy of the ion, the excitation energy of the atom and the energy of the free electron.

The absorption coefficient κ_ν is $8.06 \cdot 10^{-18} \cdot \frac{df}{dE}$.² The transition probabilities for the transitions to the ground state of the atom calculated in the way here described may be less accurate than the corresponding probabilities as calculated by MASSEY and BATES³ because the latter authors used HYLLEAAS' better

¹ J. HARGREAVES, Proc. Cambridge Phil. Soc. **15**, 75, 1929.

² A. UNSÖLD: Physik der Sternatmosphären (35,1), 1938.

³ H. S. W. MASSEY and D. R. BATES, Ap. J. **91**, 202, 1940.

approximation for the ion-eigenfunction. The degree of agreement between the two sets of results may, however, serve as an indication of the accuracy of those of our results, which cannot be compared with earlier calculations, in so far as the inaccuracy is due to the uncertainty of the ion-eigenfunction.

The eigenfunction for the negative hydrogen-ion was found by successive approximations, the Z_p -function used for an approximation being calculated by means of the preceding approximation. As a rule an eigenfunction intermediate between the two last approximations was used for the calculation of the new field function, and a corresponding average field function was used for the calculation of a new eigenfunction. In this way the oscillations in the functions from one approximation to the next were considerably damped. The eigenfunction as integrated from the origin was joined with an exponential function outside the range of the Coulomb field. The difference between the energy parameters in the inner and the outer part of the function was a measure of the error in the approximation. A new field function was not calculated until an energy parameter had been found, which made the eigenfunction satisfy the two boundary conditions simultaneously. The energy parameter—not to be confused with the ionisation potential of the ion—was in the final approximation found to be 0.093 Rydberg units.

The normalised eigenfunction P for one electron in the ion found in the way here described is given in Table I together with $r \cdot P$ and the corresponding $2Z_p = 2 - 2Y_0$. The last figure given has in most cases an uncertainty amounting to one or two units.

The eigenfunctions of the free electrons were calculated by means of successive approximations, the first being the solution of the equations without the exchange term. When one solution had been found, the exchange integrals were calculated and used for the next approximation, which could always be started as the solution without exchange sufficiently near the origin. The calculation was carried out by numerical integration. For the first integration—as for all the integrations of the ion-eigenfunctions—NUMEROV'S method¹ could be used since the

¹ B. NUMEROV: Méthode nouvelle de la détermination des orbites et le calcul des éphémérides en tenant compte des pertubations. Ch. II.

Table I.

r	P	$r \cdot P$	$2 Z_p$	r	P	$r \cdot P$
0.0	0.0000	0.000	2.000	10.0	0.0507	0.507
0.1	.1263	..	1.863	10.5	436	.458
0.2	.2291	..	1.730	11.0	374	.411
0.3	.3122	..	1.601	11.5	321	.369
0.4	.3790	..	1.481	12.0	276	.331
0.5	.4320	.216	1.365	12.5	237	.296
0.6	.4736	..	1.257	13.0	203	.264
0.7	.5057	..	1.157	13.5	175	.236
0.8	.5299	..	1.066	14.0	150	.210
0.9	.5475	..	0.980	14.5	129	.187
1.0	.5596	.560	0.902	15.0	110	.165
1.1	.5672	..	0.829	15.5	95	.147
1.2	.5711	..	0.763	16.0	81	.130
1.3	.5718	..	0.701	16.5	70	.116
1.4	.5700	..	0.645	17.0	60	.102
1.5	.5660	.849	0.592	17.5	52	91
1.6	.5604	..	0.547	18.0	44	79
1.7	.5533	..	0.502	18.5	38	70
1.8	.5452	..	0.462	19.0	33	63
1.9	.5361	..	0.427	19.5	28	55
2.0	.5263	1.053	0.392	20.0	24	48
				20.5	21	43
				21.0	18	38
2.0	0.5263	1.053	0.392	21.5	15	33
2.5	.4713	1.178	0.266	22.0	13	29
3.0	.4143	1.243	0.175	22.5	11	25
3.5	.3607	1.262	0.119	23.0	10	22
4.0	.3122	1.249	0.080	23.5	8	20
4.5	.2694	1.212	0.055	24.0	7	17
5.0	.2320	1.160	0.037	24.5	6	15
5.5	.1996	1.098	0.026	25.0	5	13
6.0	.1716	1.030	0.017	25.5	4	11
6.5	.1474	.958	0.012	26.0	4	10
7.0	.1266	.886	0.008	26.5	3	9
7.5	.1087	.815	0.005	27.0	3	8
8.0	934	.747	0.004	27.5	2	7
8.5	802	.682	0.002	28.0	2	6
9.0	688	.619	0.001	28.5	2	5
9.5	591	.561	..	29.0	2	4
10.0	507	.507	..	29.5	1	4
..	30.0	1	3

Table II.

State of Atom	Radial eigenfunction	$\frac{2 Z_p}{r} = \frac{2 - 2 Y_0}{r}$
1 S	$2 r \cdot e^{-r}$	$\left(2 + \frac{2}{r}\right) \cdot e^{-2r}$
2 S	$\frac{1}{\sqrt{2}} \cdot r \cdot e^{-\frac{r}{2}} \left(1 - \frac{1}{2} r\right)$	$\left(\frac{r^2}{4} + \frac{r}{2} + \frac{3}{2} + \frac{2}{r}\right) \cdot e^{-r}$
2 P	$\frac{1}{2\sqrt{6}} \cdot r^2 \cdot e^{-\frac{r}{2}}$	$\left(\frac{r^2}{12} + \frac{r}{2} + \frac{3}{2} + \frac{2}{r}\right) \cdot e^{-r}$
3 S	$\frac{2}{3\sqrt{3}} \cdot r \cdot e^{-\frac{r}{3}} \left(1 - \frac{2}{3} r + \frac{2}{27} r^2\right)$	$\left(\frac{8}{2187} r^4 - \frac{8}{729} r^3 + \frac{8}{81} r^2 + \frac{8}{27} r + \frac{10}{9} + \frac{2}{r}\right) \cdot e^{-\frac{2r}{3}}$
3 P	$\frac{8}{27\sqrt{6}} \cdot r^2 \cdot e^{-\frac{r}{3}} \left(1 - \frac{1}{6} r\right)$	$\left(\frac{4}{2187} r^4 + \frac{4}{729} r^3 + \frac{4}{81} r^2 + \frac{8}{27} r + \frac{10}{9} + \frac{2}{r}\right) \cdot e^{-\frac{2r}{3}}$

ratio of the second derivative of the eigenfunction to the function itself was here a function of r alone.

The normalised radial eigenfunctions of the neutral hydrogen atom and the corresponding potential functions $\frac{2 Z_p}{r} = \frac{2 - 2 Y_0}{r}$ are given in Table II for the states considered. The series expansions of the $\frac{2 Z_p}{r}$ -functions near the origin are $\frac{2}{r} - 2, \frac{2}{r} - \frac{1}{2}, \frac{2}{r} - \frac{2}{9}$ for the 1-, 2- and 3-states respectively. For the S-states the first neglected term is of the second, for the P-states of the fourth order in r . In the case considered by MASSEY and BATES with the atom in the 1 S-state they found that the total distortion of the eigenfunctions due to electrostatic and exchange forces was negligible. This result was confirmed by the examination of one function (with $E = \frac{1}{2}$), the total phase difference for $r \rightarrow \infty$ between the exact solution and the field-free solution amounting to 6 degrees only. The phase effect of the exchange forces was in this case nearly three times the effect of the electrostatic forces and of the opposite sign. Then the field-free solutions were exclusively used for the calculation of the absorption coefficients in this case, so that the whole difference

Table III.

Energy of the free electron	$\frac{df}{dE}$	α_V
0.04	1.96	$1.58 \cdot 10^{-17} \text{ cm}^2$
0.09	2.82	2.28
0.1225	2.86	2.31
0.16	2.72	2.19
0.25	2.20	1.78
0.36	1.66	1.34
0.49	1.22	0.98
0.64	0.88	0.71
0.81	0.65	0.52
1.00	0.47	0.38

found between them and MASSEY and BATES' is due to our use of another eigenfunction for the ion.

The total absorption coefficient—being the sum of the coefficients belonging to the different allowed transitions—was calculated for the limiting frequency for the absorption edges. This frequency is 1.052 Rydberg units as the binding energy of the electron in the ion is taken to be 0.052 according to HYLLEAAS. This frequency is somewhat larger than the limiting frequency for the Lyman continuum and corresponds to a wave length of 867 \AA . The energy of the free electron is then $\frac{1}{4}$ when the atom is in a 2-state and $\frac{1}{9}$ when it is in a 3-state. Further the transition probability was calculated for transitions to the 2-states only with energies of the free electron of $\frac{1}{2}$ and 1.

The eigenfunctions for the free electrons were calculated in higher approximations (2 or 3) in the two cases only, where the atom was in the 2*S*- and the 2*P*-state and the energy of the free electron was $\frac{1}{4}$. In the other cases the transition probability was so small, that a correction of it would have no interest. For the eigenfunctions of smaller energies (near the absorption limits) the exchange terms could not in a first approximation be regarded as small and the described method of successive approximations would be of little use.

Table IV.

Final state of the atom	$\left[\int_0^\infty P_A(s) \cdot P_B(s) \cdot dr \right]^2$	$\left[\int_0^\infty P_A(s) \cdot P_B(p) \cdot r dr \right]^2$		
		$\lambda = 867 \text{ \AA}$	$\lambda = 700 \text{ \AA}$	$\lambda = 506 \text{ \AA}$
1 S	0.865	0.781	—	—
2 S	0.137	3.846	0.266	0.060
3 S	0.002	0.153	—	—

Final state of the atom	$\left[\int_0^\infty P_A(s) \cdot P_B(p) \cdot r dr \right]^2$	$\left[\int_0^\infty P_A(s) \cdot P_B(s) \cdot dr \right]^2$		
		$\lambda = 867 \text{ \AA}$	$\lambda = 700 \text{ \AA}$	$\lambda = 506 \text{ \AA}$
2 P	9.175	0.264	0.023	0.004
3 P	0.066	0.038	—	—

The absorption coefficients for the transitions directly to the ground-state are given in Table III as a function of the energy of the free electron. The integrated f -sum is found to be 1.54 while for MASSEY and BATES' absorption coefficients it is found to be 1.50. The difference is unimportant, as the uncertainty due to the extrapolation of the $\frac{df}{dE}$ -curves for the integration is of the same order of magnitude. The good agreement between the two values seems to indicate, that the total absorption is better determined than the shape of the $\frac{df}{dE}$ -curve.

Table IV gives the squares of the product integrals

$$\int_0^\infty P_A(s) \cdot P_B(s) \cdot dr \quad \text{and} \quad \int_0^\infty r \cdot P_A(s) \cdot P_B(p) dr$$

and Table V the resulting absorption coefficients.

The integrated f -sum for the transitions to the 2-states is roughly about 0.6—0.7. The number of points on the curve is not quite sufficient to determine the run of the curve with certainty especially around its maximum near the Lyman limit. The probability for the transitions to the 3-states is seen to be

Table V.

Final state of the atom	$\lambda = 867 \text{ \AA}$	$\lambda = 700 \text{ \AA}$	$\lambda = 506 \text{ \AA}$
1 S	$0.38 \cdot 10^{-17} \text{ cm}^2$	—	—
2 S	0.30	0.02	0.01
3 S	0.00	—	—
2 P	1.37	0.15	0.04
3 P	0.00	—	—
Sum: 2.0 ₅			

The table gives z_ν as a function of wave length for different transitions.

negligible, so that we have probably found the whole of the theoretically expected absorption. The total f -sum is found to be 2.1—2.2. The agreement with the theoretical value of 2.0 has thus been considerably improved by taking the transitions covered by our calculations into account. The greatest part of the absorption due to transitions to the 2-states is seen to belong to the transition in which the final state of the atom is the 2 P -state.



DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 17

ON THE
PHOTO-DISINTEGRATION OF
THE DEUTERON

BY
A. PAIS



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

The photo-disintegration of the deuteron and the neutron-proton capture are discussed from the point of view of the meson theory of nuclear forces in the form proposed by MØLLER and ROSENFELD. The calculations include all first order relativistic contributions. The general expression for the photo-electric cross-section turns out to be identical in form with the corresponding quantity in the "old" theory (assuming a spherical potential well), while the photo-magnetic cross-section contains an extra term due to the charged meson fields. The theory accounts in a satisfactory way for the magnitude of the cross-sections. The discrepancy with regard to the angular distribution of the ejected particles which exists in the old theory is removed; the satisfactory agreement with experiment here found appears to be due mainly to the extra term mentioned above. For large energies, the cross-sections decrease more rapidly ($\propto \nu^{-7/2}$) than in the old theory, while there is a marked difference in angular distribution in this energy region as compared with BETHE'S "neutral" theory. The capture cross-section for thermal neutrons is proportional to ν^{-1} ; its value is in good agreement with experiment. In an Appendix, the reliability of some approximate expressions for the radial wave functions of the deuteron is discussed and the electric quadrupole transitions are given.

§ 1. Introduction.

The discovery, made by CHADWICK and GOLDHABER¹), that the deuteron can be disintegrated by γ -rays of sufficient energy, provides most valuable information about the interaction of electromagnetic radiation with nuclear systems. This effect is closely connected with the capture process of neutrons by protons which especially plays a prominent rôle in slow neutron experiments. In the earliest treatments of photo-disintegration²)³) as well as of proton-neutron capture, these effects were considered as photo-electric (*PE*) processes (interaction of the electric field of the incident wave with the nuclear system). The cross-sections thus obtained for the PE disintegration were in reasonable agreement with experiment, but there appeared to be a difference of several orders of magnitude between theoretical expectations and the measured values of the capture cross-section. This point was cleared up by the remark of FERMI⁴) that, besides the mentioned processes, one has also to take into account the photo-magnetic (*PM*) transitions due to the interaction of the magnetic field of the incident wave with the magnetic moments of the nuclear particles (cf. also BREIT and CONDON⁵); it was shown that the slow neutron capture is essentially of magnetic character and that the well-known $1/v$ law can be explained on this assumption. Thus, all experimental data known at the time could be accounted for by a theory based only on the assumption that the range of the nuclear forces is small compared to the wavelengths involved.

More recent experiments by VON HALBAN⁶⁾, however, seem to indicate a discrepancy with theoretical expectations on the angular distribution of the disintegration products. While for the *PM* effect (corresponding to a transition between the 3S -state and the 1S -state of the deuteron), this distribution is isotropic, the contribution of the *PE* effect (a $^3S \rightarrow ^3P$ transition) per unit solid angle is proportional to $\sin^2 \theta$, θ being the angle between the incident γ -ray and the ejected neutron (or proton). Therefore, from the expressions for the differential cross-section of both effects, which we shall denote by $d\Phi^{\text{el}}(\theta)$ and $d\Phi^{\text{magn}}$, we find for the ratio of the intensities at $\theta = 0$, (Φ_{\parallel}) and $\theta = \pi/2$, (Φ_{\perp}):

$$\frac{\Phi_{\parallel}}{\Phi_{\perp}} = \frac{d\Phi^{\text{magn}}}{d\Phi^{\text{magn}} + d\Phi^{\text{el}}(\pi/2)}. \quad (1)$$

For ThC'' γ -rays this ratio was calculated to be 0.29, assuming the virtual 1S -level of the deuteron to have an energy of about 10^5 eV; more detailed calculations of the *PE* cross-section which show that this cross-section had been underestimated^{3,5)}, lead to a value of 0.15. The same value has been obtained by RARITA and SCHWINGER⁷⁾ on the assumption of a spherical well potential combined with a directional coupling. On the other hand, v. HALBAN's measurements give a value of about 5%, in agreement with results obtained by CHADWICK, FEATHER and BRETSCHER⁸⁾ on the distribution of the photo-protons. In order to explain this apparent inconsistency, a better insight into the nature of nuclear forces may be deemed necessary and it is therefore of interest to discuss these problems in accordance with our present notions on the interaction between nucleons.

It is the aim of this paper to treat the photo-effect as well as proton-neutron capture from the point of view of the theory of MÖLLER and ROSENFELD⁹⁾, according to which nuclear interaction is established by a specific mixture of vector and pseudo-scalar meson fields, including charged and neutral mesons in a symmetrical way*. §§ 2—4 are devoted to the *PE* and *PM* effects, while the neutron capture is discussed in § 5.

* Recently, a discussion of the *PE* effect in the frame of the meson theory has been given by FRÖHLICH, HEITLER and KAHN¹⁰⁾, assuming the interaction to be described by a field of the vector type. However, their treatment is clearly inconsistent with the general electromagnetic properties of nuclear systems; their results will therefore not be considered here.

The admixture of a D -state with the 3S -state of the deuteron has practically no influence on the effects under consideration. The contribution of this D -state will therefore be neglected throughout and the ground state will thus be taken to be purely of the 3S -type. Moreover, it may be noted here that the inclusion of the electric quadrupole transitions will neither influence essentially the value of the total cross-section, nor the quantity defined in (1), the angular distribution to which it gives rise being proportional to $\sin^2 \theta \cos^2 \theta$ ($S \rightarrow D$ transition). This effect is treated in an Appendix.

§ 2. a) The wave equation of the deuteron.

We begin with a survey of the properties of the deuteron wave functions, representing in a slightly different form results obtained by KEMMER¹¹⁾ in a paper on the neutron-proton interaction.

The two nucleons constituting the deuteron and all quantities which refer to them are labelled with the upper indices 1 and 2, respectively; $\vec{x}^{(1)}$ and $\vec{x}^{(2)}$, for instance, represent the spatial coordinates of the first and second particles. The deuteron is described by a 16-component wave function $\psi_{E\mu}$ (μ stands for all those sets of values of the degeneracy parameters which belong to the same energy E). In the frame of reference in which the centre of gravity of the deuteron is at rest it satisfies the equation

$$\mathcal{H}_0 \psi_{E\mu}(\vec{x}) = \left[\frac{\hbar c}{i} \alpha \text{grad} + \beta Mc^2 + \mathcal{V}(r) \right] \psi_{E\mu}(\vec{x}) = E \psi_{E\mu}(\vec{x}) \quad (2)$$

with

$$\vec{x} = \vec{x}^{(1)} - \vec{x}^{(2)}, \quad r = |\vec{x}|, \quad \alpha = \vec{\alpha}^{(1)} - \vec{\alpha}^{(2)}, \quad \beta = \rho_3^{(1)} + \rho_3^{(2)}, \quad M \cong M_N \cong M_P,$$

$$\mathcal{V}(r) = (\mathbf{r}^{(1)} \mathbf{r}^{(2)}) \left[g_1^2 + g_2^2 (\vec{\sigma}^{(1)} \vec{\sigma}^{(2)}) \right] \frac{e^{-zr}}{4\pi r}, \quad z = \frac{M_m c}{\hbar},$$

M_m is the meson rest mass. The eigenvalue $+1$ (-1) of τ_3 denotes neutron (proton) states. According to KEMMER, the non-trivial proper solutions of (2) can be classified as follows:

Type Ia:		triplet state with $l = j \pm 1$,
Type Ib:	corresponds in non-relativistic	triplet state with $l = j$,
Type IIb:	approximation with	singlet state.

We now introduce the normalized spin wave functions

$$\begin{aligned}
 {}^3\chi_0 &= \frac{1}{\sqrt{2}} \left(\delta_{\sigma_3^{(1)}1} \delta_{\sigma_3^{(2)}-1} + \delta_{\sigma_3^{(1)}-1} \delta_{\sigma_3^{(2)}1} \right), \\
 {}^3\chi_1 &= \delta_{\sigma_3^{(1)}1} \delta_{\sigma_3^{(2)}1}, & {}^3\chi_{-1} &= \delta_{\sigma_3^{(1)}-1} \delta_{\sigma_3^{(2)}-1}, \\
 {}^1\chi_0 &= \frac{1}{\sqrt{2}} \left(\delta_{\sigma_3^{(1)}1} \delta_{\sigma_3^{(2)}-1} - \delta_{\sigma_3^{(1)}-1} \delta_{\sigma_3^{(2)}1} \right),
 \end{aligned}$$

and similarly the “isotopic spin wave functions” ζ and the “ ρ -wave functions” ξ :

$$\begin{aligned}
 {}^3\zeta_1 &= \delta_{t_3^{(1)},1} \delta_{t_3^{(2)},1}, \text{ etc.}; \\
 {}^3\xi_1 &= \delta_{\rho_3^{(1)}1} \delta_{\rho_3^{(2)}1}, \text{ etc.}
 \end{aligned}$$

Then, to the first order in the velocities and apart from isotopic spin dependence ($\psi^{(0)}$ is the velocity-independent part of ψ which will be called “large component”, $\psi^{(1)}$ which is of the order of v/c is the “small component”),

$$\psi = \psi^{(0)} + \psi^{(1)}$$

with

$$\begin{aligned}
 \text{Type Ia, Ib: } \psi^{(0)} &= {}^3\xi_1 \left[{}^3\chi_1 Z^1 + {}^3\chi_0 Z^0 + {}^3\chi_{-1} Z^{-1} \right], \\
 \text{Type IIb: } \psi^{(0)} &= {}^3\xi_1 {}^1\chi_0 Z^0,
 \end{aligned}
 \left. \vphantom{\begin{aligned} \text{Type Ia, Ib: } \psi^{(0)} \\ \text{Type IIb: } \psi^{(0)} \end{aligned}} \right\} (3a)$$

and

$$\begin{aligned}
 \text{Type Ia, Ib: } \psi^{(1)} &= {}^1\xi_0 \left[{}^3\chi_1 z^1 + {}^3\chi_0 z_s^0 + {}^3\chi_{-1} z^{-1} \right] + {}^3\xi_0 {}^1\chi_0 z_a^0, \\
 \text{Type IIb: } \psi^{(1)} &= {}^3\xi_0 \left[{}^3\chi_1 z^1 + {}^3\chi_0 z_s^0 + {}^3\chi_{-1} z^{-1} \right].
 \end{aligned}
 \left. \vphantom{\begin{aligned} \text{Type Ia, Ib: } \psi^{(1)} \\ \text{Type IIb: } \psi^{(1)} \end{aligned}} \right\} (3b)$$

Z and z only depend on the relative spatial coordinates; introducing polar variables ($x = r \sin \vartheta \cos \varphi$, $y = r \sin \vartheta \sin \varphi$, $z = r \cos \vartheta$), they may be expressed in the following way*

* The spherical harmonics are defined as in *loc. cit.*⁹⁾, equ. (115); we also use the same normalization prescriptions as stated there.

$$\begin{aligned}
& \text{Type I a} \quad \left\{ \begin{array}{l} Z^1 = \sqrt{(j+m-1)(j+m)} Y_{j-1}^{(m-1)} \\ Z^0 = -\sqrt{2(j+m)(j-m)} Y_{j-1}^{(m)} \\ Z^{-1} = \sqrt{(j-m-1)(j-m)} Y_{j-1}^{(m+1)} \end{array} \right\} \cdot \frac{1}{\sqrt{2j(2j-1)}} \cdot \frac{R_I(j-1)}{r} \\
& \text{Type I a} \quad \left\{ \begin{array}{l} Z^1 = \sqrt{(j-m+1)(j-m+2)} Y_{j+1}^{(m-1)} \\ Z^0 = \sqrt{2(j+m+1)(j-m+1)} Y_{j+1}^{(m)} \\ Z^{-1} = \sqrt{(j+m+1)(j+m+2)} Y_{j+1}^{(m+1)} \end{array} \right\} \cdot \frac{1}{\sqrt{2(j+1)(2j+3)}} \cdot \frac{R_I(j+1)}{r} \\
& \text{Type I b} \quad \left\{ \begin{array}{l} Z^1 = -\sqrt{(j+m)(j-m+1)} Y_j^{(m-1)} \\ Z^0 = -m\sqrt{2} Y_j^{(m)} \\ Z^{-1} = -\sqrt{(j+m+1)(j-m)} Y_j^{(m+1)} \end{array} \right\} \cdot \frac{1}{\sqrt{2j(j+1)}} \cdot \frac{R_I(j)}{r} \\
& \text{Type II b} \quad Z^0 = -Y_j^{(m)} \cdot \frac{R_{II}(j)}{r} \\
& \text{Type I a} \quad \left\{ \begin{array}{l} z^1 = \sqrt{2(j+m)(j-m+1)} Y_j^{(m-1)} \\ z_s^0 = 2m Y_j^{(m)} \\ z^{-1} = -\sqrt{2(j+m+1)(j-m)} Y_j^{(m+1)} \\ z_a^0 = 2j Y_j^{(m)} \end{array} \right\} \cdot \frac{1}{2\sqrt{2j}} \cdot \frac{A^1(j-1)}{r} \\
& \text{Type I a} \quad \left\{ \begin{array}{l} z^1 = \sqrt{2(j+m)(j-m+1)} Y_j^{(m-1)} \\ z_s^0 = 2m Y_j^{(m)} \\ z^{-1} = -\sqrt{2(j+m+1)(j-m)} Y_j^{(m+1)} \\ z_a^0 = -2(j+1) Y_j^{(m)} \end{array} \right\} \cdot \frac{1}{2\sqrt{2(j+1)}} \cdot \frac{A^2(j+1)}{r} \\
& \text{Type* I b} \quad \left\{ \begin{array}{l} z^1 = \sqrt{2(j+m-1)(j+m)} Y_{j-1}^{(m-1)} \\ z_s^0 = -2\sqrt{(j+m)(j-m)} Y_{j-1}^{(m)} \\ z^{-1} = \sqrt{2(j-m-1)(j-m)} Y_{j-1}^{(m+1)} \\ z_a^0 = 0 \end{array} \right\} \cdot \frac{-1}{2\sqrt{2j(j+1)(2j-1)}} \cdot \frac{C^1(j)}{r} \\
& \text{II b} \quad \left\{ \begin{array}{l} z^1 = \sqrt{2(j-m-1)(j-m)} Y_{j-1}^{(m-1)} \\ z_s^0 = 2\sqrt{(j+m+1)(j-m+1)} Y_{j-1}^{(m)} \\ z^{-1} = \sqrt{2(j-m+1)(j-m+2)} Y_{j-1}^{(m+1)} \\ z_a^0 = 0 \end{array} \right\} \cdot \frac{1}{2\sqrt{2(2j-1)}} \cdot \frac{H^1(j)}{r} \\
& \quad \quad \quad \left\{ \begin{array}{l} \sqrt{2(j-m+1)(j-m+2)} Y_{j+1}^{(m-1)} \\ 2\sqrt{(j+m+1)(j-m+1)} Y_{j+1}^{(m)} \\ \sqrt{2(j+m+1)(j+m+2)} Y_{j+1}^{(m+1)} \end{array} \right\} \cdot \frac{-1}{2\sqrt{2j(j+1)(2j+3)}} \cdot \frac{C^2(j)}{r} \\
& \quad \quad \quad \left\{ \begin{array}{l} \sqrt{2(j+m+1)(j+m+2)} Y_{j+1}^{(m-1)} \\ 2\sqrt{(j+m+1)(j-m+1)} Y_{j+1}^{(m)} \\ \sqrt{2(j+m+1)(j+m+2)} Y_{j+1}^{(m+1)} \end{array} \right\} \cdot \frac{1}{2\sqrt{2(2j+3)}} \cdot \frac{H^2(j)}{r}
\end{aligned}$$

* The expressions between the braces are the same for both types. The upper expressions behind the braces refer to I b-states, the lower ones to II b-states.

The large radial functions $R_s(j)$ satisfy

$$\left\{ \frac{\hbar^2}{M} \left(\frac{d^2}{dr^2} - \frac{j(j+1)}{r^2} \right) + \Gamma_{j,s} \frac{e^{-\alpha r}}{r} + E \right\} R_s(j) = 0, \quad (4)$$

where $s = \text{I or II}$ and

$$\left. \begin{aligned} \text{for type I a, I b: } \Gamma_{j,I} &= [1 - 2(-1)^{j+1}] \frac{g_1^2 + g_2^2}{4\pi}, \\ \text{for type II b: } \Gamma_{j,II} &= [1 - 2(-1)^j] \frac{g_1^2 - 3g_2^2}{4\pi}. \end{aligned} \right\} \quad (5)$$

Thus, for instance, $R_I(0)$ denotes the large radial wave function of the ground state. From (4) it follows that R_∞ , the asymptotic solution for a radial function R other than the ground state function, is given by

$$\left. \begin{aligned} R_\infty &= \lambda \sqrt{\frac{2}{\pi}} \cos(kr + \varepsilon_j); \\ \varepsilon_j &= -\frac{\pi}{2}(j+1) + \delta_j; \\ k &= \sqrt{ME}/\hbar; \end{aligned} \right\} \quad (6)$$

$$\lambda = \sqrt{\frac{M}{2\hbar^2 k}} = \sqrt{\frac{1}{\hbar v}}. \quad (7)$$

The factor $\lambda \sqrt{2/\pi}$ normalizes R_∞ in the energy scale¹²⁾; v is the nucleon velocity in the laboratory system. The phase constants δ_j are fixed by the exact solution of (4). With the exception of δ_0 , they are very small if $(\hbar^2/ME)^{1/2} \gg \alpha^{-1}$. On account of (6), the asymptotic expression for the complete large wave function may generally be written as

$$\psi_{E\mu}(r \rightarrow \infty) = B(E, \mu; \mathcal{J}, \varphi) \cdot \frac{1}{r} \cos(kr + \varepsilon_j), \quad (\psi \text{ large}). \quad (8a)$$

Furthermore, the small radial wave functions are related with the large functions by

$$\left. \begin{aligned}
 A^1(j-1) &= \left(\frac{d}{dr} - \frac{j}{r} \right) R_I(j-1) \\
 A^2(j+1) &= \left(\frac{d}{dr} + \frac{j+1}{r} \right) R_I(j+1) \\
 C^1(j) &= (j+1) \left(\frac{d}{dr} + \frac{j}{r} \right) R_I(j) \\
 C^2(j) &= j \left(\frac{d}{dr} - \frac{j+1}{r} \right) R_I(j) \\
 H^1(j) &= - \left(\frac{d}{dr} + \frac{j}{r} \right) R_{II}(j) \\
 H^2(j) &= \left(\frac{d}{dr} - \frac{j+1}{r} \right) R_{II}(j)
 \end{aligned} \right\} \frac{i\hbar}{Mc} \cdot \frac{1}{\sqrt{2j+1}} \quad (9)$$

and the asymptotic expression of the complete small wave function is

$$\psi_{E\mu}(r \rightarrow \infty) = B(E, \mu; \mathcal{J}, \varphi) \cdot \frac{i}{r} \sin(kr + \varepsilon_j), \quad (\psi \text{ small}). \quad (8b)$$

b) Interaction with electromagnetic radiation.

We now examine the effect of an irradiation of the deuteron with a monochromatic polarized γ -ray beam. In this case the deuteron wave equation is

$$i\hbar \frac{\partial \psi}{\partial t} = [\mathcal{H}_0 + (\Omega e^{-i\omega t} + \text{conj.})] \psi,$$

where¹³⁾

$$\Omega = -\{ \vec{P} \vec{\mathcal{E}} + \vec{M} \vec{\mathcal{H}} + (Q \text{ grad}) \vec{\mathcal{E}} \}. \quad (10)$$

$\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ are the electric and magnetic fields taken at the centre of gravity of the system. \vec{P} , \vec{M} and Q are the operators of electric dipole moment, magnetic dipole moment and electric quadrupole moment, respectively. For a general nuclear system, explicit expressions of these quantities in an arbitrary frame of reference have been given in a previous paper¹⁴⁾; for the deuteron, the indices (i) and (k) occurring there take the "values" (1) or (2). As in the present case \vec{P} , \vec{M} and Q of course refer to the system in which the centre of gravity is at rest, we have furthermore to replace $\vec{x}^{(1)}$ by $\vec{x}/2$ and $\vec{x}^{(2)}$ by $-\vec{x}/2$. Consequently,

$$\vec{P} = -\frac{e}{4} (\tau_{\mathbf{3}}^{(1)} - \tau_{\mathbf{3}}^{(2)}) \vec{x} - \frac{e}{8\pi\hbar c} \cdot \frac{g_1 g_2}{z} (\tau^{(1)} \mathbf{\Lambda} \tau^{(2)})_{\mathbf{3}} (\vec{\sigma}^{(1)} + \vec{\sigma}^{(2)}) \mathbf{\Lambda} \vec{x}_0 \cdot e^{-zr}, \quad (11)$$

$$\vec{M} = \left. \begin{aligned} & \frac{e}{2c} \sum_{i=1,2} \frac{1-\tau_{\mathbf{3}}^{(i)}}{2} \varrho_3^{(i)} \vec{x}^{(i)} \wedge \vec{v}^{(i)} + \mu_0 \sum_{i=1,2} \frac{1-\tau_{\mathbf{3}}^{(i)}}{2} \varrho_3^{(i)} \vec{\sigma}^{(i)} \\ & + \left(\frac{g_2}{z}\right)^2 \cdot \frac{1}{4\pi\hbar c} \cdot \frac{e}{4} \sum_{i,k=1,2} (\mathbf{r}^{(i)} \mathbf{\Lambda} \mathbf{r}^{(k)})_{\mathbf{3}} [(\vec{\sigma}^{(i)} \wedge \vec{\sigma}^{(k)}) (1-2zr^{(ik)}) \\ & + \{(\vec{\sigma}^{(i)} \wedge \vec{\sigma}^{(k)}) \vec{x}_0^{(ik)}\} \vec{x}_0^{(ik)} (1+zr^{(ik)})] \cdot \frac{e^{-zr^{(ik)}}}{r^{(ik)}} \end{aligned} \right\} (12)$$

$$Q^{mn} = \frac{e}{16} \{2 - (\tau_{\mathbf{3}}^{(1)} + \tau_{\mathbf{3}}^{(2)})\} x^m x^n, \quad (13)$$

with

$$\vec{x}_0 = \frac{\vec{x}}{r}, \quad \vec{x}_0^{(ik)} = \frac{\vec{x}^{(i)} - \vec{x}^{(k)}}{r^{(ik)}}, \quad \vec{v}^{(i)} = \frac{\hbar}{Mi} \text{grad}^{(i)}.$$

$\mu_0 = e\hbar/2Mc$ is the nuclear magneton. The first two terms of \vec{M} denote the "orbital" and spin magnetic moment of the nucleons, respectively*. It should be noted that the contribution of the (static) meson fields to Q^{mn} vanishes in the centre of gravity system. The expression (10) for the interaction operator is sufficiently accurate if the γ -ray wave-length is large compared to the "radius" of the deuteron, a condition which is well fulfilled for the whole energy region of interest.

The differential cross-section for photo-disintegration by γ -rays with a fixed direction of polarization is in a general way given by**

$$d\Phi = \pi^2 \bar{\varrho} v \sin \vartheta d\vartheta d\varphi, \quad (14)$$

where

$$\varrho = \left| \sum_{\mu} B(E, \mu; \vartheta, \varphi) \cdot (E, \mu | \Omega | o) \cdot e^{i\epsilon_j} \right|^2 \quad (15)$$

and where $\bar{\varrho}$ is meant to be the average of ϱ over the magnetic substates of the deuteron ground state (which is indicated by 0 in the above formula). The axis of the polar coordinate system is supposed to be taken in the direction of polarization of the photons.

* Details of the separation of the nucleon magnetic moment into these two parts are given in *loc. cit.*¹³; a term which is proportional to the electromagnetic field and one which may be written as a time derivative have been omitted, both being irrelevant for present purposes. It should be noted that the sign of the constant g_2 adopted here is different from that in *loc. cit.*¹³.

** Cf. *loc. cit.*¹⁵ pp. 59—61. (14) is derived in a way quite similar to the treatment of the PE effect of the hydrogen atom by BETHE¹⁶.

§ 3. Calculation of the cross-sections*.

In this paper, the photo-disintegration is treated up to the first order in the velocities, *i. e.* effects of the order v/c are taken into account. Therefore, those and only those matrix elements $(F|\Omega|0)$ will have to be considered which belong to one of the following three types

$$\int \psi_F^{(0)\dagger} \Omega \psi_0^{(0)}, \quad \int \psi_F^{(1)\dagger} \Omega \psi_0^{(0)}, \quad \int \psi_F^{(0)\dagger} \Omega \psi_0^{(1)},$$

the integral sign denoting integration over spatial coordinates and summation over spin coordinates as well.

a) The PE effect.

We shall now consider the transitions due to $\Omega_{\text{el}} = -\vec{\mathcal{E}} \vec{P}$. Taking the x -axis as the direction of propagation of the photon beam, and the z -axis as the direction of its electric vector,

$$\Omega_{\text{el}} = -EP_z.$$

The amplitude E of the fields of the wave is chosen such as to normalize the radiation to one polarized photon per sec. per cm^2 (Heaviside units are used throughout):

$$|E|^2 = \frac{\hbar \nu}{2c}. \quad (16)$$

Let us first consider

$$\Omega_{\text{el}}^{\text{nucl}} = -EP_z^{\text{nucl}} = \frac{eE}{4} (\tau_{\mathbf{3}}^{(1)} - \tau_{\mathbf{3}}^{(2)}) z.$$

As the ground state is antisymmetric in the isotopic spins and

$${}^3_{\xi_0} (\tau_{\mathbf{3}}^{(1)} - \tau_{\mathbf{3}}^{(2)}) {}^1_{\xi_0} = 2, \quad {}^1_{\xi_0} (\tau_{\mathbf{3}}^{(1)} - \tau_{\mathbf{3}}^{(2)}) {}^1_{\xi_0} = 0,$$

the final state must be symmetric with respect to $\tau_{\mathbf{3}}^{(i)}$. Taking further into account the behaviour with regard to rotations and spatial reflections and the fact that the ground state is of the

* I should like to thank Prof. C. MÖLLER for the communication of preliminary calculations on the photo-effect which have provided a valuable check of the calculations given here.

type Ia with $l = 0, j = 1$, the following states are found to combine with the ground state (behind each state the spectroscopic symbol corresponding to the non-relativistic approximation is indicated):

$$\text{I a, } l = 1, j = 2 \quad ({}^3P_2^-),$$

$$\text{I a, } l = 1, j = 0 \quad ({}^3P_0^-),$$

$$\text{I b, } l = j = 1 \quad ({}^3P_1^-),$$

while the familiar selection rule $\Delta m = 0$ holds. From (3) it is easily seen that, to the first order in the velocities, we have for all these transitions

$$(F | \Omega_{\text{el}}^{\text{nucl}} | 0) = \int \psi_F^{\dagger(0)} \Omega_{\text{el}}^{\text{nucl}} \psi_0^{(0)}. \quad (17)$$

The matrix elements are readily found to be

$$(\text{I a, } l = 1, j = 2 | \Omega_{\text{el}}^{\text{nucl}} | 0) = \frac{eE}{2} J \cdot \begin{cases} \sqrt{6}/6 & 1 \rightarrow 1 \\ \sqrt{2}/3 & 0 \rightarrow 0 \\ \sqrt{6}/6 & -1 \rightarrow -1 \end{cases}$$

$$(\text{I a, } l = 1, j = 0 | \Omega_{\text{el}}^{\text{nucl}} | 0) = -\frac{eE}{6} J \quad 0 \rightarrow 0$$

$$(\text{I b, } j = 1 | \Omega_{\text{el}}^{\text{nucl}} | 0) = -\frac{eE}{2} J \cdot \begin{cases} \sqrt{6}/6 & 1 \rightarrow 1 \\ -\sqrt{6}/6 & -1 \rightarrow -1 \end{cases}$$

Behind each expression the corresponding magnetic transition has been indicated, while

$$J = \int_0^\infty dr R_l(0) R_l(1) r.$$

$\Omega_{\text{el}}^{\text{exch}}$, the second part of Ω_{el} which, according to (11), is given by

$$\Omega_{\text{el}}^{\text{exch}} = -E P_z^{\text{exch}} = \frac{eE}{8\pi\hbar c} \cdot \frac{g_1 g_2}{z} \cdot (\mathbf{r}^{(1)} \mathbf{\Lambda} \mathbf{r}^{(2)})_{\mathbf{z}} \{ (\vec{\sigma}^{(1)} + \vec{\sigma}^{(2)}) \wedge \vec{x}_0 \}_z e^{-zr}$$

does, in our approximation, not contribute to the PE effect. This will be shown in the Appendix.

With the help of (14) and (15), the differential cross-section now can directly be obtained. As the final states are all P -states having the same radial wave function, the factor $\exp(i\varepsilon_j)$ in (15)

may be omitted. Using (7) and (2) and neglecting those terms in B that are proportional to $(v/c)^2$ we then get expressions which prove to be the same for the three possible magnetic states, so that they directly give the average value \bar{q} . Denoting by θ, ψ the direction of the ejected neutron with respect to the direction of propagation of the incoming γ -ray beam, we obtain, after averaging over all directions of polarization (which gives a factor $1/2$),

$$d\Phi^{\text{el}}(\theta) = \frac{e^2 \nu}{32c} |J|^2 \cdot \sin^2 \theta \cdot \sin \theta d\theta d\psi. \quad (18a)$$

The total cross-section is

$$\Phi^{\text{el}} = \frac{\pi e^2 \nu}{12c} |J|^2. \quad (18b)$$

This result is identical in form with that obtained in the BETHE-PEIERLS theory for a spherical well potential. Deviations from this simple formula are at most to be expected in the second order with respect to the velocities.

b) The PM effect.

Again beginning with the nucleon terms we have, noting that the first term of (12) gives no contribution,

$$\Omega_{\text{magn}}^{\text{nucl}} = E \mu_0 \sum_i \frac{1 - \tau_{\mathbf{z}}^{(i)}}{2} \varrho_3^{(i)} \sigma_y^{(i)},$$

as the magnetic vector stands in the $-y$ -direction. For transitions to states antisymmetric in the isotopic spins, this operator becomes $E \mu_0 (\varrho_3^{(1)} \sigma_y^{(1)} + \varrho_3^{(2)} \sigma_y^{(2)})/2$, while

$$\Omega_{\text{magn}}^{\text{nucl}} = -\frac{E \mu_0}{2} (\varrho_3^{(1)} \sigma_y^{(1)} - \varrho_3^{(2)} \sigma_y^{(3)})$$

if the final state is symmetric in these coordinates. It is easily seen that the matrix elements vanish in the former case; as to the latter, the only combining state is II b, $j = 0$ (1S) with $\mathcal{M} = \pm 1$. For $1 \rightarrow 0$ as well as for $-1 \rightarrow 0$, one finds

$$(\text{II b}, j = 0 | \Omega_{\text{magn}}^{\text{nucl}} | 0) = -\frac{i \mu_0 E \sqrt{2}}{2} \int_0^\infty dr R_I(0) R_{II}(0).$$

The extra magnetic moments of proton and neutron due to their proper (static) meson fields form part of the rest of (12), *viz.* those terms for which $i = k^*$. They give of course an infinite contribution and can only be managed by using a cut-off prescription. The corresponding interaction is given by

$$\Omega_{\text{magn}}^{\text{extra}} = - \left(\frac{g_2}{z} \right)^2 \frac{1}{4 \pi \hbar c} \cdot e E \lim_{\varrho=0} \sum_i \tau_{\mathbf{3}}^{(i)} \left\{ \overset{\rightarrow}{\sigma}^{(i)} (1 - 2 z \varrho) + \left(\overset{\rightarrow}{\sigma}^{(i)} \overset{\rightarrow}{\varrho}_0 \right) \overset{\rightarrow}{\varrho}_0 (1 + z \varrho) \right\}_y \cdot \frac{e^{-z\varrho}}{\varrho} \quad (19)$$

with $\varrho = |\overset{\rightarrow}{\varrho}|$ and $\overset{\rightarrow}{\varrho}_0 = \overset{\rightarrow}{\varrho}/\varrho$. Replacing $\left(\overset{\rightarrow}{\sigma}^{(i)} \overset{\rightarrow}{\varrho}_0 \right) \varrho_{0y}$ by its directional average $\sigma_y^{(i)}/3$, as would seem appropriate if the nuclear point sources are considered as the limiting case of a spherically symmetrical distribution, and cutting off by putting

$$\frac{2}{3} \frac{g_2^2}{4 \pi \hbar c} \frac{M}{M_m} \lim_{\varrho=0} \left(\frac{4}{z\varrho} - 5 \right) e^{-z\varrho} \quad (20 a)$$

equal to a finite quantity μ , gives

$$\Omega_{\text{magn}}^{\text{extra}} = - \mu_0 \mu E \sum_i \tau_{\mathbf{3}}^{(i)} \sigma_y^{(i)}.$$

μ is the "extra magnetic moment" in units μ_0 . Experimentally, a slight dissymmetry between extra proton and neutron moment is found, the discussion of which, however, falls outside the scope of the present considerations¹³). Formally, we may account for it by writing instead of the last expression

$$\Omega_{\text{magn}}^{\text{extra}} = \mu_0 E \sum_i \left\{ (\mu_P - 1) \frac{1 - \tau_{\mathbf{3}}^{(i)}}{2} + \mu_N \frac{1 + \tau_{\mathbf{3}}^{(i)}}{2} \right\} \sigma_y^{(i)},$$

thus using the empirical value μ_P (μ_N) of the magnetic proton (neutron) moment instead of μ . The introduction of a term proportional to $(\sigma_y^{(1)} + \sigma_y^{(2)})$ which is involved in this change on account of $\mu_P + \mu_N - 1 \neq 0$ is irrelevant, since it does not give rise to any transitions. The only allowed state is again the ¹S-state in which case the operator becomes

* Other terms of higher order which also contribute to these extra magnetic moments have to be discarded according to the prescription given in *loc. cit.*¹³).

$$-\frac{1}{2}\mu_0(\mu_P - \mu_N - 1)E(\sigma_y^{(1)} - \sigma_y^{(2)}).$$

Consequently,

$$(\text{II b, } j = 0 | \Omega_{\text{magn}}^{\text{extra}} | 0) = -\frac{i\sqrt{2}}{2}\mu_0(\mu_P - \mu_N - 1)E \int_0^\infty dr R_I(0) R_{II}(0), \quad (20\text{b})$$

and the total contribution of the nucleons is

$$(\text{II b, } j = 0 | \Omega_{\text{magn}}^{\text{nucl}} + \Omega_{\text{magn}}^{\text{extra}} | 0) = -\frac{i\sqrt{2}}{2}\mu_0(\mu_P - \mu_N)E \int_0^\infty dr R_I(0) R_{II}(0). \quad (21)$$

Finally, the terms from the second part of (12) with $i \neq k$ must be considered. The corresponding interaction, $\Omega_{\text{magn}}^{\text{exch}}$, is

$$\begin{aligned} \Omega_{\text{magn}}^{\text{exch}} = & \frac{eE}{2} \frac{g_2^2}{4\pi\hbar c} (\mathbf{T}^{(1)} \mathbf{A} \mathbf{T}^{(2)})_{\mathbf{3}} \left[(\vec{\sigma}^{(1)} \wedge \vec{\sigma}^{(2)})_y \left(\frac{1}{z^2 r} - \frac{2}{z} \right) \right. \\ & \left. + \left\{ (\vec{\sigma}^{(1)} \wedge \vec{\sigma}^{(2)}) \vec{x}_0 \right\} x_{0y} \left(\frac{1}{z^2 r} + \frac{1}{z} \right) \right] e^{-zr}. \end{aligned}$$

As ${}^3\zeta_0(\mathbf{T}^{(1)} \mathbf{A} \mathbf{T}^{(2)})_{\mathbf{3}} {}^1\zeta_0 = -2i$, ${}^1\zeta_0(\mathbf{T}^{(1)} \mathbf{A} \mathbf{T}^{(2)})_{\mathbf{3}} {}^1\zeta_0 = 0$, the final states must be symmetric in the isotopic spins. The only states combining with the ground state turn out to be

Type II b, $j = 0$, (1S),

Type II b, $j = 2$, (1D).

The matrix elements are found to be again of the type (17). They are

$$\begin{aligned} & \left. \begin{array}{l} 1 \rightarrow 0 \\ -1 \rightarrow 0 \end{array} \right\} (\text{II b, } j = 0 | \Omega_{\text{magn}}^{\text{exch}} | 0) \\ & = \frac{2i\sqrt{2}}{3} \frac{g_2^2}{4\pi\hbar c} \frac{M}{M_m} \mu_0 E \int_0^\infty dr R_I(0) R_{II}(0) \left(\frac{4}{zr} - 5 \right) e^{-zr}, \end{aligned} \quad (22)$$

$$\begin{aligned} & \left. \begin{array}{l} 1 \rightarrow 0 \\ -1 \rightarrow 0 \end{array} \right\} (\text{II b, } j = 2 | \Omega_{\text{magn}}^{\text{exch}} | 0) \\ & = -\frac{2i\sqrt{10}}{15} \frac{g_2^2}{4\pi\hbar c} \frac{M}{M_m} \mu_0 E \int_0^\infty dr R_I(0) R_{II}(2) \left(\frac{1}{zr} + 1 \right) e^{-zr}. \end{aligned} \quad (23)$$

The obvious similarity in form of (20 a, b) and (22) becomes clear if one remembers that, for an $S \rightarrow S$ transition, $\left\{ (\vec{\sigma}^{(1)} \wedge \vec{\sigma}^{(2)}) \vec{x}_0 \right\} x_{0y}$ should be replaced by its directional average which is $(\vec{\sigma}^{(1)} \wedge \vec{\sigma}^{(2)})_y / 3$.

Having verified that the contribution of (23) is very small, we have ignored this transition.

The differential cross-section is computed from (21) and (22) in the same way as (18 a) was found; averaging over the initial magnetic substates gives a factor 2/3. The result is

$$d\Phi^{\text{magn}} = \frac{1}{12} \mu_0^2 \frac{\nu}{c} |K|^2 \cdot \sin \theta d\theta d\psi, \quad (24 a)$$

with

$$K = (\mu_P - \mu_N) \int_0^\infty dr R_I(0) R_{II}(0) - \frac{4}{3} \frac{g_2^2}{4\pi \hbar c} \frac{M}{M_m} \int_0^\infty dr R_I(0) R_{II}(0) \left(\frac{4}{zr} - 5 \right) e^{-zr}.$$

Therefore,

$$\Phi^{\text{magn}} = \frac{\pi}{3} \mu_0^2 \frac{\nu}{c} |K|^2. \quad (24 b)$$

§ 4. Numerical evaluations.

Approximate expressions for the radial wave function of the ground state and the ¹S-state have been obtained by HULTHÉN^{17, 18}), assuming the nuclear potential to be of the form as determined by (4) and (5). They are:

$$R_I(0) = \sqrt{\varrho z} e^{-\beta z r} \{ (1 - e^{-zr}) - c_1 (1 - e^{-zr})^2 \}, \quad \beta = \frac{M}{M_m} \left(\frac{E_0}{Mc^2} \right)^{1/2}, \quad (25)$$

$$R_{II}(0) = \lambda \sqrt{\frac{2}{\pi}} \cdot \frac{1}{1 - c_2} \{ (1 - e^{-zr}) - c_2 (1 - e^{-zr})^2 \} \sin(kr + \delta). \quad (26)$$

ϱ is a normalizing factor and is given by loc. cit.¹⁷) equ. (35). The constants c_1 and c_2 are determined by variational methods. For nucleon energies not much larger than zero, $c_2 = 0,349$ while, taking $E_0 = 2,16 \text{ MeV}$,

α	z	β	ϱ	c_1	$\frac{g_2^2}{4\pi \hbar c}$
200	$0,52 \cdot 10^{13} \text{ cm.}$	0,442	4,66	0,370	0,065
300	$0,78 \cdot 10^{13} \text{ cm.}$	0,294	2,48	0,365	0,095

α denoting the ratio between meson and electron mass. g_2 and the phase δ in (26) depend upon the energy E'_0 of the virtual ¹S-level, for which we have taken $E'_0 = 5 \cdot 10^4 \text{ eV}$. With $\gamma = (ME'_0)^{1/2}/\hbar = 3,48 \cdot 10^{11} \text{ cm}^{-1}$,

$$\cos \delta = \frac{\gamma}{\sqrt{\gamma^2 + k^2}}, \quad \sin \delta = \frac{k}{\sqrt{\gamma^2 + k^2}}.$$

The P -function is, assuming as usual the interaction between proton and neutron to be negligible in this state,

$$R_I(1) = \lambda \sqrt{\frac{2}{\pi}} \left(-\cos kr + \frac{\sin kr}{kr} \right).$$

Inserting all this in (18 b) and (24 b) we get (it should be borne in mind that e is expressed in Heaviside units)

$$\Phi^{\text{el}} = \frac{\rho}{3} \cdot \frac{e^2}{\hbar c} \cdot \left(\frac{Mc}{\hbar} \right)^2 \cdot \frac{\hbar \nu}{Mc^2} \cdot \frac{z}{k} \cdot \frac{1}{k^4} \left| (1-c_1) P_0^2 + (2c_1-1) P_1^2 - c_1 P_2^2 \right|^2 \quad (27)$$

$$\Phi^{\text{magn}} = \left. \begin{aligned} & \frac{\rho}{12} \cdot \frac{1}{(1-c_2)^2} \cdot \frac{e^2}{\hbar c} \cdot \frac{\hbar \nu}{Mc^2} \cdot \frac{z}{k} \cdot \frac{1}{k^2} \cdot \frac{\gamma^2}{\gamma^2 + k^2} \left| (\mu_P - \mu_N) F(X) \right. \\ & \left. + \frac{16}{3} \cdot \frac{g_2^2}{4\pi \hbar c} \cdot \frac{M}{M_m} \cdot F(Y) \right|^2 \end{aligned} \right\} \quad (28)$$

where the following abbreviations have been used:

$$P_n = \left\{ 1 + (\beta + n)^2 \left(\frac{z}{k} \right)^2 \right\}^{-1},$$

$$F(X) = (1-c_1)(1-c_2)X_0 - (4c_1c_2 - 3c_1 - 3c_2 + 2)X_1 \\ + (6c_1c_2 - 3c_1 - 3c_2 + 1)X_2 - (4c_1c_2 - c_1 - c_2)X_3 + c_1c_2X_4,$$

with

$$X_n = \left\{ 1 + (\beta + n) \frac{z}{\gamma} \right\} P_n.$$

$F(Y)$ is understood to be obtained from $F(X)$ by replacing X_n by Y_n , where

$$Y_n = \frac{k}{z} \left[\text{arc tg} \left\{ (\beta + n) \frac{z}{k} \right\} + \frac{k}{2\gamma} \ln \left\{ (\beta + n)^2 + \frac{k^2}{z^2} \right\} \right] + \frac{5}{4} X_{n+1}.$$

The PE cross-section decreases with increasing z , as is shown by the following table:

$\hbar \nu$ in MeV	Φ^{el} in 10^{-28} cm 2	
	$\alpha = 200$	$\alpha = 300$
2,64	15	12
4	34	23
6,2	24	18
10	16	11
17	9	5,5

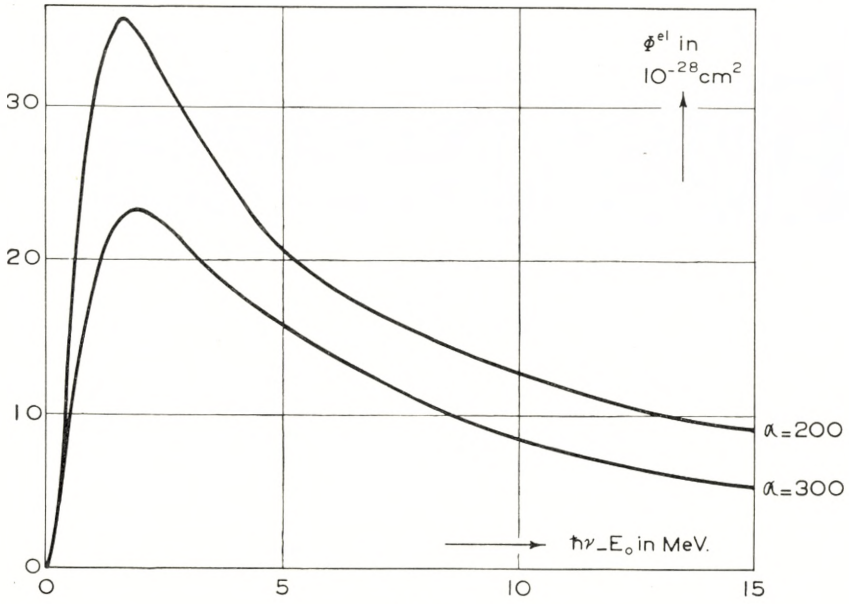


Fig. 1.

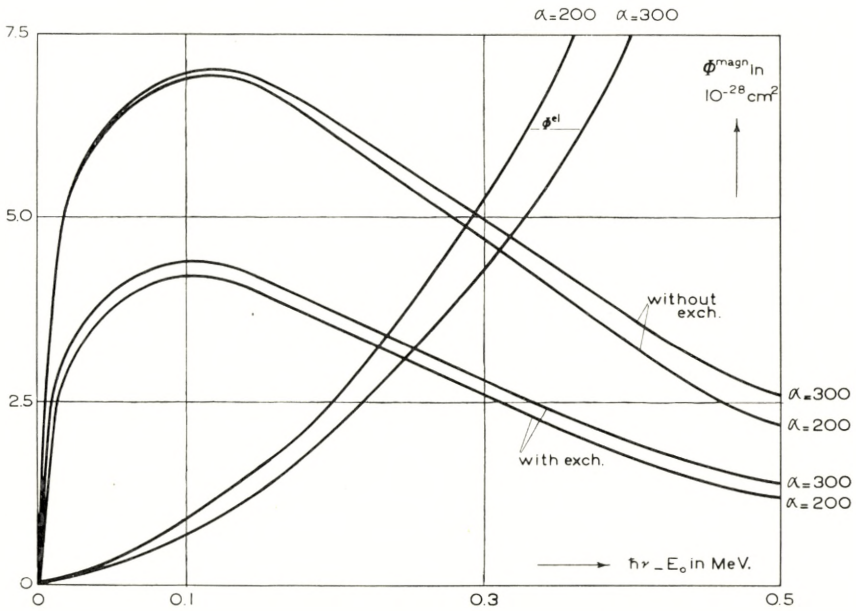


Fig. 2.

In Fig. 1, Φ^{el} has been plotted against $\hbar\nu - E_0 = 2E$. From the considerations in the Appendix, Note 1, it will become clear that one should be careful in drawing quantitative conclusions from this graph.

The magnetic cross-section is given in Fig. 2 (curves marked "with exch."); we have taken $\mu_P = 2,78$, $\mu_N = -1,93$. In particular, one has for the ThC'' γ -rays, $\hbar\nu = 2,64 \text{ MeV}$:

$$\Phi^{\text{magn}} = \begin{cases} 1,3 \cdot 10^{-28} & \text{for } \alpha = 200, \\ 1,5 \cdot 10^{-28} & \text{for } \alpha = 300. \end{cases}$$

It follows that for this energy

$$\frac{\Phi_{\parallel}}{\Phi_{\perp}} = \begin{cases} 0,055, & \alpha = 200, \\ 0,075, & \alpha = 300 \end{cases}$$

which shows that the present theory gives a good account of the angular distribution. In order to understand better the origin of the difference between this and the "old" theory, we have also computed the value of $\Phi_{\parallel}/\Phi_{\perp}$ which would be obtained by omitting the term due to the meson fields in (28); for doing so, Φ^{magn} , too, would have the same general form as in the old theory (though it should be remembered that other expressions for the radial wave-functions are used). The curves of Fig. 2 marked "without exch." refer to Φ^{magn} as calculated by omitting $F(Y)$. We then find, for the ThC'' γ -rays, $\Phi^{\text{magn}} = 2,4 \cdot 10^{-28}$ for $\alpha = 200$, and $2,8 \cdot 10^{-28}$ for $\alpha = 300$, which would lead to 10 % ($\alpha = 200$) and 13 % ($\alpha = 300$) for $\Phi_{\parallel}/\Phi_{\perp}$. Thus, the present result is essentially due to the existence of meson exchange currents*.

One can, therefore, not expect to get satisfactory values for (1) by means of a "neutral" theory, as this does not exhibit exchange phenomena, notwithstanding the influence on the angular distribution of a strong directional coupling which may be inherent in such a theory, as is *e. g.* the case in the theory put forward by BETHE²¹). As a matter of fact, it can be seen from the form of the radial wave-functions of the states involved

* A discussion of the obtained values in connection with the magnitude of Φ^{el} is given in the Appendix, Note 1.

in the transitions due to this coupling that the angular distribution will not change appreciably for energies not much larger than E_0 .

From Fig. 2 it appears that with increasing energy Φ^{magn} initially increases much more rapidly than Φ^{el} . Indeed, it can be seen from (27) and (28) that for $k \ll \kappa$, $\Phi^{\text{el}} \propto k^3$ and thus increases like $E^{3/2}$, while $\Phi^{\text{magn}} \propto k(\gamma^2 + k^2)^{-1}$; from the latter result one infers, moreover, that Φ^{magn} attains its maximum at the "resonance value" $k \cong \gamma$, corresponding to $\hbar\nu - E_0 = 0,1 \text{ MeV}$, *i. e.* twice the value of the energy of the ^1S -level. The angular distribution just above the threshold should apparently be nearly isotropic.

In a recent paper*, MYERS and VAN ATTA²⁷⁾ report the results of photo-disintegration experiments in which X-rays are used with energies ranging from 0—0,25 MeV above the threshold; the major part of the intensity lies within 0,1 MeV of this limiting energy (for which they find $2,183 \pm 0,012 \text{ MeV}$). The ratio of the intensities at 90° and 0° appears to be $1,15 \pm 0,10$. We should like to point out that no comparison with theoretical results is possible without a detailed knowledge of the X-ray spectrum employed: writing the intensity ratio under consideration as $1 + x$, it is seen that, in the energy region $\hbar\nu - E_0 = 0 - 0,1 \text{ MeV}$, x increases proportionally to E and, thus, varies rapidly for the energies concerned. For reference, we give the values of x at 0,1 MeV: 0,31 (0,19) for $\alpha = 200$ and 0,23 (0,15) for $\alpha = 300$ (the values in brackets are obtained if the contributions of the meson exchange currents are omitted).

Calculations on the magnetic effect have mostly been performed for vanishing range of the nuclear forces^{5,7)}. For comparison, we shall give the numerical result in this limiting case on the present theory. It is easily seen that $F(Y) = 0$ for $\kappa^{-1} = 0$. As $g_2^2 \kappa^{-1}$ practically does not depend on κ^9 , the exchange effect therefore vanishes for zero range. Furthermore, only X_0 is now different from zero, this quantity being independent of κ , while $q\kappa$ in the limit tends to a finite value. Φ^{magn} then turns out to be $3,2 \cdot 10^{-28}$, as compared with $3,3 \cdot 10^{-28}$ found by RARITA and SCHWINGER⁷⁾. It is to be noted that the dependence

* I have not been able to see this article myself; an abstract of its contents has kindly been communicated to me by dr. L. HULTHÉN.

of Φ^{magn} on range is different from that of Φ^{el} . To this point we shall come back in § 5.

As to the absolute value of the photo-disintegration cross-section, there is reasonable agreement with the measured values, *viz.* $5 \cdot 10^{-28}$ (CHADWICK and GOLDHABER¹)) and $9 \cdot 10^{-28} \text{ cm}^2$ (v. HALBAN⁶)) for $\hbar\nu = 2,64$, and $11,6 \cdot 10^{-28} \text{ cm}^2$ for $6,2 \text{ MeV}$ (ALLEN and SMITH¹⁹); cf. especially the Appendix, Note 1. The cross-section reaches a maximum at about 4 MeV and then decreases rapidly. In fact, it is easily seen from (27) and (28) that for very large energies both Φ^{el} and Φ^{magn} decrease as $\nu^{-3/2}$, *i. e.* more rapidly than in the old theory ($\sim \nu^{-3/2}$).

Theories which, in contrast to the mixed theory, involve a strong directional coupling of the dipole interaction type give rise to an angular distribution of a quite different kind. This has been calculated by RARITA and SCHWINGER²⁰) for the $Li + H$ γ -ray energy ($17,5 \text{ MeV}$) for which in all theories the magnetic effects are negligible. It is seen that the "symmetrical" theory gives a total cross-section of $3,8 \cdot 10^{-28}$ and an angular distribution such that $\Phi_{\parallel}/\Phi_{\perp} = 0,01$ while, in the "neutral" theory²¹) these quantities are $7,7 \cdot 10^{-28}$ and $0,27$, respectively. Here, the emission in the forward direction is due to electric transitions induced by the non-central forces which lead to a 3D contribution to the ground state. While the total cross-sections are seen not to differ much from the value given here on mixed theory, the angular distributions cannot be directly compared with (18a), as in our approximation the non-central forces do not come into consideration. Indeed, it is an essential feature of the mixed theory, distinguishing it from all other current meson theories, that the tensor interaction responsible for the S - D coupling is of non-static nature. Thus, the matrix elements of the corresponding transitions are of higher order in v/c (which for $\hbar\nu = 17,5 \text{ MeV}$ is $\sim 0,1$) compared with those given here, so that their contribution, even for this energy, will be relatively small; therefore $\Phi_{\parallel}/\Phi_{\perp}$ would, according to the mixed theory, seem to be of the same order as in the symmetrical theory with directional coupling and would at any rate be much smaller than in the neutral theory.

In the foregoing, it has been tacitly assumed that the centre of gravity system of the deuteron may be identified with the

actual system of measurement. In the high energy region a correction is necessary, however, as here the photon momentum may not be neglected.

§ 5. Capture of neutrons by protons.

Although the old theory could not account for the *PM* disintegration, the agreement of the theoretically found capture cross-section for thermal neutrons with experiment was satisfactory. As the latter process is entirely of magnetic origin in the energy region concerned and its probability is intimately connected with Φ^{magn} , it might be feared that the change of magnitude of Φ^{magn} as compared with the old theory would affect the capture cross-section in an unfavourable way. However, this is not the case, due to the circumstance that, for thermal energies, the influence of the exchange terms is considerably less than for the energies of interest in the discussion of the photo-effect.

The cross-section for this process can immediately be inferred from (27) and (28). We have in fact, calling the cross-sections for "electric" and "magnetic" capture Φ_c^{el} and Φ_c^{magn} ,

$$\Phi_c^{\text{el}} = \frac{9}{2} \left(\frac{\nu}{kc} \right)^2 \Phi^{\text{el}}, \quad \Phi_c^{\text{magn}} = \frac{3}{2} \left(\frac{\nu}{kc} \right)^2 \Phi^{\text{magn}}.$$

We are especially interested in the behaviour of these expressions in the region of thermal neutron energies; in this case, $\gamma^2 + k^2$ may be replaced by γ^2 . It is seen that for these energies $\Phi_c^{\text{el}} \propto k$ and, thus, may be neglected compared with Φ_c^{magn} which is $\propto k^{-1}$. For very small k

$$\Phi_c^{\text{magn}} = \frac{g}{8} \cdot \frac{1}{(1-c_2)^2} \cdot \frac{e^2}{\hbar c} \cdot \left(\frac{\hbar \nu}{Mc^2} \right)^3 \cdot \left(\frac{Mc}{\hbar} \right)^2 \cdot \left(\frac{k}{z} \right)^3 \cdot \frac{1}{k^4} \cdot \left| \left(\mu_P - \mu_N \right) F(\bar{X}) + \frac{16}{3} \cdot \frac{g_2^2}{4\pi \hbar c} \cdot \frac{M}{M_m} F(\bar{Y}) \right|^2$$

with

$$\bar{X}_n = \frac{1}{(\beta+n)^2} \left\{ 1 + (\beta+n) \frac{z}{\gamma} \right\}$$

$$\bar{Y}_n = \frac{z}{\gamma} \ln(\beta+n) + \frac{5}{4} \bar{X}_{n+1}.$$

(The term of \bar{Y}_n containing an *arc tg* need not be written down, as these various terms cancel each other). For $\hbar\nu$ one may take E_0 . The following numerical results have been obtained:

v in cm/sec.	$\mathcal{D}_c^{\text{magn}}$ in 10^{-24} cm ²		"Old" th.	Exp.
	$\alpha = 200$	$\alpha = 300$		
$2,2 \cdot 10^5$	0,23 (0,39)	0,26 (0,39)	0,35	0,27 ²²⁾
$2,5 \cdot 10^5$	0,20 (0,34)	0,23 (0,34)	0,31	0,31 ²³⁾ .

The values in brackets are obtained by omitting the exchange term. In the last column but one, the values according to the old theory are given; cf. *loc. cit.*¹²⁾, equ. (95). The agreement with experiment is satisfactory.

We have considered in some detail the dependence of this effect on κ . Just as for *PM* disintegration, it appears that for $\kappa^{-1} = 0$ the contribution due to the exchange currents vanishes and that only \bar{X}_0 differs from zero. For $v = 2,2 \cdot 10^5$ cm/sec., $\mathcal{D}_c^{\text{magn}} = 0,39 \cdot 10^{-24}$ in this case. Further, by disregarding the exchange effect, one obtains the range dependence due to the form of the radial wave-functions and, thus, to the Yukawa potential employed in the present calculations. It then appears that, for small values of κ^{-1} , $\mathcal{D}_c^{\text{magn}}$ is practically constant* and then decreases very slowly. As, for very small κ , $q \propto \kappa^{-3}$, it can be seen from the analytical expression of the quantity considered that for $\kappa \rightarrow 0$ it tends to a finite value differing from zero.

* This has also been found for the case of a MORSE potential²⁵⁾. However, in the present case, calculations up to $20 \cdot 10^{-13}$ cm. show a steady decrease of the capture cross-section, whereas in *loc. cit.* a sharp increase is found at $6 \cdot 10^{-13}$ cm.

Appendix.

Note 1. On the *PE* effect. This has been calculated using the operator $-\vec{\mathcal{E}}\vec{P}$. But as

$$\vec{\mathcal{A}}\dot{\vec{P}} = \vec{\mathcal{E}}\vec{P} + \frac{d}{cdt} (\vec{\mathcal{A}}\vec{P})$$

and as the second term on the right has vanishing matrix elements for the transitions concerned, because of energy conservation, $-\vec{\mathcal{A}}\dot{\vec{P}} = \vec{\Omega}_{el}$ may be taken just as well. This will be done here; in the centre of gravity system¹³⁾

$$\dot{\vec{P}} = \frac{e}{2} \sum_i (1 - \tau_{\mathbf{3}}^{(i)}) \vec{\alpha}^{(i)} + \frac{e}{4\pi\hbar c} (\mathbf{r}^{(1)} \mathbf{A} \mathbf{r}^{(2)})_{\mathbf{3}} \{g_1^2 + g_2^2 (\vec{\sigma}^{(1)} \vec{\sigma}^{(2)})\} \vec{x}_0 \cdot e^{-zr}. \quad (29)$$

The first term gives

$$-\frac{Ae}{2} [(1 - \tau_{\mathbf{3}}^{(1)}) \alpha_z^{(1)} + (1 - \tau_{\mathbf{3}}^{(2)}) \alpha_z^{(2)}],$$

where $A = -icE\nu^{-1}$ is the amplitude of the vector potential. With the help of (9), the matrix elements are found to be

$$\left. \begin{aligned} (\text{I a, } l = 1, j = 2 | \vec{\Omega}_{el}^1 | 0) &= -\frac{ieA\hbar}{Mc} J_1 \cdot \begin{cases} \sqrt{6}/6 & 1 \rightarrow 1 \\ \sqrt{2}/3 & 0 \rightarrow 0 \\ \sqrt{6}/6 & -1 \rightarrow -1 \end{cases} \\ (\text{I a, } l = 1, j = 0 | \vec{\Omega}_{el}^1 | 0) &= \frac{ieA\hbar}{3Mc} J_1 \cdot \begin{cases} & 0 \rightarrow 0 \end{cases} \\ (\text{I b, } j = 1 | \vec{\Omega}_{el}^1 | 0) &= \frac{ieA\hbar}{Mc} J_1 \cdot \begin{cases} \sqrt{6}/6 & 1 \rightarrow 1 \\ -\sqrt{6}/6 & -1 \rightarrow -1 \end{cases} \end{aligned} \right\} (30)$$

with

$$J_1 = \int_0^\infty dr R_I(1) \left(\frac{d}{dr} - \frac{1}{r} \right) R_I(0).$$

The second term of (29) is treated in the same way; the matrix elements are obtained from (30) by replacing

$$\frac{\hbar}{Mc} J_1 \quad \text{by} \quad -\frac{g_1^2 + g_2^2}{2\pi\hbar c} J_2,$$

where

$$J_2 = \int_0^\infty dr e^{-zr} R_I(1) R_I(0).$$

Writing down the equations for $R_I(0)$ and $R_I(1)$ in accordance with (4) and (5) and multiplying the first with $R_I(1)r$, the second with $R_I(0)r$, it is after subtraction and partial integration easy to see that

$$\frac{\nu}{2c} J = -\frac{\hbar}{Mc} J_1 + \frac{g_1^2 + g_2^2}{2\pi\hbar c} J_2.$$

For J , cf. p. 13. Therefore,

$$(F | \bar{\Omega}_{\text{el}} | 0) = -(F | \vec{\mathcal{E}} \vec{P}_{\text{nuc}} | 0) \quad (31)$$

and the following expression for Φ^{el} is equivalent with (18 b):

$$\Phi^{\text{el}} = \frac{\pi e^2 c}{3\nu} \left| \frac{\hbar}{Mc} J_1 - \frac{g_1^2 + g_2^2}{2\pi\hbar c} J_2 \right|^2. \quad (32)$$

It can also be seen more directly that the matrix elements corresponding to P_{exch} vanish in our approximation; for these are all proportional to

$$\frac{\nu}{zc} \cdot A \cdot \frac{g_1 g_2}{\hbar c} J_2$$

and thus are obviously of higher order in the velocities than the matrix elements corresponding to the second term of (29).

As (18 b) and (32), of course, only would give identical numerical results if exact explicit expressions for the radial wave-functions are used, this provides a check as to the reliability of the approximate expressions for these functions proposed by various authors. Previously¹⁵⁾, we had employed for $R_I(0)$ WILSON's result²⁵⁾ and it appeared that then the cross-section obtained from (32) for $\hbar\nu = 2,64 \text{ MeV}$ is 4,5 times the corresponding quantity obtained from (18 b), if $M_m = 0,1 M$; it is thus quite impossible in this case to predict anything with regard to such a sensitive effect as the angular distribution. Using (25), this ratio becomes 0,7 for the same energy.

In order to understand better the origin of these discrepancies we have computed the ratio of the two expressions for Φ^{el} at the photo-electric threshold: $\hbar\nu = E_0$. Here, the Bessel-function representation for $R_I(1)$ is certainly accurate so that, in this limiting case, the remaining deviation should be entirely ascribed to the ground state function. The threshold ratio is found to be 1,5 for WILSON'S function and 0,94 if (25) is used. HULTHÉN'S expression is therefore a much better approximation. The larger deviations for energies greater than E_0 will be partly due to the inaccuracy of $R_I(1)$; in this connection, it should be remarked that the values of the matrix elements may be very sensitive even for small changes in the radial functions.

From the above it seems reasonable to assume that the values for Φ^{el} given in § 4 are too large for energies not much greater than E_0 . The agreement of the theoretical values for the cross-section with experiment will therefore be better if more accurate approximations for the wave-functions are used, while the results with regard to $\Phi_{\parallel}/\Phi_{\perp}$ would remain satisfactory. In fact, it follows from the results stated for 2,64 MeV that this quantity would lie between 5,5 % and 8 % for $\alpha = 200$ and between 7,5 % and 10,5 % for $\alpha = 300$. As to the meson rest mass, one might infer from this result that $\alpha = 200$ is a more probable value than $\alpha = 300$.

Finally, it should be noted that (31) also holds in a pure vector or pure pseudoscalar meson theory, provided the dipole interaction potential (including cut-off) may be considered a perturbation and the contribution of the 3D wave-function to the ground state may be disregarded.

Note 2. The electric quadrupole effect. In accordance with (10) and (13) and remembering the assumption on the direction of propagation and polarization of the γ -rays, the operator of the quadrupole transitions may be written as

$$\Omega_{\text{quadr}} = -\frac{i\nu eE}{16c} [2 - (\tau_{\mathbf{3}}^{(1)} + \tau_{\mathbf{3}}^{(2)})] xz.$$

The only allowed transitions are to D -states that are anti-symmetric in isotopic spin:

$$\text{I a, } l = 2, j = 3 \quad ({}^3D_3)$$

$$\text{I a, } l = 2, j = 1 \quad ({}^3D_1)$$

$$\text{I b, } l = j = 2 \quad ({}^3D_2).$$

There is therefore no interference with the electric or magnetic dipole transitions. For these states, $\Omega_{\text{quadr}} = -i\nu e E x z / 8c$ and

$$(\text{I a, } l = 2, j = 3 | \Omega_{\text{quadr}} | 0) = -\frac{i\nu e E}{60c} \cdot G \cdot \begin{cases} \sqrt{5}/2 & 1 \rightarrow 2 \\ -\sqrt{6}/4 & 1 \rightarrow 0 \\ 1 & 0 \rightarrow 1 \\ -1 & 0 \rightarrow -1 \\ \sqrt{6}/4 & -1 \rightarrow 0 \\ -\sqrt{5}/2 & -1 \rightarrow -2 \end{cases}$$

$$(\text{I a, } l = 2, j = 1 | \Omega_{\text{quadr}} | 0) = \frac{i\nu e E}{80c} \cdot G \cdot \begin{cases} 1 & 1 \rightarrow 0 \\ 1 & 0 \rightarrow 1 \\ -1 & 0 \rightarrow -1 \\ -1 & -1 \rightarrow 0 \end{cases}$$

$$(\text{I b, } l = j = 2 | \Omega_{\text{quadr}} | 0) = \frac{i\nu e E}{240c} \cdot G \cdot \begin{cases} \sqrt{10} & 1 \rightarrow 2 \\ -\sqrt{15} & 1 \rightarrow 0 \\ -\sqrt{5} & 0 \rightarrow 1 \\ -\sqrt{5} & 0 \rightarrow -1 \\ -\sqrt{15} & -1 \rightarrow 0 \\ \sqrt{10} & -1 \rightarrow -2 \end{cases}$$

with

$$G = \int_0^\infty dr R_I(2) R_I(0) r^2.$$

The contributions of the different magnetic substates of the ground state appear to be equal, as in the electric dipole case. Averaging over directions of polarization gives a factor 1/2. The result is

$$d\Phi_{\text{quadr}} = \frac{e^2}{512} \cdot \left(\frac{\nu}{c}\right)^3 |G|^2 \sin^2 \theta \cos^2 \theta \cdot \sin \theta d\theta d\psi,$$

$$\Phi_{\text{quadr}} = \frac{\pi e^2}{960} \left(\frac{\nu}{c}\right)^3 |G|^2.$$

This expression would also have been obtained in the old theory; the electric dipole and quadrupole cross-sections thus both have the same general form as in the old theory.

For $R_I(2)$ we have to take $\lambda (kr)^{1/2} J_{3/2}(kr)$. Estimations show Φ_{quadr} to be at most of the order of 10^{-29} cm²; as this effect has, furthermore, no bearing on the quantity (1), these transitions are of no practical interest. The electric quadrupole capture is of course negligible, since a jump of l from 2 to 0 would be involved, which is highly improbable for small energies²⁶).

I am much indebted to prof. L. ROSENFELD for his interest in this work and for many helpful discussions.

Instituut voor theoretische Natuurkunde, Rijks-Universiteit Utrecht, Holland.

References.

1. J. CHADWICK and M. GOLDHABER, *Nature* **134**, 237, 1934.
2. H. A. BETHE and R. PEIERLS, *Proc. Roy. Soc. A* **148**, 146, 1935.
3. H. MASSEY and C. MOHR, *Proc. Roy. Soc. A* **148**, 206, 1935; cf. also H. HALL, *Phys. Rev.* **49**, 401, 1936.
4. E. FERMI, *Phys. Rev.* **48**, 570, 1935.
5. G. BREIT and E. U. CONDON, *Phys. Rev.* **49**, 904, 1936; **51**, 56, 1937.
6. H. v. HALBAN, *Nature* **141**, 644, 1938.
7. W. RARITA and J. SCHWINGER, *Phys. Rev.* **59**, 436, 1941.
8. J. CHADWICK, N. FEATHER and E. BRETSCHER, *Proc. Roy. Soc. A* **163**, 366, 1937.
9. C. MØLLER and L. ROSENFELD, *D. Kgl. Danske Vidensk. Selskab, Math.-fys. Medd.*, **XVII**, 8, 1940.
10. H. FRÖHLICH, W. HEITLER and B. KAHN, *Proc. Roy. Soc. A* **174**, 85, 1940.
11. N. KEMMER, *Helv. Phys. Acta* **10**, 47, 1937.
12. H. A. BETHE and R. F. BACHER, *Rev. Mod. Phys.* **8**, 82, 1936.
13. C. MØLLER and L. ROSENFELD, *D. Kgl. Danske Vidensk. Selskab, Mat.-fys. Medd.* **XX**, 12, 1943.
14. A. PAIS, *Physica* **9**, 407, 1942.
15. A. PAIS, Thesis, Utrecht 1941.
16. H. A. BETHE, *Ann. d. Phys.* **4**, 443, 1930.
17. L. HULTHÉN, *Arkiv. f. Mat. Astr. och. Fys.* **28 A**, fasc. 5, 1942.
18. L. HULTHÉN, *ibid.* **29 B**, fasc. 1, 1942.
19. J. A. ALLEN and N. M. SMITH, *Phys. Rev.* **59**, 618, 1941.
20. W. RARITA & J. SCHWINGER, *Phys. Rev.* **59**, 556, 1941.
21. H. A. BETHE, *Phys. Rev.* **57**, 260, 390, 1940.
22. O. R. FRISCH, H. v. HALBAN and J. KOCH, *D. Kgl. Danske Vid. Selskab, Math.-fys. Medd.* **XV**, 10, 1937.
23. E. AMALDI and E. FERMI, *Phys. Rev.* **50**, 899, 1936.
24. P. M. MORSE, J. B. FISK and L. I. SCHIFF, *Phys. Rev.* **50**, 748, 1936.
25. A. H. WILSON, *Proc. Cambr. Phil. Soc.* **34**, 365, 1938.
26. Cf. N. F. MOTT and H. S. W. MASSEY, *Theory of Atomic Collisions*, Oxford 1933, p. 25.
27. F. E. MYERS and L. C. VAN ATTA, *Phys. Rev.* **61**, 19, 1942.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 18

CONTRIBUTION TO
THE THEORY OF ANALYTIC ALMOST
PERIODIC FUNCTIONS

ON THE BEHAVIOUR OF AN ANALYTIC ALMOST
PERIODIC FUNCTION IN THE NEIGHBOURHOOD OF A
BOUNDARY FOR ITS ALMOST PERIODICITY

BY

HARALD BOHR



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

CONTENTS

Chapter I. Introductory remarks	3
— II. Functions almost periodic in a strip $\langle a, \infty \rangle$	12
— III. Functions almost periodic in a strip $\langle a, \beta \rangle$, where $\beta < \infty$..	17
— IV. A gap theorem on the almost periodicity of Dirichlet series ..	31

CHAPTER I.

Introductory remarks.

We begin by recalling some of the fundamental notions and theorems in the theory of the almost periodic functions of a real as well as of a complex variable. We wish, however, essentially to confine ourselves to what is indispensable in the following¹.

A (complex) function $F(t)$ of a real variable t , continuous in $-\infty < t < \infty$, is called almost periodic, if to any given $\varepsilon > 0$ there exists a relatively dense set of translation numbers $\tau = \tau(\varepsilon)$, i. e. of numbers τ which satisfy the inequality

$$|F(t+\tau) - F(t)| \leq \varepsilon \quad \text{in} \quad -\infty < t < \infty.$$

Here, a set of real numbers is called relatively dense, if there exists a length L such that any interval of length L contains at least one number of the set.

An almost periodic function $F(t)$ is bounded and uniformly continuous in $-\infty < t < \infty$, and the sum (and product) of two almost periodic functions proves again to be almost periodic.

¹ In its main features, the theory of the almost periodic functions was developed by the author in three papers in *Acta Mathematica* (vol. 45, 46, 47) under the common title "Zur Theorie der fastperiodischen Funktionen". Especially the last of these articles dealing with the functions of a complex variable is of importance for the present paper. Furthermore, we shall also make use of some of BOCHNER's results given in his important paper (Beiträge zur Theorie der fastperiodischen Funktionen, I. Teil, *Math. Ann.* vol. 96). However, it is not supposed that the papers are known to the reader and, therefore, the theorems applied will be directly formulated. For further information, cf. one of the monographs, A. S. BESICOVITCH: *Almost periodic functions*, Cambridge 1932; J. FAVARD: *Leçons sur les fonctions presque-périodiques*, Paris 1933; H. BOHR: *Fastperiodische Funktionen*, Berlin 1932.

For an arbitrary fixed real τ , let us consider the quantity

$$v(\tau) = \text{u. b. } |F(t+\tau) - F(t)|, \\ -\infty < t < \infty$$

This quantity, by the author originally denoted as the minimum error (corresponding to the given τ) of the function, was studied in detail by BOCHNER as a function of τ and, now, the function $v(\tau)$, $-\infty < \tau < \infty$, is usually called the translation function of the given function $F(t)$. The function $v(\tau)$ is again an almost periodic function and its own translation function. A set $\{F(t)\}$ of almost periodic functions can be "majorised", if there exists an almost periodic function $F_0(t)$, which is denoted a majorant of the set $\{F(t)\}$, with the property that the translation function $v(\tau) = v_{F_0}(\tau)$ of any given function $F(t)$ of the set satisfies for all τ the inequality

$$v(\tau) \leq v_0(\tau),$$

where $v_0(\tau)$ denotes the translation function of $F_0(t)$. A necessary and sufficient condition that the set of almost periodic functions $\{F(t)\}$ can be majorised is that the functions of the set are "uniformly" uniformly continuous and uniformly almost periodic. Here, a set of almost periodic functions is said to be uniformly almost periodic, if to any $\varepsilon > 0$ there exists a relatively dense set of numbers τ which are translation numbers corresponding to ε for any function of the set. A finite set of almost periodic functions can always be majorised. From this it becomes obvious that the sum of a finite number of almost periodic functions is again almost periodic.

A principal theorem in the theory of almost periodic functions of a real variable states that the class of all almost periodic functions is identical with the class of functions which, uniformly for all t , can be approximated by finite sums of the form $\sum_1^N a_n e^{i\lambda_n t}$, where the coefficients a_n are complex numbers, while the exponents λ_n are real numbers. With any almost periodic function $F(t)$ there is associated a Fourier series

$$F(t) \approx \sum A_n e^{iA_n t},$$

where the exponents A_n form the countably infinite set of values of λ for which the mean value

$$a(\lambda) = M\{F(t)e^{-i\lambda t}\} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T F(t)e^{-i\lambda t} dt$$

differs from zero, while the coefficients A_n are the corresponding values of $a(A_n)$. Starting from the Fourier series of an almost periodic function, by suitable summation methods finite sums of the form $\sum_1^N B_n e^{iA_n t}$ can be deduced (all exponents of which are Fourier exponents of $F(t)$) which converge uniformly to $F(t)$.

If $F(t)$ is almost periodic with the Fourier series $\sum A_n e^{iA_n t}$, and p is an arbitrary positive number, the (possibly empty) series $\sum' A_n e^{iA_n t}$ consisting of just the terms $A_n e^{iA_n t}$ in the original series with integral multipla of $\frac{2\pi}{p}$ as exponents, is the Fourier series of a continuous, purely periodic function $P(t)$. This function $P(t)$, which was especially considered by BOCHNER, will be denoted as the "periodic component of $F(t)$ belonging to the period p ". This component can also be determined from the sequence of almost periodic functions

$$F_n(t) = \frac{F(t+p) + F(t+2p) + \dots + F(t+np)}{n},$$

as this sequence $F_n(t)$ converges for $n \rightarrow \infty$ to $P(t)$, uniformly in the whole interval $-\infty < t < \infty$. From the limit equation $P(t) = \lim F_n(t)$ it results immediately that

$$\text{u. b. } |P(t)| \leq \text{u. b. } |F(t)|, \quad -\infty < t < \infty$$

Let us now briefly recall some notions and theorems concerning the almost periodic functions of a complex variable.

A function $f(s) = f(\sigma + it)$, analytic in a vertical strip $\alpha < \sigma < \beta$ ($-\infty \leq \alpha < \beta \leq \infty$), is called almost periodic in (α, β) , if to any $\varepsilon > 0$ there exists a relatively dense set of translation numbers $\tau = \tau(\varepsilon)$ satisfying the inequality

$$|f(s + i\tau) - f(s)| \leq \varepsilon \quad \text{in the strip } \alpha < \sigma < \beta.$$

In other words, we require that for any fixed σ in the interval $\alpha < \sigma < \beta$ the function $F_\sigma(t) = f(\sigma + it)$ is an almost periodic

function of the real variable t , and that the functions of the set $\{F_\sigma(t)\}$ corresponding to $\alpha < \sigma < \beta$ are uniformly almost periodic. But, generally, this set cannot be majorised, since the functions (on account of the behaviour near the boundary of the strip) need not be "uniformly" uniformly continuous functions of t .

A function $f(s)$, analytic in $\alpha < \sigma < \beta$ ($-\infty \leq \alpha < \beta \leq \infty$), is called almost periodic in $[\alpha, \beta]$, if it is almost periodic in every substrip (α_1, β_1) where $\alpha < \alpha_1 < \beta_1 < \beta$. We also use mixed brackets and thus speak of a function almost periodic in $[\alpha, \beta)$.

With each function $f(s)$, almost periodic in $[\alpha, \beta]$, is associated a Dirichlet series

$$f(s) \sim \sum A_n e^{A_n s}$$

with real exponents A_n which, for any fixed σ in $\alpha < \sigma < \beta$, gives the Fourier series of the almost periodic function $F_\sigma(t) = f(\sigma + it)$ of the real variable t . The set of all functions, almost periodic in $[\alpha, \beta]$, is identical with the set of the functions which can be approximated uniformly in $[\alpha, \beta]$ (i. e. uniformly in every substrip $\alpha_1 < \sigma < \beta_1$) by finite sums of the form $\sum a_n e^{\lambda_n s}$, where the λ_n are real numbers. If $f_1(s)$ and $f_2(s)$ are two analytic functions which both are almost periodic in $[\alpha, \beta]$, their sum $f_1(s) + f_2(s)$ is again almost periodic in $[\alpha, \beta]$. A corresponding general theorem does not hold for functions almost periodic in (α, β) . Thus, the functions e^s and $e^{s\sqrt{2}}$ are both almost periodic (even purely periodic) in $(-\infty, \infty)$, whereas their sum $f(s)$ is almost periodic in $(-\infty, \infty]$, but not in $(-\infty, \infty)$; for any real $\tau \neq 0$ it is even valid that $|f(s + i\tau) - f(s)| \rightarrow \infty$ for $\sigma \rightarrow \infty$, uniformly in t . Another simple example is given by the two geometrical series

$$f_1(s) = \sum_{n=0}^{\infty} e^{ns} = \frac{1}{1 - e^s} \quad \text{and} \quad f_2(s) = \sum_{n=0}^{\infty} e^{\sqrt{2}ns} = \frac{1}{1 - e^{\sqrt{2}s}},$$

which both are almost periodic (even purely periodic) in $(-\infty, 0)$, while their sum is almost periodic in $(-\infty, 0]$, but not in $(-\infty, 0)$; for, as the function $f(s) = f_1(s) + f_2(s)$ has poles in (and only in) all points of the two arithmetic progressions $2m\pi i$ and $\sqrt{2}m\pi i$,

there does not exist any number $\tau \neq 0$ for which the difference $f(s+i\tau) - f(s)$ is bounded in the whole half plane $-\infty < \sigma < 0$ (as the set of poles is not transformed into itself by any translation) so that $f(s)$ for no ε has other translation numbers in $(-\infty, 0)$ than the trivial $\tau = 0$.

It is the aim of this paper to investigate the behaviour of a function $f(s)$, almost periodic in a strip (α, β) (and not only in $[\alpha, \beta]$), in the immediate neighbourhood of one of the boundaries of the strip, for instance the right one. We shall assume the strip cut off (if necessary) on the left in order to avoid any difficulties on the left boundary of the strip.

For our purpose it will be convenient also to introduce the notion of a function almost periodic in $\{\alpha, \beta\}$ (and correspondingly in (α, β) etc.). For $-\infty < \alpha < \beta < \infty$, the function $f(s)$ is called almost periodic in $\{\alpha, \beta\}$, if $f(s)$ is continuous in the closed strip $\alpha \leq \sigma \leq \beta$ and analytic and almost periodic in the open strip (α, β) . Besides, it is clear that every translation number $\tau = \tau(\varepsilon)$ corresponding to $f(s)$ in the open strip (α, β) as a matter of course is a translation number $\tau(\varepsilon)$ for $f(s)$ in the closed strip $\{\alpha, \beta\}$, as the inequality

$$|f(s+i\tau) - f(s)| \leq \varepsilon \quad \text{in} \quad \alpha < \sigma < \beta,$$

on account of the continuity of $f(s)$ in $\alpha \leq \sigma \leq \beta$, involves the inequality

$$|f(s+i\tau) - f(s)| \leq \varepsilon \quad \text{in} \quad \alpha \leq \sigma \leq \beta.$$

By simple conclusions (known from the theory of the almost periodic functions of a real variable) we realize immediately that a function almost periodic in $\{\alpha, \beta\}$ is bounded and uniformly continuous in $\alpha \leq \sigma \leq \beta$. Thus, if $f(s)$ is almost periodic in $\{\alpha, \beta\}$, the functions $F_\sigma(t) = f(\sigma + it)$ ($\alpha \leq \sigma \leq \beta$) are not only uniformly almost periodic, but also "uniformly" uniformly continuous, i. e. the set $\{F_\sigma(t)\}$ ($\alpha \leq \sigma \leq \beta$) can be majorised. From this, it readily results that the sum of two functions $f_1(s)$ and $f_2(s)$, both almost periodic in $\{\alpha, \beta\}$, again is almost periodic in $\{\alpha, \beta\}$. We add, without going into details, that it is easy to prove that, if $f_1(s)$ is almost periodic in $\{\alpha, \beta\}$ and $f_2(s)$ is almost periodic in $\{\alpha, \beta\}$, the sum $f_1(s) + f_2(s)$ is almost periodic in $\{\alpha, \beta\}$ and not only in (α, β) .

But it is important to call attention to the fact that the sum of two functions, almost periodic in $\{\alpha, \beta\}$, generally is almost periodic only in $\{\alpha, \beta\}$, but not in $\{\alpha, \beta\}$; thus, the two functions mentioned above, $\frac{1}{1-e^s}$ and $\frac{1}{1-e^{i/2s}}$, are both almost periodic, for instance in $\{-1, 0\}$, whereas the sum is not.

In the present paper, we shall study the functions, almost periodic in a strip of the type $\{\alpha, \beta\}$, especially the unbounded functions of this type.

Of special importance for our investigation of a function almost periodic in $\{\alpha, \beta\}$ is the set F of all translation numbers τ (where no ε is prescribed), i. e. the set of all real numbers τ for which the difference $f(s+i\tau)-f(s)$ is bounded in $\alpha \leq \sigma < \beta$. While, for a function almost periodic in $\{\alpha, \beta\}$, this set F consists of all real numbers, this is not necessarily the case for a function almost periodic in $\{\alpha, \beta\}$; thus for the function $\frac{1}{1-e^{s^2}}$ purely periodic in $\{-1, 0\}$, the set F consists of the numbers $2m\pi$. It will be shown below that the set F consists of all real numbers only in the trivial case, where the function $f(s)$ itself is bounded in $\{\alpha, \beta\}$. Let us call the function

$$v(\tau) = \text{u. b. } |f(s+i\tau) - f(s)|, \\ s \text{ in } \{\alpha, \beta\}$$

which is defined in the set F , the translation function of $f(s)$ in $\{\alpha, \beta\}$. The set F contains together with τ also $-\tau$, and we have $v(-\tau) = v(\tau)$; moreover F contains together with τ_1 and τ_2 also $\tau_1 + \tau_2$, and the inequality $v(\tau_1 + \tau_2) \leq v(\tau_1) + v(\tau_2)$ is valid. The set F is a module, as it contains together with τ_1 and τ_2 also $\tau_1 - \tau_2$ and it contains other numbers than 0 (as a subset it contains for instance the relatively dense set of all numbers $\tau = \tau(1)$); we shall denote F the "translation module" of the function $f(s)$ in $\{\alpha, \beta\}$.

Moreover, the process mentioned above of separating the periodic component with a given period from an almost periodic function can be transferred from functions of a real variable to functions of a complex variable. For a function $f(s)$, almost periodic in $\{\alpha, \beta\}$, and an arbitrarily chosen $p > 0$, we find that the sequence of the functions (evidently almost periodic in $\{\alpha, \beta\}$)

$$f_n(s) = \frac{f(s + ip) + f(s + 2ip) + \dots + f(s + nip)}{n}$$

converges for $n \rightarrow \infty$ in the whole strip $\{\alpha, \beta\}$ and even uniformly in each substrip $\{\alpha, \gamma\}$, where $\alpha < \gamma < \beta$, to a function $p(s)$ purely periodic with the period ip in $\{\alpha, \beta\}$ whose Dirichlet series (Laurent series) $\sum B_m e^{\frac{2\pi}{p}ms}$ consists just of those terms $A_n e^{A_n s}$ in the Dirichlet series of the given function $f(s)$ the exponents A_n of which are integral multipla of $\frac{2\pi}{p}$. The transition from the real to the complex case is immediate, if we only observe that uniform convergence of the sequence $f_n(s)$ on the two lines $\sigma = \alpha$ and $\sigma = \gamma$ involves that $f_n(s)$ converges uniformly in the whole strip $\alpha \leq \sigma \leq \gamma$, as, according to the theorem of PHRAGMEN-LINDELÖF, u. b. $|f_{n_1}(s) - f_{n_2}(s)|$ remains unchanged, whether s varies in the whole strip $\alpha \leq \sigma \leq \gamma$ or only on the two boundaries $\sigma = \alpha$ and $\sigma = \gamma$ of the strip. Furthermore, if $f(s)$ is bounded in $\{\alpha, \beta\}$, the component $p(s)$ is also bounded, and the inequality

$$\text{u. b. } |p(s)| \underset{s \text{ in } \{\alpha, \beta\}}{\leq} \text{u. b. } |f(s)|$$

is valid.

After this introducing Chapter I, the present paper is divided in three Chapters.

In Chapter II, the case $\beta = \infty$, i. e. the functions almost periodic in $\{\alpha, \infty\}$, is treated. If the function $f(s)$ is bounded in $\{\alpha, \infty\}$, its behaviour for $\sigma \rightarrow \infty$ is extremely simple in consequence of well-known theorems on almost periodic analytic functions. But also for unbounded functions, almost periodic in $\{\alpha, \infty\}$, the situation is very perspicuous in view of the fact that the translation module F here always proves to be discrete, i. e. forms an arithmetic progression; in fact, a general "splitting theorem" holds which states that every such function $f(s)$ can be written (and essentially only in one way) as a sum of a function $p(s)$, unbounded and periodic in $\{\alpha, \infty\}$, and a func-

tion $b(s)$, bounded and almost periodic in $\{\alpha, \infty\}$. Moreover, also the converse is valid, viz. that the sum of a function $p(s)$, unbounded and periodic in $\{\alpha, \infty\}$, and a function $b(s)$, bounded and almost periodic in $\{\alpha, \infty\}$, is always almost periodic in $\{\alpha, \infty\}$.

Chapter III deals with the case, where β is finite; here, we may suppose that $\beta = 0$ so that we have to do with functions which are almost periodic in a strip $\{\alpha, 0\}$. The functions which are bounded and almost periodic in $\{\alpha, 0\}$ are only shortly discussed, as the main object is the study of the functions unbounded in $\{\alpha, 0\}$. For such an unbounded function, we begin proving that its translation module F cannot contain all real numbers. Subsequently, we distinguish between the case where the set F is discrete and that where F is everywhere dense. The first case, where F is an arithmetic progression, does not cause any difficulties; a general splitting theorem is valid here (completely analogous to that holding in the case of $\beta = \infty$), as $f(s)$ can be splitted—and practically uniquely—into a function $p(s)$ purely periodic in $\{\alpha, 0\}$ and a function $b(s)$ bounded and almost periodic in $\{\alpha, 0\}$. Next, the other (essentially more difficult) case is considered, where the translation module F is everywhere dense; in this case, the line $\sigma = 0$ is always an essentially singular line for the analytic function $f(s)$. The main question is, whether also here a general splitting theorem holds, analogous to that valid in the other cases. It is proved that this is not the case. In a decisive way we use a “gap theorem” concerning Dirichlet series $\sum a_n e^{\lambda_n s}$, convergent for $\sigma < 0$, where the exponents form an increasing sequence of positive numbers which increase “very strongly” to the infinite; this theorem states that the function $f(s)$ represented by such a series always is almost periodic in the whole strip $(-\infty, 0)$, and not only in $(-\infty, 0]$.

Finally, in Chapter IV, the proof of this gap theorem is given. Here, the treatment of Dirichlet series with strongly increasing exponents is extended somewhat further than necessary for the proper purpose of this paper. Thus, in order to throw light on the nature of the methods used, a new proof of a special case of the so-called HADAMARD gap theorem for Dirichlet series $\sum a_n e^{\lambda_n s}$ with the convergence half plane $\sigma < 0$

is given which states that the convergence line $\sigma = 0$ always is an essentially singular line for the analytic function represented by the series, if the exponents increase rapidly enough. In the case where the λ are integers (and we are concerned with a power series) this simple proof has already been communicated by the author in a paper¹ written in Danish.

¹ Om den Hadamard'ske "Hulsætning", Matematisk Tidsskrift, B (1919).
See also the paper: Om Potensrækker med Huller, Matematisk Tidsskrift, B (1942).

CHAPTER II.

Functions almost periodic in a strip $\{\alpha, \infty\}$.

We shall now investigate, how an analytic function $f(s)$ which is almost periodic in $\{\alpha, \infty\}$ behaves for $\sigma \rightarrow \infty$.

We begin by recalling the following theorem which was proved in the third paper in *Acta Mathematica* cited above. If $f(s)$ is almost periodic in $\{\alpha, \infty\}$, a necessary and sufficient condition that $f(s)$ is bounded for $\sigma \rightarrow \infty$, i. e. bounded in $\{\alpha, \infty\}$, is that all the Dirichlet exponents λ_n of the function are ≤ 0 ; if this condition is satisfied, the function is almost periodic not only in $\{\alpha, \infty\}$, but also in $\{\alpha, \infty\}$, and for $\sigma \rightarrow \infty$ uniformly in t it approaches a definite limit (viz. the constant term in the Dirichlet development of the function); in this case the function is called regular in the point $\sigma = \infty$. In view of a later application, we observe that the difference between two functions, bounded and almost periodic in $\{\alpha, \infty\}$, again is almost periodic in $\{\alpha, \infty\}$, and not only in $\{\alpha, \infty\}$.

Thus, the functions almost periodic in a strip $\{\alpha, \infty\}$, which are bounded for $\sigma \rightarrow \infty$, are simply the almost periodic functions regular in $\sigma = \infty$; an investigation is therefore only demanded in the case, where the function $f(s)$, almost periodic in $\{\alpha, \infty\}$, is unbounded for $\sigma \rightarrow \infty$ (and, therefore, among its Dirichlet exponents has positive ones).

As the most simple type of such a function, we have obviously a function unbounded and purely periodic in $\{\alpha, \infty\}$.

Let us begin by proving the following theorem: If $p(s)$ and $b(s)$ are two functions, almost periodic in $\{\alpha, \infty\}$, of which $p(s)$ is purely periodic, while $b(s)$ is bounded in $\{\alpha, \infty\}$, their sum

$$f(s) = p(s) + b(s)$$

is again almost periodic in $\{\alpha, \infty\}$.

The theorem is not quite trivial, as it is not valid that the sum of any two functions almost periodic in (α, ∞) is again almost periodic in $\{\alpha, \infty\}$ (but only in $\{\alpha, \infty\}$); however, it is easy enough to prove. It is the task to show that to any arbitrarily given $\varepsilon > 0$ a relatively dense set of real numbers τ exists for which

$$|f(s + i\tau) - f(s)| \leq \varepsilon \quad \text{in the whole strip } \alpha \leq \sigma < \infty.$$

As $b(s)$ tends to a limit for $\sigma \rightarrow \infty$ (uniformly in t), to the given ε we can first determine a number $\gamma = \gamma(\varepsilon)$ so that any real number τ is a translation number $\tau(\varepsilon)$ for the function $b(s)$ in the half plane $\{\gamma, \infty\}$; from this, it follows that the function $b(s)$ has quite the same translation numbers $\tau(\varepsilon)$, whether we consider it in the half plane $\{\alpha, \infty\}$ or in the strip $\{\alpha, \gamma\}$. In consequence of the earlier cited majorising property of the almost periodic functions $b(\sigma + it)$ ($\alpha \leq \sigma \leq \gamma$) the set of translation numbers $\tau(\varepsilon)$ of $b(s)$ in $\{\alpha, \gamma\}$ has, however, a relatively dense intersection with any arithmetic progression—because, as is well-known, this is the case for the translation numbers of a single almost periodic function of a real variable. If, as difference in this arithmetic progression just $p > 0$ is chosen, where ip is a period of the given periodic function $p(s)$, any number τ in the previously mentioned relatively dense intersection is a translation number of the sum $f(s) = p(s) + b(s)$ in $\{\alpha, \infty\}$, as

$$p(s + i\tau) = p(s) \quad \text{and} \quad |b(s + i\tau) - b(s)| \leq \varepsilon \quad \text{in } \alpha \leq \sigma < \infty.$$

As we shall see, we have thereby actually exhausted all possibilities for a function almost periodic and unbounded in $\{\alpha, \infty\}$, since the following inverse theorem is valid.

Splitting theorem: *Every function $f(s)$ unbounded and almost periodic in $\{\alpha, \infty\}$ can—and essentially only in one way—be written as a sum*

$$f(s) = p(s) + b(s),$$

where $p(s)$ is periodic and unbounded in $\{\alpha, \infty\}$, while $b(s)$ is bounded and almost periodic in $\{\alpha, \infty\}$.

If we have a splitting of $f(s)$ as that stated in the theorem, and if ip is a period of the periodic term $p(s)$, the number p must necessarily belong to the translation module F of the function $f(s)$ in $\{\alpha, \infty\}$, as in $\alpha \leq \sigma < \infty$ the inequality

$$|f(s+ip) - f(s)| \leq 2B$$

holds, where B means u. b. $|b(s)|$ in $\alpha \leq \sigma < \infty$. It is, therefore, natural—and may also be interesting in itself—to study primarily this translation module F , i. e. to try to determine the numbers τ for which the difference $f(s+i\tau) - f(s)$ is bounded in $\{\alpha, \infty\}$. Since the Dirichlet development of this difference is determined as the difference between the Dirichlet developments of $f(s+i\tau)$ and $f(s)$, i. e. given by

$$f(s+i\tau) - f(s) \sim \sum A_n (e^{iA_n\tau} - 1) e^{A_ns},$$

the theorem quoted in the beginning of this Chapter, however, involves immediately the following necessary and sufficient condition for the (anyhow in $\{\alpha, \infty\}$) almost periodic function $f(s+i\tau) - f(s)$ to be bounded in $\{\alpha, \infty\}$: In the Dirichlet series mentioned above must not occur any term with a positive exponent, i. e. for any positive Dirichlet exponent A_n of the given function $f(s)$ (necessarily occurring, because $f(s)$ is unbounded in $\{\alpha, \infty\}$) it must be valid that $e^{iA_n\tau} - 1 = 0$, i. e.

$$A_n\tau \equiv 0 \pmod{2\pi} \quad \text{for any } A_n > 0.$$

Thus, the numbers τ in the translation module F of $f(s)$ are just the numbers τ which are multipla of all the numbers $\frac{2\pi}{A_n}$, where A_n runs through the positive Dirichlet exponents of $f(s)$. Hereby (only applying that any τ has to be an integral multiple of one of these numbers $\frac{2\pi}{A_{n_0}}$) the translation module appears to be discrete, i. e. consists of all the numbers of an arithmetic progression νq ($q > 0$, $\nu = 0, \pm 1, \dots$). After this, turning to the exponents A_n , we notice (as $\tau = q$ satisfies all the congruences above) that all positive exponents A_n must be multipla of the number $\frac{2\pi}{q}$ and, further (as $\tau = q$ is the smallest positive solution

of the congruences) that $\frac{2\pi}{q}$ is the greatest common divisor of the positive exponents \mathcal{A}_n .

It is now easy to accomplish the proof of the splitting theorem, since it can be demonstrated that, as a period of the periodic term in a splitting of the desired kind, we may even use the number of least absolute value which may be taken into consideration, viz. the number iq , where q is the smallest positive number in the translation module F . As $p(s)$ we may use the periodic component (which was introduced in Chapter I) of $f(s)$ in $\{\alpha, \infty)$ belonging to the period iq . For, as the Dirichlet development of this function $p(s)$, purely periodic in $\{\alpha, \infty)$, consists of the terms $A_n e^{i\mathcal{A}_n s}$ in the Dirichlet development of $f(s)$, for which \mathcal{A}_n is a multiple of $\frac{2\pi}{q}$, the Dirichlet development of $p(s)$ coincides as regards the terms with positive exponents (even with exponents ≥ 0) with the Dirichlet development of $f(s)$. Therefore, the Dirichlet development of the difference $b(s) = f(s) - p(s)$, almost periodic in $\{\alpha, \infty]$, can only contain terms with negative exponents; it follows that the function $b(s)$ actually is a function, bounded and almost periodic in $\{\alpha, \infty)$.

From the proof of the splitting theorem given above it is furthermore easy to decide to what degree the splitting is unique. Let us assume that

$$f(s) = p(s) + b(s)$$

is the "standard splitting" stated in the above proof, the period of $p(s)$ being the number iq where q is the smallest positive number in the translation module F , and $p(s)$ just being the periodic component of $f(s)$ belonging to that period iq ; moreover, let us assume that

$$f(s) = p^*(s) + b^*(s)$$

is another arbitrary splitting of $f(s)$ in $\{\alpha, \infty)$ of the kind stated in the theorem. As for the period ip of the periodic term $p^*(s)$ it holds that p is a number of the translation module F , i. e. a number of the form νq , the difference $\pi(s) = p^*(s) - p(s) = b(s) - b^*(s)$ must be a function, bounded and

periodic in $\{\alpha, \infty\}$, with a period of the form $i\nu q$. Conversely, however, it is valid that, if $\pi(s)$ is a function bounded and periodic in $\{\alpha, \infty\}$ with a period of the form $i\nu q$, we can use the function

$$p^*(s) = p(s) + \pi(s)$$

as a periodic splitting term; for, if we write $f(s)$ in the form

$$f(s) = (p(s) + \pi(s)) + (b(s) - \pi(s)),$$

the first term $p(s) + \pi(s)$ is periodic in $\{\alpha, \infty\}$, while the other term $b(s) - \pi(s)$ is almost periodic and bounded in $\{\alpha, \infty\}$, as the difference of two functions, almost periodic and bounded in $\{\alpha, \infty\}$.

Thus it is clear that the splitting is "essentially" unique.



CHAPTER III.

Functions almost periodic in a strip $\{\alpha, \beta\}$, where $\beta < \infty$.

Obviously it may be assumed, otherwise only applying the translation $s = s' + \beta$, that $\beta = 0$. In this chapter, the behaviour of a function, almost periodic in $\{\alpha, 0\}$ ($-\infty < \alpha < 0$), in the immediate neighbourhood to the left of the imaginary axis $\sigma = 0$ will be investigated.

Here, the situation is more complicated than in the case $\beta = \infty$, due to the fact that a function bounded and almost periodic in $\{\alpha, 0\}$ does not show a similar simple behaviour for $\sigma \rightarrow 0$ as a function bounded and almost periodic in $\{\alpha, \infty\}$ for $\sigma \rightarrow \infty$.

The functions, bounded and almost periodic in $\{\alpha, 0\}$, however, shall be treated briefly, since we are especially interested in the unbounded functions. If $f(s)$ is bounded and almost periodic in $\{\alpha, 0\}$ it has—already because it is bounded—according to a theorem by FATOU, for $\sigma \rightarrow 0$ a limit function $F(t) = f(it)$ in the sense that $f(\sigma + it)$ for $\sigma \rightarrow 0$ approaches a limit $F(t)$ for any t in $-\infty < t < \infty$ except in a set E of measure zero¹. Thus, if $\tau = \tau(\varepsilon)$ is an arbitrary translation number of $f(s)$ in $\{\alpha, 0\}$, i. e.

$$|f(s + it) - f(s)| \leq \varepsilon \quad \text{in} \quad \alpha \leq \sigma < 0,$$

¹ In its usual formulation FATOU's theorem deals with a function $f(z)$ bounded and analytic in a circle $|z| < 1$, and it states that, for almost all points z_0 on the boundary $|z| = 1$, $f(z)$ approaches a limit $f(z_0)$, if z tends to z_0 along the radius vector. When transforming the circle $|z| < 1$ by a linear substitution into the half plane $\sigma < 0$, we do not just arrive at the assertion mentioned in the text, as the radii vectori of the circle are not transformed into horizontal straight lines in the half plane, but into certain arcs of circles orthogonal to the boundary $\sigma = 0$. It is, however, obvious that a function $f(s)$, bounded and analytic in $\sigma < 0$, which approaches a limit g , if s approach-

we have for all t , except in a set of measure zero $E' = E'(\tau)$ (namely the sum of the set E and the set obtained from E by the translation $-\tau$), the inequality

$$|F(t+\tau) - F(t)| \leq \varepsilon.$$

From this it follows especially that the limit function $F(t) = f(it)$ is a function almost periodic in STEPANOFF'S sense, as for any of the mentioned translation numbers $\tau(\varepsilon)$ of $f(s)$ in $\{\alpha, 0\}$ holds the inequality

$$\text{u. b. } \int_{-\infty < t < \infty}^t |F(t+\tau) - F(t)| dt \leq \varepsilon.$$

However, we shall not study this limit function in detail, as such a study—analogue for instance to the study of a limit function of a function bounded and analytic in the unity circle $|z| < 1$ —lies beyond the scope of this paper.

In order to obtain information on the different possibilities and to have some conveniently simple examples at our disposal—before turning towards our proper problem, viz. the investigation of the functions almost periodic and unbounded in $\{\alpha, 0\}$ —we shall, however, mention some typical examples of functions bounded and almost periodic in $\{\alpha, 0\}$.

Example 1. It may, of course, happen that a function, bounded and almost periodic in $\{\alpha, 0\}$, is almost periodic also in $\{\alpha, 0\}$, although it cannot be continued analytically across the line $\sigma = 0$. This is the case, for instance, with the purely periodic function $\sum_1^{\infty} \frac{1}{n^2} e^{n^1 s}$ (as the unity circle is the natural boundary of the power series $\sum \frac{1}{n^2} z^{n^1}$).

Example 2. The function

$$g(s) = e^{\frac{e^s + 1}{e^s - 1}},$$

crosses a point $s_0 = it_0$ of the boundary along such an arc of a circle, also converges to g , if s tends to s_0 along the tangent $t = t_0$, simply because the inequality $|f(s)| < K$ for $\sigma < 0$ involves that $f'(s) = O\left(\frac{1}{\sigma}\right)$ for $\sigma \rightarrow 0$, while the vertical segment between the tangent and the circle is $O(\sigma^2)$ for $\sigma \rightarrow 0$.

also bounded in the half plane $\sigma < 0$ and periodic with the period $2\pi i$, is regular on the whole boundary $\sigma = 0$ except the points $2\pi im$ ($m = 0, \pm 1, \dots$); that the function $\varphi(s)$ is bounded for $\sigma < 0$, viz. $|\varphi(s)| < 1$, follows from the fact that $|e^s| < 1$ for $\sigma < 0$ and the exponent $\frac{e^s + 1}{e^s - 1}$ therefore has a negative real part in $\sigma < 0$ (as by the linear function $u = \frac{w + 1}{w - 1}$ the unity circle $|w| < 1$ is transferred to the half plane $\Re(u) < 0$). If s ranges over the segment $\sigma = 0$ ($0 < t < 2\pi$), the function e^s ranges over the unity circle and therefore $\frac{e^s + 1}{e^s - 1}$ over the whole imaginary axis; thus $\varphi(s)$ will circulate an infinite number of times on the unity circle for $t \rightarrow 0$ and $t \rightarrow 2\pi$. We have called attention to this bounded function $\varphi(s)$, because, for any sufficiently small ε , in any case for $\varepsilon < 1$, it has no other transiation numbers $\tau = \tau(\varepsilon)$ in $\{\alpha, 0\}$ (α an arbitrary negative number) than just the numbers (periods) $2\pi m$. For, if τ is an arbitrary number $\pm 2\pi m$, in the difference

$$\varphi(s + i\tau) - \varphi(s)$$

the first term $\varphi(s + i\tau)$ converges to $\varphi(i\tau)$ for $s \rightarrow 0$, while the other term $\varphi(s)$ can be made to converge to an arbitrarily assigned value of the unity circle by letting s approach 0 in a convenient way from inside the half plane $\sigma < 0$; this implies, however, that $|\varphi(s + i\tau) - \varphi(s)|$ obtains values greater than ε in the half plane $\sigma < 0$ (by the way in every half circle $|s| < \delta$, $\sigma < 0$).

Example 3. Already in the introductory Chapter we mentioned two functions, almost periodic in $\{\alpha, 0\}$, viz.

$$f_1(s) = \frac{1}{1 - e^s} \quad \text{and} \quad f_2(s) = \frac{1}{1 - e^{\sqrt{2}s}},$$

the sum of which is not almost periodic in $\{\alpha, 0\}$; however, these functions are not bounded. It is of interest that the mentioned conditions can also occur for two functions $f_1(s)$ and $f_2(s)$ bounded and almost periodic in $\{\alpha, 0\}$; hereby we have especially shown that there exist functions, for instance the sum $f_1(s) + f_2(s)$, which are almost periodic in $\{\alpha, 0\}$ and bounded

in $\{\alpha, 0\}$, but not almost periodic in $\{\alpha, 0\}$, in contrast to the case $\beta = \infty$. As an example we may use the two functions

$$f_1(s) = \varphi(s) \quad \text{and} \quad f_2(s) = \varphi(\sqrt{2}s),$$

where $\varphi(s)$ is the function of example 2. We realize that the sum $f(s) = f_1(s) + f_2(s)$, considered for instance in $\{-1, 0\}$, has no translation number $\tau \neq 0$ for any $\varepsilon < 1$. In fact, if $\tau \neq 0$ is an arbitrarily given number, we can obviously choose a number t_0 in such a way that one and only one of the four numbers $t_0, t_0 + \tau, \sqrt{2}t_0, \sqrt{2}(t_0 + \tau)$ is a multiple of 2π (if τ itself is a multiple of 2π , we may, for instance, use $t_0 = \pi\sqrt{2}$; otherwise, we may use either the number $t_0 = 2\pi$ or, if $\sqrt{2}(2\pi + \tau)$ just is a multiple of 2π , the number $t_0 = 4\pi$). Therefore, if we make s approach the boundary point it_0 from the half plane $\sigma < 0$, three of the four functions $f_1(s), f_1(s + i\tau), f_2(s), f_2(s + i\tau)$ will approach definite limits, while by making s tend to it_0 in a suitable way from the half plane $\sigma < 0$ we may arrive at any number on the unity circle as a limit of the fourth function; hence it is excluded (just as in example 2) that the modulus of the difference

$$f(s + i\tau) - f(s) = f_1(s + i\tau) + f_2(s + i\tau) - f_1(s) - f_2(s)$$

remains smaller than a number $\varepsilon < 1$ in the whole half plane $\sigma < 0$.

Now we begin the investigation of the functions, almost periodic and unbounded in $\{\alpha, 0\}$. In order to demonstrate at once that the situation here is essentially different from that of the functions almost periodic and unbounded in $\{\alpha, \infty\}$, it may be emphasized that the sum of a function $b(s)$, bounded and almost periodic in $\{\alpha, 0\}$, and a function $p(s)$, unbounded and periodic in $\{\alpha, 0\}$, is not necessarily almost periodic in $\{\alpha, 0\}$. To this purpose, we need only consider the sum

$$f(s) = p(s) + b(s),$$

where $b(s)$ is the function $\varphi(s)$, bounded and almost periodic—even purely periodic—in $\{\alpha, 0\}$, given in example 2, which has no other translation number for any $\varepsilon < 1$ than the numbers

$2\pi m$, while $p(s)$ is the periodic function $\frac{1}{1 - e^{i\sqrt{2}s}}$ with the period $\sqrt{2}\pi i$ and the poles $\sqrt{2}m\pi i$. For, if τ is a translation number of $f(s)$ in $\{\alpha, 0\}$ (belonging to some ε or other), we must have $\tau = \sqrt{2}m\pi$ and, therefore, $p(s + i\tau) = p(s)$; but none of these numbers τ , except the trivial $\tau = 0$, is a translation number for $b(s)$ belonging to any $\varepsilon < 1$.

It is one of our main problems to decide to what extent a converse theorem would hold, i. e. to what degree we also here have a general splitting theorem analogous to that valid for $\beta = \infty$. As we shall see at the end of the Chapter, this is not always the case, although the theorem holds in an important special case.

Let us begin with the remark that a function $f(s)$, unbounded and almost periodic in $\{\alpha, 0\}$, actually assumes values with arbitrarily large modulus already in a bounded part of the plane, precisely speaking, that there exists a length L such that the function is not bounded in any rectangle of the form $\alpha \leq \sigma < 0$, $t' < t < t' + L$. This is obviously valid for an L chosen thus that, in every interval $t' < t < t' + L$, there exists at least one translation number τ of the function $f(s)$ in $\{\alpha, 0\}$ belonging for instance to $\varepsilon = 1$; for, if $f(s)$ was bounded (say $|f(s)| < K$) in only one of these rectangles, it would be bounded (viz. $|f(s)| < K + 1$) in the whole strip $\alpha \leq \sigma < 0$. Hence, denoting a boundary point $s_0 = it_0$ as an "infinity point" of $f(s)$, if $f(s)$ is unbounded in every half circle $\sigma < 0$, $|s - s_0| < \delta$, we conclude, that $f(s)$ certainly has infinity points on the boundary $\sigma = 0$ (as $f(s)$ is bounded in every substrip $\alpha \leq \sigma \leq \gamma < 0$), and that these infinity points form a relatively dense set. That the existence of one infinity point s_0 involves the existence of a relatively dense set of infinity points, follows also from the fact that, together with s_0 , in any case all points $s_0 + i\tau$ must be infinity points, where τ runs through the (relatively dense) translation module F of the function. The set of these infinity points of $f(s)$ on the boundary $\sigma = 0$ forms of course a closed set. This set can very well consist of all the points of the boundary; this is the case for the function $f(s) = \sum e^{n_1 s}$, purely periodic in $(-\infty, 0)$, where it holds for any fixed $t = 2\pi r$ (r rational) that $|f(\sigma + it)| \rightarrow \infty$ for $\sigma \rightarrow 0$.

For the more detailed study of a function, unbounded and

almost periodic in $\{\alpha, 0\}$, an investigation of the translation module F is of special importance. In the case $\beta = \infty$ considered in the foregoing chapter, the translation module was always discrete, i. e. it consisted of the numbers of an arithmetic progression. This may of course also happen when β is finite, for instance for the function $f(s) = \sum_0^{\infty} e^{ns} = \frac{1}{1-e^s}$ (purely periodic in $(-\infty, 0)$) with the poles $2m\pi i$, where F just consists of the numbers $2m\pi$. However, it may here also occur, that F is everywhere dense on the line $-\infty < t < \infty$. This is, for instance, the case for the above mentioned function $f(s) = \sum_1^{\infty} e^{n!s}$ (also purely periodic in $(-\infty, 0)$), since any rational multiple of 2π surely belongs to F , as for $\tau = \frac{m}{q} 2\pi$ in the whole half plane $\sigma < 0$ the inequality

$$|f(s+i\tau) - f(s)| = \sum_1^q \left| e^{n!(s+i\tau)} - e^{n!s} \right| \leq 2q$$

holds. In the case where the translation module F is everywhere dense, the infinity points must also be everywhere dense, as follows from a previous remark; therefore they must constitute the whole boundary (as the set is closed), and especially the boundary $\sigma = 0$ must always be an essentially singular line for the function.

The investigation of the functions, unbounded and almost periodic in $\{\alpha, 0\}$, naturally falls into two cases corresponding to a discrete translation module and to an everywhere dense translation module, respectively. Before starting this investigation, it will first be proved that the translation module cannot consist of all real numbers. We observe that this is not a special case of a general theorem concerning arbitrary unbounded analytic functions, but a typical theorem for unbounded almost periodic functions; thus, the trivial not almost periodic analytic function $f(s) = s$ is unbounded in an arbitrary strip (α, β) , while the difference $f(s+i\tau) - f(s)$ is bounded (in fact constant $= i\tau$) for every fixed τ .

Theorem. *The translation module of a function $f(s)$, unbounded and almost periodic in $\{\alpha, 0\}$, cannot consist of all real numbers.*

It has to be proved that a function $f(s)$, almost periodic in $\{\alpha, 0\}$, which has every number τ as translation number (for some or other ϵ), must necessarily be bounded.

A decisive step in the proof is the demonstration that the translation function

$$v(\tau) = \text{u. b. } |f(s + i\tau) - f(s)|, \\ s \text{ in } \{\alpha, 0\}$$

here defined for all τ , is a bounded function of τ . To show this, it is obviously sufficient to prove that $v(\tau)$ is bounded in the interval $0 < \tau < L$, where L is chosen in such a way that any interval of the length L contains a number τ' , which is a translation number of $f(s)$ in $\{\alpha, 0\}$ corresponding to $\epsilon = 1$; for if $v(\tau) < k$ in $0 < \tau < L$, due to the inequality $v(\tau_1 + \tau_2) \leq v(\tau_1) + v(\tau_2)$, the function $v(\tau)$ is obviously $< k + 1$ for all τ . That the translation function $v(\tau)$ is bounded in a finite interval $0 < \tau < L$ can be demonstrated in the following way: We consider $v(\tau)$ in the interval $-L < \tau < L$ and denote by E_n the (measurable) set in $-L < \tau < L$ in whose points $v(\tau) < n$. As $E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots$ and as any τ in $-L < \tau < L$ belongs to the set E_n for n sufficiently large, the measure $m(E_n)$ of the set E_n will converge to the length $2L$ of the whole interval, for $n \rightarrow \infty$. Therefore, we may determine N such that

$$m(E_N) > \frac{3}{2}L.$$

Then, for any τ in $0 < \tau < L$ the inequality

$$v(\tau) < 2N$$

will be valid. This is evidently proved, if we have shown that any given τ in $0 < \tau < L$ can be written in the form $\tau = \tau_1 - \tau_2$, where both τ_1 and τ_2 belong to the set E_N (and τ_1 lies in the interval $0 < \tau < L$). This is possible, as a consequence of the fact that the set E_N and the set $E'_N = E'_N(\tau)$ obtained from E_N by translating it by τ must have a common point between 0 and L (which point then is equal to τ_1 as well as to $\tau + \tau_2$, where τ_1 and τ_2 both belong to E_N), the set E_N as well as the set E'_N having an intersection with the interval $0 < \tau < L$ the measure of which is greater than $\frac{L}{2}$.

Now it is easy to finish the proof, viz. to show that $f(s)$ must be bounded in $\{\alpha, 0\}$. We have only to choose an arbitrary number \mathcal{A} which is incommensurable with all the Dirichlet exponents of $f(s)$ different from zero (forming a countably infinite set) and to consider the periodic component $p(s)$ of $f(s)$ in $\{\alpha, 0\}$ belonging to the period $ip = \frac{2\pi}{\mathcal{A}}i$. As no Dirichlet exponent of $f(s)$ which is different from zero is a multiple of \mathcal{A} , this periodic component is simply a constant c , viz. the constant term in the Dirichlet development of $f(s)$. On the other hand, however, $p(s)$ can be determined (in the whole strip $\alpha \leq \sigma < 0$) by

$$p(s) = \lim_{n \rightarrow \infty} \frac{f(s+ip) + f(s+2ip) + \cdots + f(s+nip)}{n};$$

hence, it results that the difference $p(s) - f(s)$ must be bounded in the whole strip $\{\alpha, 0\}$, viz. that its modulus must be $\leq G = \text{u. b. } v(\tau)$, as we have

$$p(s) - f(s) = \lim_{n \rightarrow \infty} \frac{(f(s+ip) - f(s)) + (f(s+2ip) - f(s)) + \cdots + (f(s+nip) - f(s))}{n},$$

where the modulus of each of the differences $f(s+ivp) - f(s)$ is $\leq G$ in $\{\alpha, 0\}$. From $p(s) = c$ and $|p(s) - f(s)| \leq G$ in $\{\alpha, 0\}$, it finally results that

$$|f(s)| \leq |c| + G \quad \text{in } \{\alpha, 0\}$$

and so we have proved that $f(s)$ is bounded.

In the following we shall first treat the case where the translation module F is discrete, and then the case where F is everywhere dense on the line $-\infty < t < \infty$.

The translation module is discrete.

In this case the situation proves to be highly analogous to that for $\beta = \infty$ (where the translation module always is discrete), as the following splitting theorem is valid.

Splitting theorem: *A function $f(s)$, unbounded and almost periodic in $\{\alpha, 0\}$, with a discrete translation module can always—and practically in one way only—be written as a sum*

$$f(s) = p(s) + b(s),$$

where $p(s)$ is a function, unbounded and purely periodic in $\{\alpha, 0\}$, while $b(s)$ is a function, bounded and almost periodic in $\{\alpha, 0\}$.

Let us denote the numbers in the translation module F which, in consequence of the assumption, form an arithmetic progression, by $\tau = \nu q$ ($q > 0$, $\nu = 0, \pm 1, \dots$). It is obvious that, in any splitting of $f(s)$ of the kind mentioned in the theorem, each period of the periodic term must necessarily have the form $i\nu q$. It will be proved that, as a period ip of the periodic term, we may even use the number iq , where q is the smallest positive number which might be taken into consideration, for we shall prove (in analogy to the case $\beta = \infty$) that the periodic component of $f(s)$ in $\{\alpha, 0\}$ belonging to the period iq is a possible $p(s)$. When proving this, it would not be convenient (as in the case $\beta = \infty$) to use the Dirichlet developments, because here (in contrast to the case $\beta = \infty$) we have no simple criterion, whether a Dirichlet development just represents a function bounded in $\{\alpha, 0\}$. We have to operate with $p(s)$, determined in $\{\alpha, 0\}$ as a mean value, i. e. by the limit equation

$$p(s) = \lim_{n \rightarrow \infty} \frac{f(s + iq) + f(s + 2iq) + \dots + f(s + niq)}{n}.$$

We have to show that the function $b(s)$ —obtained by subtracting from $f(s)$ this function $p(s)$ of the period iq , purely periodic in $\{\alpha, 0\}$ —is not only (of course) almost periodic in $\{\alpha, 0\}$, but actually almost periodic in $\{\alpha, 0\}$ and moreover is bounded in $\{\alpha, 0\}$.

However, it is plain that the difference $b(s) = f(s) - p(s)$ is almost periodic in the whole strip $\{\alpha, 0\}$. For, as any translation number τ of $f(s)$ in $\{\alpha, 0\}$ lies in the translation module F , i. e. has the form νq , the number $i\tau$ is a period of $p(s)$, and therefore any translation number $\tau(\epsilon)$ of $f(s)$ in $\{\alpha, 0\}$ is also a translation number $\tau(\epsilon)$ of $b(s)$ in $\{\alpha, 0\}$.

In order to prove that the function $b(s) = f(s) - p(s)$ is bounded in $\{\alpha, 0\}$ we show primarily that the translation function $v(\tau)$ of $f(s)$ in $\{\alpha, 0\}$, which here is defined for $\tau = \nu q$ ($\nu = 0, \pm 1, \dots$) only, is bounded, i. e. that

$$v(\tau) \leq K \quad \text{for all } \tau = \nu q \quad (\nu = 0, \pm 1, \dots).$$

To this purpose we consider the set of translation numbers τ of $f(s)$ belonging, for instance, to $\varepsilon = 1$. As each of these numbers is a multiple of q , the relative density of the set formed by these numbers $\tau(1)$ implies that there exists a positive integer M such that among M arbitrary consecutive multiples of q there exists at least one which is a $\tau(1)$. If now k denotes the greatest of the M numbers $v(\nu q)$ ($\nu = 1, 2, \dots, M$), for all $\tau = \nu q$ the inequality $v(\tau) \leq k + 1$ obviously holds. Having thus proved the inequality $v(\tau) \leq K$ for all $\tau = \nu q$, it is plain that the function $b(s)$ is bounded in the whole strip $\{\alpha, 0\}$, since for any point s in $\{\alpha, 0\}$ the limit equation

$$b(s) = f(s) - p(s) = \lim_{n \rightarrow \infty} \frac{(f(s) - f(s + iq)) + (f(s) - f(s + 2iq)) + \dots + (f(s) - f(s + inq))}{n}$$

is valid, where the modulus of each occurring difference

$$f(s) - f(s + i\nu q)$$

is $\leq K$; hence also $|b(s)| \leq K$.

Subsequently, it is easy to decide to what degree the mentioned splitting is unique and, as we shall see, the result is quite analogous to that found for $\beta = \infty$. Let

$$f(s) = p(s) + b(s)$$

be the "standard splitting" given in the proof above, in which $p(s)$ is the periodic component of $f(s)$ belonging to the period iq , where q is the smallest positive number in the translation module F , and let

$$f(s) = p^*(s) + b^*(s)$$

be an arbitrary splitting of $f(s)$ in $\{\alpha, 0\}$ of the kind mentioned in the theorem. As the periodic term $p^*(s)$ certainly has a

number of the form imq as a period (m a positive integer), the difference $\pi(s) = p^*(s) - p(s) = b(s) - b^*(s)$ must necessarily be a function, bounded and periodic in $\{\alpha, 0\}$ with a period of the form imq . Conversely, however, it also holds that for any function $\pi(s)$, bounded and periodic in $\{\alpha, 0\}$, with a period of the form imq , we may use as a periodic splitting term the function

$$p^*(s) = p(s) + \pi(s)$$

periodic in $\{\alpha, 0\}$, i. e. the function

$$b^*(s) = b(s) - \pi(s)$$

is not only (of course) bounded in $\{\alpha, 0\}$ and almost periodic in $\{\alpha, 0\}$, but also almost periodic in the whole strip $\{\alpha, 0\}$. Evidently, this is proved, when we have shown that for any $\varepsilon > 0$ the function $b(s)$ has in $\{\alpha, 0\}$ a relatively dense set of translation numbers (not only, as we already know, of the form νq , but also) of the form νmq . To see this, we only need to apply that $b(s)$ has a relatively dense set of translation numbers of the form νq belonging to $\frac{\varepsilon}{m}$, and that these latter translation numbers multiplied by m are translation numbers of $b(s)$ belonging to ε itself.

The translation module F is everywhere dense.

We have already mentioned that there exist functions, $f(s)$, unbounded and almost periodic in $\{\alpha, 0\}$, whose translation module is everywhere dense on the line $-\infty < t < \infty$; the periodic function $f(s) = \sum_1^\infty e^{n!s}$ considered in $\{-1, 0\}$, for instance, is of this type. Furthermore, we have seen that any function $f(s)$ of this type has the boundary $\sigma = 0$ as an essentially singular line, and even that every point on the boundary is an infinity point. Our main result concerning these functions is comprised in the following (negative) theorem:

Theorem. *There exist functions $f(s)$, almost periodic and unbounded in a strip $\{\alpha, 0\}$, which cannot be splitted into a sum*

$$f(s) = p(s) + b(s),$$

where $p(s)$ is purely periodic in $\{\alpha, 0\}$, while $b(s)$ is bounded and almost periodic in $\{\alpha, 0\}$.

Moreover, we shall prove the somewhat further going theorem that such a splitting is not always possible, even if we only demand that the function $b(s)$ (which on account of the equality $b(s) = f(s) - p(s)$ automatically is almost periodic in $\{\alpha, 0\}$) be bounded in $\{\alpha, 0\}$, but not that it be almost periodic in the whole strip $\{\alpha, 0\}$.

In order to construct a "counter-example" which is suited to prove the correctness of the assertion made in the theorem, we shall first look for a general type of examples concerning functions unbounded and almost periodic in a strip $\{\alpha, 0\}$ and having an everywhere dense translation module. Starting from the simple example $f(s) = \sum_1^{\infty} e^{n!s}$ (periodic and therefore quite unadapted to our proper purpose), it is obvious to think of almost periodic functions with Dirichlet exponents which form a sequence strongly increasing to the infinity, i. e. ordinary Dirichlet series

$$f(s) = \sum_1^{\infty} a_n e^{\lambda_n s} \quad (0 < \lambda_1 < \lambda_2 < \dots \rightarrow \infty)$$

with so-called "gaps", i. e. with very large intervals between the exponents. Actually, the general theorem holds that every such series represents a function $f(s)$, unbounded and almost periodic in $(-\infty, 0)$ (and not only in $(-\infty, 0]$), with an everywhere dense translation module, if the series has the half plane $\sigma < 0$ as convergence half plane and is divergent on the boundary line $\sigma = 0$. We shall postpone the formulation of this "gap theorem" and its proof to the next Chapter. Here, we shall confine ourselves—as this is sufficient for our present purpose—to mentioning that it results from this gap theorem that every Dirichlet series $\sum e^{\lambda_n s}$ with coefficients $a_n = 1$ for which

$$0 < \lambda_1 < \lambda_2 < \dots \rightarrow \infty \quad \text{and} \quad \lambda_{n+1} > e^{\lambda_n}$$

represents a function $f(s)$, unbounded and almost periodic in $(-\infty, 0)$, and, therefore, also for instance in $\{-1, 0\}$, with an everywhere dense translation module.

Within this class of functions $f(s)$ we shall attempt to determine one which may be called highly "aperiodic" in the sense that by splitting off a function $p(s)$ purely periodic in $\{-1, 0\}$ we can never obtain a function $b(s)$ bounded in $\{-1, 0\}$.

As we shall see, we have a function of this type in every function $f(s)$ of our class with rationally independent exponents λ_n . The task is to show that such a function $f(s)$ cannot be written in the form

$$f(s) = p(s) + b(s),$$

where $p(s)$ is purely periodic in $\{-1, 0\}$ and $b(s)$ bounded in $\{-1, 0\}$ (and of course almost periodic in $\{-1, 0\}$). We give an indirect proof and, consequently, suppose that such a representation exists. As the Dirichlet exponents of the function $f(s)$ are rationally independent, at the most one of them can have the form $\frac{2\pi}{p}\nu$, where ip is a period of $p(s)$. In the following we may assume that $f(s)$ has no Dirichlet exponents of this form, as, if $A = \frac{2\pi}{p}\nu$ was such a Dirichlet exponent, we should only subtract e^{As} on both sides of the equation, exactly speaking we should replace $f(s)$ by $f(s) - e^{As}$, and $b(s)$ by $b(s) - e^{As}$, whereby $f(s) - e^{As}$ as $f(s)$ is unbounded in $\{-1, 0\}$, and $b(s) - e^{As}$ as $b(s)$ is bounded in $\{-1, 0\}$. We now write the equation $f(s) = p(s) + b(s)$ in the form

$$b(s) = -p(s) + f(s)$$

and we consider this equation in the strip $\{-1, -\epsilon\}$ where $0 < \epsilon < 1$. The periodic term $-p(s)$ is obviously just the periodic component of $b(s)$ in $\{-1, -\epsilon\}$ belonging to the period ip , because $f(s)$ has no Dirichlet exponents of the form $\frac{2\pi}{p}\nu$. Hence, we have the inequality

$$\text{u. b. } |p(s)| \underset{s \text{ in } \{-1, -\epsilon\}}{\leq} \text{u. b. } |b(s)|.$$

Consequently, if B denotes upper bound of $|b(s)|$ in the whole strip $\{-1, 0\}$, the absolute value of $p(s)$ is $\leq B$ in the whole strip $\{-1, 0\}$. However, this contradicts the fact that $f(s)$ is unbounded in $\{-1, 0\}$, as it would involve that in the whole strip $\{-1, 0\}$ the inequality

$$|f(s)| \leq |p(s)| + |b(s)| \leq 2B$$

was valid. Hereby, the proof of our theorem is fulfilled.

CHAPTER IV.

A gap theorem concerning the almost periodicity of Dirichlet series.

In this Chapter, we shall only deal with Dirichlet series in the classical sense, i. e. with series

$$\sum_1^{\infty} a_n e^{\lambda_n s}, \text{ where } 0 < \lambda_1 < \lambda_2 < \dots \rightarrow \infty.$$

We shall even consider such series only, whose exponents increase strongly to the infinite, from which it follows in particular that the series is absolutely convergent in the whole convergence half plane of the series, which may be supposed to be the half plane $\sigma < 0$. Then, for $\sigma < 0$, the series represents an analytic function $f(s)$ which is almost periodic in $(-\infty, 0]$ and has the given series as its Dirichlet development.

The so-called HADAMARD'S gap theorem for Dirichlet series states, generally speaking, that the convergence line $\sigma = 0$ always is an essentially singular line of the function $f(s)$ represented by the series, if the sequence of the exponents increases rapidly enough. In order to illustrate the kind of reflections made below by an especially simple case, we shall begin this Chapter by proving HADAMARD'S gap theorem in a rather extreme case, viz. the case where the exponents increase so strongly that the ratio of an exponent and the foregoing one is greater than a constant > 3 . As we shall see, the theorem can then be proved in a particularly simple way.

A special case of Hadamard's gap theorem. *A function $f(s)$ represented by a Dirichlet series $\sum a_n e^{\lambda_n s}$ with the convergence half plane $\sigma < 0$ whose exponents satisfy the inequality*

$$\frac{\lambda_{n+1}}{\lambda_n} > k > 3 \quad (\text{for } n > n_0)$$

has the convergence line $\sigma = 0$ as an essentially singular line. If $\sum |a_n|$ is divergent, all the points of the line $\sigma = 0$ are infinity points of $f(s)$.

The proof is based on the following well-known (and easily provable) theorem of VIVANTI and LANDAU: If for a Dirichlet series $f(s) = \sum a_n e^{\lambda_n s}$ with the convergence half plane $\sigma = 0$ it is valid that in a point $s_0 = it_0$ of the boundary all the terms $a_n e^{i\lambda_n t_0}$ are positive from a certain step, the point s_0 is always a singular point of the function $f(s)$ —whether the series is convergent or divergent in the point. We shall use this theorem in the well-known, somewhat more comprehensive formulation where the assumption that all the terms $a_n e^{i\lambda_n t_0}$ are positive from a certain step is replaced by the weaker assumption that all the terms $a_n e^{i\lambda_n t_0}$ from a certain step lie in a fixed angle $< \pi$, for instance in an angle $-\frac{\pi}{2} + d < \nu < \frac{\pi}{2} - d$ (where $0 < d < \frac{\pi}{2}$). Moreover, we shall use the following simple remark: if on the line $-\infty < t < \infty$ there lie intervals of a fixed length $\beta < \alpha$, periodically with a period $\alpha > 0$, then every interval with a length $> \alpha + \beta$ in its interior contains at least one of the mentioned intervals of the length β .

In order to prove that all points of the line $\sigma = 0$ are singular points of the function $f(s)$ it is, of course, sufficient to prove that the singular points lie everywhere dense on the line, that is to say that there exists a singular point in every interval $t_1 < t < t_2$ on the line $\sigma = 0$. In consequence of the theorem mentioned above, this is certainly the case, if in the arbitrarily given interval $t_1 < t < t_2$ there exists a point t_0 such that all the terms $a_n e^{i\lambda_n t_0}$ from a certain step lie in the fixed angle $-\frac{\pi}{2} + d < \nu < \frac{\pi}{2} - d$, where we assume d chosen so small (which is possible because of the inequality $\frac{\lambda_{n+1}}{\lambda_n} > k > 3$ for $n > n_0$) that

$$\frac{\pi - 2d}{\lambda_n} > \frac{3\pi - 2d}{\lambda_{n+1}} \quad \text{for } n > n_0.$$

Concerning the n -th term $a_n e^{i\lambda_n t}$, the intervals I_n on the t -axis, in whose points the amplitude of the term lies in the angle $-\frac{\pi}{2} + d < v < \frac{\pi}{2} - d$, have the length $\frac{\pi - 2d}{\lambda_n}$, and they repeat themselves periodically with the period $\frac{2\pi}{\lambda_n}$. In consequence of a remark above, every interval of a length greater than

$$\frac{2\pi}{\lambda_n} + \frac{\pi - 2d}{\lambda_n} = \frac{3\pi - 2d}{\lambda_n}$$

therefore certainly contains one of the mentioned intervals I_n . By virtue of the inequality written above, for $n > n_0$ each interval I_n contains an interval I_{n+1} . Now, we can immediately complete the proof. We have only to choose $N > n_0$ so great that the given interval $t_1 < t < t_2$ contains an interval I_N . Inside this interval we have then to determine an interval I_{N+1} , inside that again an interval I_{N+2} etc. If t_0 denotes the common point of the sequence of intervals thus determined, all the terms $a_n e^{i\lambda_n t_0}$ for $n \geq N$ are situated in the angle $-\frac{\pi}{2} + d < v < \frac{\pi}{2} - d$, and the point $s_0 = it_0$ is therefore a singular point of $f(s)$.

In the case where $\sum |a_n|$ is divergent, it is moreover clear that the point $s = it_0$ thus obtained is an infinity point of $f(s)$ (as the above consideration shows that $|f(\sigma + it_0)| \rightarrow \infty$ for $\sigma \rightarrow 0$); thus, if $\sum |a_n|$ is divergent, in every interval $t_1 < t < t_2$ there exist infinity points of $f(s)$, i. e. the boundary consists of nothing but infinity points.

We shall now formulate and prove the main theorem of this chapter.

An almost periodic gap theorem. *If $\lambda_1 < \lambda_2 < \dots$ is a sequence of positive numbers, which (for the sake of simplicity) we shall suppose to be > 1 , and which are increasing so strongly to the infinite that*

$$\lambda_{n+1} > e^{k\lambda_n} \quad \text{for all } n,$$

where k is a positive constant, then every Dirichlet series

$$\sum_{n=1}^{\infty} a_n e^{\lambda_n s}$$

belonging to this sequence of exponents and convergent for $\sigma < 0$ —and therefore also absolutely convergent for $\sigma < 0$ —represents a function $f(s)$ analytic in $\sigma < 0$ which is almost periodic in the whole half plane $(-\infty, 0)$ and not only in $(-\infty, 0]$.

Moreover, the translation module F of $f(s)$ in $(-\infty, 0)$ is everywhere dense on the t -axis, i. e. in every interval $t_1 < t < t_2$ exists a number τ such that the difference $f(s + i\tau) - f(s)$ is bounded in the whole half plane $\sigma < 0$.

If $\sum |a_n|$ is convergent, the theorem is trivial, and therefore we may suppose in the proof that $\sum |a_n|$ is divergent; as a consequence of a remark given above it is then valid—as the condition which is now imposed on the exponents λ_n is much stronger than the former one—that the boundary $\sigma = 0$ consists of nothing but infinity points of the function $f(s)$.

Proof: We choose a fixed positive $c < k$. As the series $\sum a_n e^{-\lambda_n c}$ is convergent, there exists a constant K such that $|a_n| e^{-\lambda_n c} < K$ for all n ; we may suppose $K = 1$ (otherwise we only divide all the coefficients a_n by K), i. e.

$$|a_n| < e^{\lambda_n c} \text{ for all } n.$$

Next, we choose a number c' such that

$$c < c' < k$$

and write

$$\varepsilon_n = e^{-\lambda_n c'} \text{ for all } n.$$

Especially it is valid that the series $\sum \varepsilon_n e^{\lambda_n c}$ and thereby a fortiori the series $\sum \varepsilon_n |a_n|$ is convergent.

Let us consider the exponential factor $e^{i\lambda_n t}$. Periodically with a period $\frac{2\pi}{\lambda_n}$, on the t -axis there lie intervals I_n of the length $\frac{2\varepsilon_n}{\lambda_n}$ in whose points t the exponent $\lambda_n t$ differs from an integral multiple of 2π less than ε_n ; in these points it is furthermore valid that

$$|e^{i\lambda_n t} - 1| < \varepsilon_n.$$

For large n , say for $n > n_0$, every one of our intervals I_n contains one of our intervals I_{n+1} , as for large n the inequality

$$\frac{2 \varepsilon_n}{\lambda_n} > \frac{2 \pi}{\lambda_{n+1}} + \frac{2 \varepsilon_{n+1}}{\lambda_{n+1}}$$

is valid; this is evident, as the right side is $< \frac{4 \pi}{\lambda_{n+1}}$ and as for large n (because $c' < k$) we have

$$\lambda_{n+1} > \frac{2 \pi}{\varepsilon_n} \lambda_n = 2 \pi \lambda_n e^{\lambda_n c'}.$$

From this, one of the assertions made in the theorem follows at once, viz. that in every given interval $t_1 < t < t_2$ there exists a number τ for which the difference $f(s + i\tau) - f(s)$ is bounded in the whole half plane $\sigma < 0$. We have only to determine $N > n_0$ so large that the interval $t_1 < t < t_2$ contains one of our intervals I_N ; then, we determine in it an interval I_{N+1} ; in it again an interval I_{N+2} etc. The common point of the sequence of intervals $I_N, I_{N+1}, I_{N+2}, \dots$ is then a point in $t_1 < t < t_2$ for which $f(s + i\tau) - f(s)$ is bounded in $\sigma < 0$, as for every s in $\sigma > 0$ the inequality holds

$$\begin{aligned} |f(s + i\tau) - f(s)| &\leq 2 \sum_{n=1}^{N-1} |a_n| + \sum_{n=N}^{\infty} |a_n| \cdot |e^{i\lambda_n \tau} - 1| \\ &\leq 2 \sum_{n=1}^{N-1} |a_n| + \sum_{n=N}^{\infty} |a_n| \varepsilon_n < \infty. \end{aligned}$$

However, we must proceed somewhat more cautiously when proving our main assertion, viz. that $f(s)$ is almost periodic in $(-\infty, 0)$, i. e. that to any arbitrarily given ε , which we may suppose to be < 1 , there exists a length $L = L(\varepsilon)$ such that, in every interval of the length L , there exists a number τ for which

$$|f(s + i\tau) - f(s)| \leq \varepsilon \text{ in the whole half plane } \sigma < 0.$$

We determine a number $N = N(\varepsilon) > n_0$ such that

$$\sum_{N+1}^{\infty} |a_n| \varepsilon_n < \frac{\varepsilon}{2} \quad \text{and} \quad \lambda_{N+1} > \frac{16 \pi}{\varepsilon} N \lambda_N e^{\lambda_N c}.$$

We can fulfill the last condition, because $c < k$ and $\lambda_{n+1} > e^k \lambda_n$. Now we split the function $f(s)$ in the half plane $\sigma < 0$ into a beginning $B_N(s)$ and a remainder $R_N(s)$, namely

$$B_N(s) = \sum_1^N a_n e^{\lambda_n s} \quad \text{and} \quad R_N(s) = \sum_{N+1}^{\infty} a_n e^{\lambda_n s}.$$

Here, the function $B_N(s)$ is of course almost periodic in $(-\infty, 0)$, namely even almost periodic in $(-\infty, \infty]$. We determine a length $L (> 1)$ such that every interval of the length L contains a translation number τ of $B_N(s)$ in $(-\infty, 0)$ belonging to $\frac{\varepsilon}{4}$. Then, this length will be a usable length $L(\varepsilon)$ of $f(s)$ in $(-\infty, 0)$. To prove this we primarily estimate the differential coefficient $B'_N(s)$ in $\sigma < 0$ and find

$$|B'_N(s)| = \left| \sum_{n=1}^N a_n \lambda_n e^{\lambda_n s} \right| \leq \sum_{n=1}^N \lambda_n |a_n| \leq N \lambda_N e^{\lambda_N c}.$$

If we set

$$l_N = \frac{\frac{\varepsilon}{4}}{N \lambda_N e^{\lambda_N c}},$$

it holds for every t_0 in the interval $\tau - l_N < t < \tau + l_N$ around one of our translation numbers $\tau = \tau\left(\frac{\varepsilon}{4}\right)$ of $B_N(s)$ in $(-\infty, 0)$, that in $(-\infty, 0)$

$$|B_N(s + it_0) - B_N(s + i\tau)| \leq l_N N \lambda_N e^{\lambda_N c} = \frac{\varepsilon}{4};$$

consequently the interval $\tau - l_N < t < \tau + l_N$ consists of nothing but translation numbers of $B_N(s)$ in $(-\infty, 0)$ belonging to $\frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$. In particular, we find (as $L > 1$ and $l_N < \frac{1}{2}$) that every interval $\gamma < t < \gamma + L$ of the length L contains a whole interval i_N of the length l_N whose points all are translation numbers of $B_N(s)$ in $(-\infty, 0)$ belonging to $\frac{\varepsilon}{2}$. However, we have

$$l_N = \frac{\varepsilon}{4 N \lambda_N e^{\lambda_N c}} > \frac{4\pi}{\lambda_{N+1}},$$

and our interval i_N of the length l_N contains therefore at least one of the intervals I_{N+1} mentioned above. Now we proceed as before; we determine an interval I_{N+2} inside I_{N+1} , in it again an interval I_{N+3} , and so on. The common point of the

sequence of intervals $I_N, I_{N+1}, I_{N+2}, \dots$ is called τ . This number τ lies in the interval $\gamma < t < \gamma + L$, and it is further a translation number of $f(s)$ in $(-\infty, 0)$ belonging to ε . In fact, the inequality $|e^{i\lambda_n t} - 1| < \varepsilon_n$ is valid for t lying in an interval I_n ; hence for $\sigma < 0$ we get

$$\begin{aligned} |f(s+i\tau) - f(s)| &\leq |B_N(s+i\tau) - B_N(s)| + |R_N(s+i\tau) - R_N(s)| \\ &\leq \frac{\varepsilon}{2} + \sum_{N+1}^{\infty} |a_n| |e^{i\lambda_n \tau} - 1| \leq \frac{\varepsilon}{2} + \sum_{N+1}^{\infty} |a_n| \varepsilon_n \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus, the theorem is established.



DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, Nr. 19

ON HOMOGENEOUS
GRAVITATIONAL FIELDS IN THE
GENERAL THEORY OF RELATIVITY
AND THE CLOCK PARADOX

BY

C. MØLLER



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark
Bianco Lunos Bogtrykkeri A/S

1. Introduction and statement of the problem.

When the behaviour of clocks is treated according to the principles of the special theory of relativity, without making due allowance for the principles of the general theory, a well-known paradox can arise, which was already mentioned in EINSTEIN'S original paper¹⁾ and later was discussed in detail by LANGEVIN²⁾, LAUE³⁾, and LORENTZ⁴⁾. With a slight simplification of the usual representation*, the problem may be stated as follows. Consider two identically constructed clocks, C_1 and C_2 , one of which, say C_1 , is permanently situated at rest at a point A on the positive X -axis of a definite Lorentz frame of reference K , while C_2 is moving with constant velocity $-v$ in the direction of the X -axis (see Fig. 1). At the moment of coincidence between C_1 and C_2 , the readings of the two clocks are compared. After having travelled with constant velocity for a long time, C_2 for a short time is attacked by a constant force F which brings it to rest at the origin 0 of K and starts it back to A with reversed velocity v . At the moment of the second encounter, the clocks C_1 and C_2 are compared again. Let Δt_1 and Δt_2 denote the measurements on the two clocks of the time elapsed between the two encounters. Now, assuming that the force F is so large that the time during which C_2 is accelerated is negligible compared with the time of travel at the constant velocity v , we have, according to the special theory of relativity, the formula**

* Usually, the two clocks are assumed to be initially and finally at rest, which necessitates the further introduction of a force at the beginning and at the end of the experiment.

** Throughout this paper, we shall use a time unit which makes the velocity of light equal to unity. The transition to ordinary units is then performed by replacing in our formulae all time variables t , velocities v , accelerations g , and gravitational potentials Φ by ct , $\frac{v}{c}$, $\frac{g}{c^2}$, $\frac{\Phi}{c^2}$, respectively, where c is the velocity of light in ordinary units.

$$\mathcal{A}t_2 = \mathcal{A}t_1 \sqrt{1-v^2} \quad (1)$$

which shows that C_2 will register a smaller number of divisions than C_1 at the end of the indicated experiment.

The paradox in question now arises, if we introduce a frame of reference k moving together with C_2 in such a way that C_2 is permanently situated at the origin of k . Since the motion of C_1 with respect to k then is similar to the motion of C_2 with respect to K , it seems that an observer in k should arrive at the conclusion that $\mathcal{A}t_1$ must be smaller than $\mathcal{A}t_2$ and must be given by the formula

$$\mathcal{A}t_1 = \mathcal{A}t_2 \sqrt{1-v^2} \quad (2)$$

in contradiction to (1). In the papers quoted above, it was pointed out, however, that the equation

$$d\tau = dt \sqrt{1-u^2} \quad (3)$$

connecting the proper time $d\tau$ of a clock moving with the velocity u in a given system of reference with the time dt of this system is valid only if the frame of reference is a system of inertia like K . The application of (3) in K thus leads to the correct formula (1), while the application of (3) in k which leads to formula (2) is not justified, since k is accelerated in the middle of the experiment and, therefore, does not constitute a simple system of inertia during this interval.

In the space-time continuum introduced by MINKOWSKI, the two events marked by the first and second encounters of the clocks are represented by two points connected by the world lines of C_1 and C_2 , of which the first mentioned is a straight line. Since the lengths of these world lines, on account of (3), are proportional to the proper times $\mathcal{A}t_1$ and $\mathcal{A}t_2$ of the two clocks, the statement expressed by (1) may be considered a special case of the general statement that a straight line connecting two points in Minkowski space is of greater length than any other curve (of everywhere time-like character) connecting the two points.

Thus, it was clear that the discussion of the indicated experiment could not lead to any difficulties for the special theory of relativity, since this theory does not make any statement at

all regarding the behaviour of clocks in accelerated systems like k . The paradox arose again, however, in the general theory of relativity, according to which a treatment of the behaviour of C_1 , from the point of view of an observer in k , must be possible. Neglecting the short interval during which k is no system of inertia, we then find again the formula (2) for the time increase of C_1 measured with the time scale of k and, at first sight, it is difficult to understand how it is possible to account for the difference between (2) and (1) by consideration of the short interval in which k is accelerated. The whole question was clarified by EINSTEIN⁵⁾ who pointed out that, during this interval, the distant masses of the universe are accelerated relative to k , and thus temporarily create a gravitational field which influences the time rates of the clocks in such a way that the total time increase of C_1 measured in the time scale of k is again given by (1).

In his paper just quoted, EINSTEIN did not give any explicit calculations, but it is clear beforehand that the result of a calculation must be as stated above. In fact, since Δt_1 and Δt_2 are proportional to the lengths of the world lines of C_1 and C_2 and these lengths, according to the basic assumptions of the general theory of relativity, are independent of the space-time coordinates used in their evaluation, it is obvious that we shall get the same value for $\frac{\Delta t_1}{\Delta t_2}$ whether the calculation is performed in K or in k . Nevertheless, it is instructive to calculate directly the time increase of C_1 during the existence of the gravitational field in k . For small values of v , this has been done by TOLMAN⁶⁾ who assumed that terms in v higher than the second can be neglected. In order to account for the lack of symmetry between the treatment given to the clock C_1 , which was at no time subjected to any force, and that given to the clock C_2 , which was subjected to the force F in the middle of the experiment, TOLMAN introduces a temporary homogeneous gravitational field in the description where C_1 is taken as the moving clock and C_2 as the one which remains at rest. This gravitational field is allowed to act on C_1 and C_2 in such a way as to produce the desired change in velocity of C_1 , while C_2 remains at rest on account of the force F . By means of the well-known formula

for the relative rates of two clocks situated at points of different potential in a weak static gravitational field, TOLMAN then finds for the total increase in time of C_1 and C_2 during the considered experiment the relation

$$\Delta t_1 = \Delta t_2 \left(1 + \frac{1}{2}v^2\right) \quad (4)$$

which, for small v , is in accordance with (1).

Apart from the restriction to the case of small v , this treatment does not seem to us to be complete, since it remains to be shown that the transformation from K to the accelerated system k leads to a system of space-time coordinates in which the components of the metrical tensor are constant in time and are of the form corresponding to the gravitational field explicitly introduced by TOLMAN. In the present paper, we shall investigate this point more closely and without making the assumption of a small velocity v . It is shown that the accelerated frame of reference k may be defined in such a way that the gravitational field in k is static in the sense of the general theory of relativity. The equations by which the space-time coordinates of k are expressed as functions of the coordinates of the system K during the whole experiment are explicitly written down. By means of these equations, the behaviour of the clocks C_1 and C_2 may easily be treated from the alternative standpoints of the observers in the two systems K and k , thus leading to a complete solution of the clock paradox.

2. Uniformly accelerated frames of reference and homogeneous gravitational fields.

In a general discussion of the clock paradox, we need a formula connecting the space-time coordinates X , Y , Z , and T of a Lorentz frame K with the coordinates x , y , z , and t of a "uniformly accelerated" frame of reference k . If the direction of acceleration is chosen as x -axis, the desired transformation must have the form

$$\left. \begin{aligned} x &= f(X, T), & y &= Y, & z &= Z \\ t &= h(X, T), \end{aligned} \right\} \quad (5)$$

f and h being functions of X and T , only.

Taking for f and h the expressions

$$f = X - \frac{1}{2}g T^2, \quad h = T, \tag{6}$$

where g is a constant, (5) represents the ordinary transformation to accelerated axes which, at least for small velocities, might be regarded as a reasonable change of coordinates. A free particle in k has then a constant acceleration $-g$, just like a particle in a constant Newtonian field of gravitation. The gravitational field in k , however, is not static in the sense of the general theory of relativity, since the components of the metrical tensor are varying with t .

In fact, introducing (5) and (6) into the expression

$$ds^2 = dX^2 + dY^2 + dZ^2 - dT^2 \tag{7}$$

for the line element in Minkowski space, we get

$$ds^2 = dx^2 + dy^2 + dz^2 + 2gt dx dt - dt^2(1 - g^2 t^2), \tag{8}$$

i. e. the non-vanishing components of the metrical tensor defined by the general expression*

$$ds^2 = g_{ik} dx^i dx^k, \quad (x^i) = (x, y, z, t) \tag{9}$$

are

$$\left. \begin{aligned} g_{11} = g_{22} = g_{33} = 1, \quad g_{44} = -(1 - g^2 t^2) \\ g_{14} = g_{41} = gt. \end{aligned} \right\} \tag{10}$$

Even the geometry in physical space defined by the three-dimensional line element

$$\left. \begin{aligned} d\sigma^2 &= \sum_{i,k=1}^3 \gamma_{ik} dx^i dx^k \\ \gamma_{ik} &= g_{ik} - \frac{g_{i4} g_{k4}}{g_{44}} \end{aligned} \right\} \tag{11}$$

is seen to vary with t .

From (10) and (11), we get

$$d\sigma^2 = \frac{dx^2}{1 - g^2 t^2} + dy^2 + dz^2 \tag{12}$$

* Here, the usual convention is made regarding the summation over dummy indices from 1 to 4.

in accordance with the fact that the measuring rods in k are subjected to a Lorentz contraction.

The gravitational field in the frame of reference defined by (6) has, therefore, not much resemblance with the gravitational fields assumed in the previous discussions of the clock paradox. Our first task will be, if possible, to choose the functions f and h in (5) in such a way that the gravitational field in k is static. The expression for the element of interval in the new coordinates will then be of the form

$$ds^2 = A \cdot dx^2 + dy^2 + dz^2 - D \cdot dt^2, \quad (13)$$

where A and D are functions of x , only. This expression may be further simplified by taking as coordinate $\int \sqrt{A} dx$ instead of x so that the line element takes the form

$$ds^2 = dx^2 + dy^2 + dz^2 - D \cdot dt^2. \quad (13')$$

If the desired transformation is at all possible, the functions g_{ik} defined by (9) and (13') must satisfy EINSTEIN'S field equations for an empty space

$$G_i^k \equiv R_i^k - \frac{1}{2} \delta_i^k R = 0, \quad (14)$$

where R_i^k is the contracted Riemann-Christoffel tensor, and $R \equiv R_i^i$ is obtained from R_i^k by further contraction. The components of G_i^k have been calculated by DINGLE⁷⁾ for a general line element of the form

$$ds^2 = A(dx^1)^2 + B(dx^2)^2 + C(dx^3)^2 - D(dx^4)^2 \quad (15)$$

with A , B , C , and D being any functions of the coordinates.

Using DINGLE'S formula in the special case of (13'), we get simply

$$G_2^2 = G_3^3 = -\frac{1}{2D} \left[D'' - \frac{(D')^2}{2D} \right] = -\frac{1}{D^{1/2}} (D^{1/2})'',$$

where the accents indicate differentiation with respect to x , and all other components G_i^k vanish identically.

The equations (14) thus reduce to the single equation

$$(D^{1/2})'' = 0 \quad (16)$$

with the general solution

$$D = a(1 + gx)^2 \tag{16'}$$

containing two arbitrary constants, a and g .

By adequate choice of the time variable, the constant a may be made equal to one, giving for the line element (13') the expression

$$ds^2 = dx^2 + dy^2 + dz^2 - (1 + gx)^2 dt^2. \tag{17}$$

The functions g_{ik} , defined by (9) and (17), which were found as solutions of the equations (14), may now by a simple calculation be shown also to satisfy the more strict conditions

$$R^i_{klm} = 0, \tag{18}$$

where R^i_{klm} is the uncontracted Riemann-Christoffel tensor. This means that the geometry in the space-time continuum corresponding to (17) is pseudo-Euclidean and that the line element (17) may be brought into the simple form (7) by a suitable transformation of the type (5). Apart from an arbitrary Lorentz transformation, which does not change the form (7), this transformation is uniquely determined.

Before we write down explicitly this transformation, which inversely gives the transition from an inertial system K to the desired frame of reference k , we note that the gravitational field in k , according to (17), is uniform in that part of the space for which gx is a small quantity. In fact, neglecting all terms of higher order in gx than the first, we have

$$g_{44} = -1 - 2gx, \tag{19}$$

and the Newtonian gravitational potential Φ_w , which, in the case of "weak" fields, is defined by the equation⁸⁾

$$g_{44} = -1 - 2\Phi_w, \tag{20}$$

has therefore the simple form

$$\Phi_w = gx. \tag{21}$$

The line element (17) has, however, well defined physical consequences for large values of gx also, so that the gravitational field defined by (17) is a generalization of the "weak" uniform

field postulated in previous discussions of the clock paradox. The only necessary restriction regarding the values of x is the condition $x > -\frac{1}{g}$.

The geometry of physical space in k is Euclidean, x , y , and z being Cartesian coordinates. The time variable t is the time measured by a standard clock situated at rest at the origin $x = 0$. The increase of time $d\tau$ of a standard clock situated at any other place is given by the formula

$$d\tau = \sqrt{-g_{44}} dt = (1 + gx) dt, \quad (22)$$

$d\tau$ thus being zero in the singular plane $x = -\frac{1}{g}$.

Turning now to the explicit derivation of the transformation connecting the space-time variables of the two systems K and k , we start with the system k and try to find a transformation by which the gravitational field of k is "transformed away". This may be effected by introduction of a frame of reference consisting of material points which are allowed to fall freely in the gravitational field of k . The world line of a free particle is a geodesic given by the equations

$$\frac{d^2 x^i}{d\tau^2} + \Gamma_{kl}^i \frac{dx^k}{d\tau} \frac{dx^l}{d\tau} = 0, \quad (23)$$

where $d\tau = \frac{1}{i} ds$ is the proper time of the particle and Γ_{kl}^i denote the ordinary Christoffel three index symbols. The values of Γ_{kl}^i in the case of (17) may also be taken from DINGLE'S paper⁷⁾, and we get

$$\Gamma_{44}^1 = g(1 + gx), \quad \Gamma_{14}^4 = \Gamma_{41}^4 = \frac{g}{1 + gx},$$

all other components being zero.

The equations (23) with $i = 1, 2, 3$ are then simply

$$\left. \begin{aligned} \frac{d^2 x}{d\tau^2} + g(1 + gx) \left(\frac{dt}{d\tau}\right)^2 &= 0 \\ \frac{d^2 y}{d\tau^2} = \frac{d^2 z}{d\tau^2} &= 0 \end{aligned} \right\} \quad (24)$$

and from (17) we get, as a first integral of (23),

$$\left(\frac{dt}{d\tau}\right)^2 = \frac{1 + \left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2}{(1 + gx)^2}. \quad (25)$$

For a particle at rest (or small velocity), (24) reduces to

$$\begin{aligned} \frac{d^2x}{dt^2} &= -g(1 + gx) \\ \frac{d^2y}{dt^2} &= \frac{d^2z}{dt^2} = 0. \end{aligned}$$

If the gravitational potential Φ is defined by the equation

$$\frac{d^2\vec{x}}{dt^2} = -\text{grad } \Phi, \quad \vec{x} = (x, y, z),$$

we thus get

$$\Phi = gx + \frac{1}{2}g^2x^2, \quad (26)$$

an expression which may be regarded as the generalization of (21) to the case of strong fields.

The equation (20) is seen to hold also in this case, since we get from (17) and (26)

$$g_{44} = -(1 + 2\Phi). \quad (27)$$

Finally, (22) may be written

$$d\tau = \sqrt{1 + 2\Phi} dt \quad (28)$$

which, for small Φ , reduces to the well-known formula⁸⁾ for weak fields

$$d\tau = (1 + \Phi_w) dt. \quad (29)$$

Returning now to the general equations (24) and (25), we see that the motion of the particle in the directions of the y - and z -axes is uniform if the proper time τ is used as time scale. We are here only interested in the case where the velocities are zero at $t = \tau = 0$, so that we have the solutions

$$y = y_0, \quad z = z_0, \quad (30)$$

y_0 and z_0 being the initial values of y and z .

From (25) and (24) we then get

$$\frac{dt}{d\tau} = \frac{\sqrt{1 + \left(\frac{dx}{d\tau}\right)^2}}{1 + gx} \quad (31)$$

and

$$\frac{d^2x}{d\tau^2} + \frac{g}{1 + gx} \left(\frac{dx}{d\tau}\right)^2 = -\frac{g}{1 + gx}$$

which may also be written

$$\frac{d^2}{d\tau^2} (1 + gx)^2 = -2g^2. \quad (32)$$

When the initial velocity is zero and x_0 denotes the initial value of x , we get by integration of (32)

$$x = \frac{1}{g} \left\{ \sqrt{(1 + gx_0)^2 - g^2 \tau^2} - 1 \right\}. \quad (33)$$

Introduction of (33) into (31) gives

$$\frac{dt}{d\tau} = \frac{1 + gx_0}{(1 + gx_0)^2 - g^2 \tau^2}$$

which by integration yields

$$t = \int_0^\tau \frac{dt}{d\tau} d\tau = \frac{1}{2g} \ln \frac{1 + gx_0 + g\tau}{1 + gx_0 - g\tau}. \quad (34)$$

From (33) and (34) it follows that a free particle initially at rest at some point in k will move with increasing velocity in the direction of the negative x -axis, later the velocity will decrease and, finally, the particle will come to rest again in the singular plane $x = -\frac{1}{g}$ at the time $t = \infty$ or $\tau = x_0 + \frac{1}{g}$.

We now get the desired transformation, if we put $x_0 = X$, $y_0 = Y$, $z_0 = Z$, and $\tau = T$ in the equations (30), (33), and (34), X , Y , Z , T then being the space-time coordinates of a freely falling frame of reference K which, at the time $t = T = 0$, coincides with the system k . In this way, we get

$$\left. \begin{aligned} x &= \frac{1}{g} \left\{ \sqrt{(1 + gX)^2 - g^2 T^2} - 1 \right\} \\ y &= Y, \quad z = Z \\ t &= \frac{1}{2g} \ln \frac{1 + gX + gT}{1 + gX - gT} \end{aligned} \right\} \quad (35)$$

By a simple calculation, it may be verified that the line element (17) is really brought into the form (7) by the transformation (35), showing that the system K is actually a system of inertia.

Any fixed point in k with constant coordinates x , y , and z is moving relative to K in accordance with the equations

$$\left. \begin{aligned} X &= \frac{1}{g} \left\{ \sqrt{(1 + gx)^2 + g^2 T^2} - 1 \right\} \\ Y &= y, \quad Z = z \end{aligned} \right\} \quad (36)$$

obtained by solving (35) with respect to X , Y , Z .

This motion is, according to the laws of the special theory of relativity, identical with the motion of a particle of rest mass m subjected to a constant force $\frac{mg}{1 + gx}$ in the direction of the X -axis in a system of inertia, *i. e.* (36) represents the "hyperbolic motion"⁹⁾ of a "uniformly accelerated" particle with acceleration

$$\gamma = \frac{g}{1 + gx}. \quad (37)$$

On account of the dependence of γ on x , the distance, measured by an observer in K , between two fixed points in the frame k will not, in general, be constant in time. Since, however, the same distance is constant when measured by a comoving meter stick, the system k deserves the name of a uniformly accelerated rigid frame of reference, and the transformation (35) plays a similar part as does the Lorentz transformation in the case of a rigid frame moving with constant velocity.

Since the variables x and t , defined by (35), must be real, we shall have to confine ourselves to the consideration of events satisfying the condition

$$-(1 + gX) < gT < 1 + gX. \quad (38)$$

For later use, we also write down the Lorentz transformation connecting the space-time coordinates of two systems of inertia with the relative velocity v

$$\left. \begin{aligned} x &= \frac{X - X_0 - v(T - T_0)}{\sqrt{1 - v^2}} \\ y &= Y, \quad z = Z \\ t - t_0 &= \frac{T - T_0 - v(X - X_0)}{\sqrt{1 - v^2}}. \end{aligned} \right\} \quad (39)$$

In (39), the space and time variables have been chosen in such a way that the origin $x = 0$ of the system k at the time $t = t_0$ corresponds to the coordinate $X = X_0$ and the time $T = T_0$ in K .

3. The clock paradox.

a. In the first part of this section, we shall treat the problem from the point of view of an observer in K . While the clock C_1 is permanently situated at rest at the point A on the positive X -axis, C_2 at the beginning is travelling with constant velocity $-v$ in the direction of the X -axis. At the point B , the clock C_2 is subjected to a constant force F , which brings it to rest at the origin O and starts it back to B with reversed velocity. At the time of arrival in B , C_2 will have regained the velocity v which it retains during the travel from B to A . Let us assume for simplicity that the coincidence of C_2 with O takes place at the time $T = 0$ and that the proper time τ of C_2 is also zero at this moment. Since the problem is then completely symmetrical with respect to this event, we only need explicitly to consider the behaviour of C_2 during its travel from O to B and onwards to A .

Let T' and T'' be the times, measured in the time scale of the system K , during which C_2 travels from O to B and from B to A , respectively, and let τ' and τ'' be the corresponding proper times measured by the clock C_2 itself. The motion of C_2 from B to O and back to B will be a hyperbolic motion given by the equation

$$X = \frac{1}{g} \left\{ \sqrt{1 + g^2 T^2} - 1 \right\}, \quad (40)$$

where the constant g is connected with the force F and the rest mass m of C_2 by the relation

$$F = mg. \tag{41}$$

According to (40), the velocity $u = \frac{dX}{dT}$ is given by

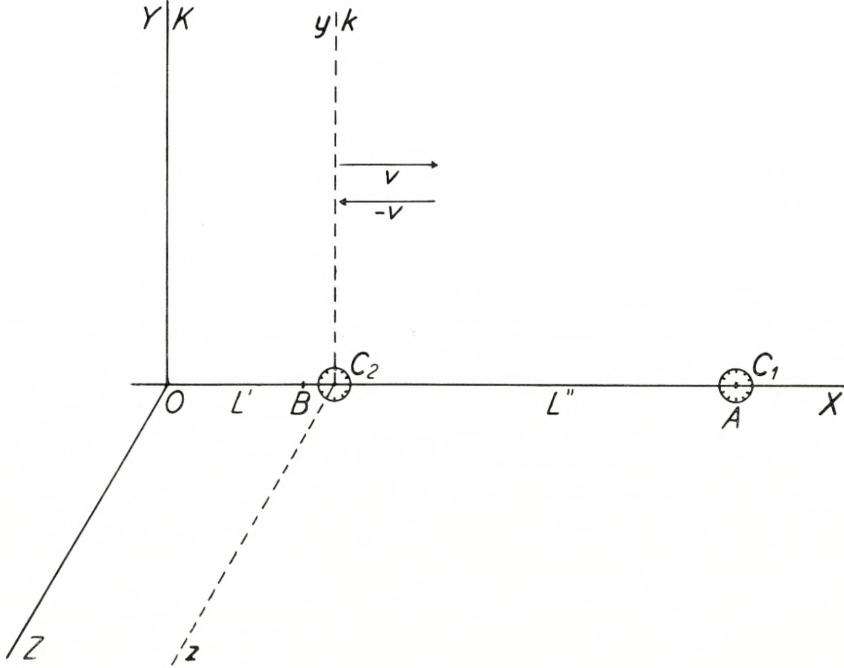


Fig. 1.

$$u = \frac{gT}{\sqrt{1 + g^2 T^2}} \tag{42}$$

and, since $u = v$ for $T = T'$, we have

$$gT' = \frac{v}{\sqrt{1 - v^2}}. \tag{43}$$

Introducing (42) into (3), we get by integration

$$gT' = \sinh g\tau'. \tag{44}$$

The corresponding relation between T'' and τ'' is, according to the well-known formula from the special theory of relativity,

$$T'' = \frac{\tau''}{\sqrt{1-v^2}}. \quad (45)$$

From (43) and (44) we further obtain

$$\left. \begin{aligned} v &= \operatorname{tgh} g\tau' \\ \frac{1}{\sqrt{1-v^2}} &= \operatorname{cosh} g\tau'. \end{aligned} \right\} (46)$$

Now, let $A'_K t_1$ denote the number of divisions registered by C_1 during the travel of C_2 from O to B , as judged by an observer in K , and let $A''_K t_1$ be the corresponding number during the period of uniform motion of C_2 from B to A . We then have

$$A'_K t_1 = T' \quad \text{and} \quad A''_K t_1 = T'' \quad (47)$$

and, for the total time elapsed between the two encounters of C_1 and C_2 , measured by C_1 and C_2 , respectively, we get

$$\left. \begin{aligned} At_1 &= 2(A'_K t_1 + A''_K t_1) = 2(T' + T'') \\ At_2 &= 2(\tau' + \tau''). \end{aligned} \right\} (48)$$

When the applied force F is chosen so large that T' and τ' , given by (43) and (44), become negligible, the connection between At_1 and At_2 , according to (48) and (45), is again given by the simple formula (1).

If L' and L'' denote the distances OB and BA , measured with the measuring rods of the system K , we get from (40) and (43)

$$L' = \frac{1}{g} (\sqrt{1+g^2 T'^2} - 1) = \frac{1}{g} \left(\frac{1}{\sqrt{1-v^2}} - 1 \right) \quad (49)$$

while, obviously,

$$L'' = vT''. \quad (50)$$

We shall now introduce a frame of reference k moving together with C_2 and we may take C_2 as the origin of k . While the motion of the origin is, thus, completely determined, the motion of any other fixed point of k may, beforehand, be chosen

arbitrarily. In the previous discussions of the clock paradox, it has, however, tacitly been assumed that k should be a rigid frame of reference. According to the considerations in Section 2, it is then clear that the transformation connecting the space-time variables of K and k must be given by (35) during the accelerated motion of C_2 from B to O and back. The motion of the origin $x = 0$ relative to K is then, on account of (36), identical with the motion of C_2 given by (40), and the time variable t is simply the proper time of the clock C_2 .

For all events satisfying the conditions $-\tau' < t < \tau'$, the connection between the coordinates of K and k is, thus, given by (35). For $t > \tau'$, the system k is a simple system of inertia, and the corresponding space-time transformation is obtained from (39) by putting

$$t_0 = \tau', \quad T_0 = T', \quad \text{and} \quad X_0 = L'. \tag{51}$$

Similarly, we have for $t < -\tau'$ the transformations (39) with reversed signs of v , τ' and T' .

In the following, we shall use the equations (35) and (39) in a somewhat different form. Solving the last equation (35) with respect to gT and introducing into the first equation, we get, if we omit the trivial transformations of the y and z variables,

$$\left. \begin{aligned} gT &= (1 + gX) \operatorname{tgh} gt \\ 1 + gX &= (1 + gx) \operatorname{cosh} gt \end{aligned} \right\} \tag{52}$$

for

$$-\tau' < t < \tau'.$$

By a similar procedure, we get from (39) and (51) the transformation

$$\left. \begin{aligned} T - T' &= (t - \tau') \sqrt{1 - v^2} + v(X - L') \\ X - L' &= \frac{x + v(t - \tau')}{\sqrt{1 - v^2}} \end{aligned} \right\} \tag{53}$$

for

$$t > \tau'.$$

For $t < -\tau'$, the corresponding transformation (53') is obtained from (53) by reversing the signs of v , T' , and τ' .

In spite of the great difference in form between the equations (52) and (53), they are easily seen to be identical for $t = \tau'$. For this particular value of t , the equations (52) reduce to

$$\begin{aligned} gT &= (1 + gX) \operatorname{tgh} g\tau' \\ 1 + gX &= (1 + gx) \cosh g\tau' \end{aligned}$$

which, by means of (43), (46), and (49) may be written

$$\begin{aligned} T &= v(X - L') + T' \\ X - L' &= \frac{x}{\sqrt{1 - v^2}}, \end{aligned}$$

in accordance with (53) for $t = \tau'$.

On account of the symmetry inherent in our problem, a similar result would be obtained for $t = -\tau'$, so that the correlation of the coordinates x, y, z, t and the physical events is performed in a continuous way by the equations (52), (53), and (53'). Also the velocity of any fixed point in k relative to K varies continuously at $t = \tau'$ (and $-\tau'$). From (36) and (52) we get for constant values of x, y, z

$$\left(\frac{dX}{dT}\right)_x = \frac{gT}{\sqrt{(1 + gx)^2 + g^2 T^2}} = \operatorname{tgh} gt. \quad (54)$$

On the other hand, $\left(\frac{dX}{dT}\right)_x$ is equal to v and $-v$ for $t > \tau'$ and $t < -\tau'$, respectively, which, on account of (46), is seen to be in accordance with (54) for t equal to τ' and $-\tau'$.

While, thus, the velocities of the different points of k vary continuously, it is clear that the accelerations must be discontinuous for $t = \tau'$ and $-\tau'$, since the force F is assumed to set in abruptly. This is also the reason for the sudden change in the gravitational potential from the value zero to the value given by (26) at these moments.

The system k defined by (52), (53), and (53') thus seems to be the most natural frame of reference to be used in the discussion of the clock paradox. The applicability of this system of coordinates is only restricted by the condition that (38) must be satisfied for $-\tau' < t < \tau'$, *i. e.* for

$$-v(1 + gX) < gT < v(1 + gX), \quad (55)$$

on account of (52) and (46). Since v is smaller than one, a comparison of (38) and (55) shows that this condition is satisfied for all events which take place at points $X > -\frac{1}{g}$.

b. We shall now treat the problem from the point of view of an observer in k , according to which C_2 is permanently situated at rest at the origin o of k , while C_1 at the beginning is travelling with constant velocity v . The first encounter between C_1 and C_2 takes place at the time $t = -\tau' - \tau''$. At $t = -\tau'$, C_1 has arrived at a point b on the positive x -axis with the coordi-

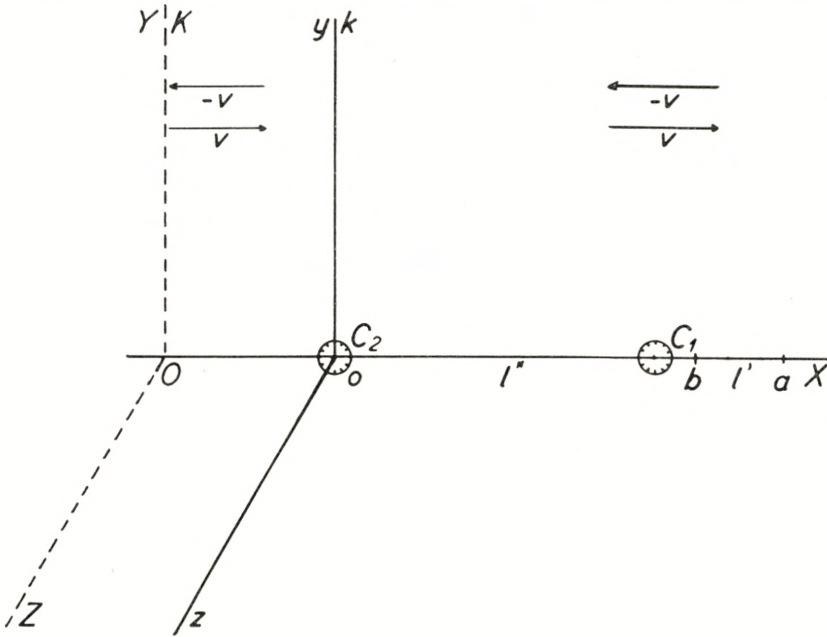


Fig. 2.

nate $x = l''$. During the time $-\tau' < t < \tau'$, C_1 is subjected to the gravitational field which brings it to rest at the time $t = 0$ at a point a in the distance l' from b , and starts it back to o with reversed motion. In spite of this gravitational field everywhere present during this period, C_2 remains at rest on account of the force F which just counterbalances the gravitational force.

The behaviour of the clock C_1 is now simply obtained from (52) and (53) if we remember that the X -coordinate of C_1 has the constant value $X = L' + L''$.

From the second equation (53) we then get

$$l'' = L'' \sqrt{1 - v^2}, \tag{56}$$

since l'' is the value of x for $t = \tau'$.

Further, since the x -value of C_1 at $t = 0$ is $l' + l''$, we get from the second equation (52)

$$l' + l'' = L' + L'' \quad (57)$$

and, therefore,

$$l' = L' + L'' (1 - \sqrt{1 - v^2}). \quad (58)$$

On account of the Lorentz contraction factor in (56), the distance travelled by C_1 with constant velocity v relative to k will, thus, be shorter than the distance which C_2 travels with constant velocity in K . Nevertheless, the total distances travelled by the two clocks along the x -axes will be equal. In the extreme case of $v \rightarrow c$, we have simply $l'' \rightarrow 0$ and $l' \rightarrow L' + L''$.

If $A'_k t_1$ denotes the number of divisions registered by C_1 during the travel from a to b (or from b to a), we get from the first equation (52), by putting $X = L' + L''$, $t = \tau'$, and $T = A'_k t_1$,

$$A'_k t_1 = \frac{1}{g} [1 + g(L' + L'')] \operatorname{tgh} g \tau' = T' + v^2 T'' \quad (59)$$

by means of (44), (46), (49), and (50).

For the corresponding number of divisions $A''_k t_1$ registered by C_1 during the period of travel with constant velocity, we have, according to the special theory of relativity,

$$A''_k t_1 = \tau'' \sqrt{1 - v^2}. \quad (60)$$

This formula is also easily obtained from the first equation (53) if we remember that $A''_k t_1$ is the increase in T for $X = L' + L''$ during the interval $\tau' < t < \tau' + \tau''$. On account of (45), we may also write

$$A''_k t_1 = T'' (1 - v^2). \quad (61)$$

Although, thus, $A''_k t_1$ is smaller than $A'_k t_1$ in (47), it follows from (59), and (61) that the total time elapsed between the two encounters of C_1 and C_2 measured by C_1 and C_2 , respectively, is again given by

$$\left. \begin{aligned} A t_1 &= 2(A'_k t_1 + A''_k t_1) = 2(T' + T'') \\ A t_2 &= 2(\tau' + \tau'') \end{aligned} \right\} (48')$$

in accordance with the expressions (48) derived from the standpoint of an observer in K .

It is interesting to note that $\mathcal{A}'_k t_1$ remains finite in the limiting case of very large forces F , where τ' , Γ' , and $\mathcal{A}'_K t_1$ vanish, since $\mathcal{A}'_k t_1$ in (59) contains a term which only depends upon v and Γ'' . It is just this term which is essential for the solution of the clock paradox.

Since $\mathcal{A}t_2$ in any case is smaller than $\mathcal{A}t_1$, and $\mathcal{A}''_k t_1$, according to (60), is smaller than τ'' , $\mathcal{A}'_k t_1$ must be greater than τ' , *i. e.* the clock C_1 goes faster than C_2 during this period. From the point of view of an observer in k , the reason for this difference in rate is to be sought mainly in the difference in gravitational potential Φ at the places of the two clocks. The behaviour of C_1 , however, will in general not be like that of a clock at rest at the point $x = l' + l'' = L' + L''$, even if Γ' and τ' are made small by use of a large force F . In fact, the number of divisions registered by a clock at rest during the time $\mathcal{A}t = \tau'$ is, according to (26), (28), or (22), given by

$$(\mathcal{A}'_k t_1)_0 = [1 + g(L' + L'')] \cdot \tau', \tag{62}$$

a number which is greater than $\mathcal{A}'_k t_1$ in (59), since we have

$$\frac{\text{tgh } g\tau'}{g} < \tau'.$$

From (17) and (26), we get the expression

$$\left. \begin{aligned} d\tau &= dt \sqrt{(1 + gx)^2 - \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2 - \left(\frac{dz}{dt}\right)^2} \\ &= dt \sqrt{1 + 2\Phi - u^2} \end{aligned} \right\} \tag{63}$$

for the proper time of a particle moving with velocity u in the gravitational potential Φ . This general formula, which comprises the special formulae (3) and (28), clearly shows that $\mathcal{A}'_k t_1$ in general must be smaller than $(\mathcal{A}'_k t_1)_0$, since C_1 during the time in question falls freely with increasing velocity from the place $x = L' + L''$ towards smaller values of x , *i. e.* smaller values of the potential Φ .

Only in the case $v \ll 1$ considered by TOLMAN, where $\text{tgh } g\tau'$ is equal to $g\tau'$, apart from terms of the third order in v (cf.

(46)), it is allowed to treat C_1 as a clock at rest during the period of acceleration, since the difference between $\mathcal{A}'_k t_1$ and $(\mathcal{A}'_k t_1)_0$ is then of higher order in v . Even in this case, where gt may be treated as a small quantity, the equations (52), however, do not reduce to the transformations given by (6). If we neglect terms of higher order in gt , we obtain instead

$$T = (1 + gX) t = (1 + gx) t$$

$$X = x + \frac{1}{2} gt^2 (1 + gx).$$

To get the transformation (6), we should, thus, have to replace the factor $1 + gx$ by 1 and this would mean neglect of just those terms which, in the preceding discussion, have been seen to be essential for the treatment of the clock paradox.

4. Rigid frames of reference in arbitrary motion.

In Section 2, it was shown that the transformation (35) is essentially determined by the condition that the gravitational field of the accelerated system k should be static, and the line element and the gravitational potential in the transformed system are given by (17) and (26), respectively. Since the motion of the origin of k in this case is a hyperbolic motion, the applicability of the transformation (35) in the preceding discussion is confined to the case where the clock C_2 is subjected to a constant force during the period of acceleration. For any other motion the gravitational field in the comoving system will not be static. Anyway, it is always possible to choose the time variable t in the transformations (5) in such a way that the line element takes the form (13), where A and D in general are functions of both variables x and t . If we want the system k to be a rigid frame of reference, A must, however, be independent of t , so that the line element may be brought into the simple form (13') by a suitable choice of the variable x . Then, the spacial geometry is again Euclidean, x , y , and z being Cartesian coordinates.

Using DINGLE's general formulae⁷⁾, one finds that EINSTEIN's field equations (14) in this case reduce to the single equation

$$\frac{\partial^2}{\partial x^2} (D^{1/2}) = 0$$

which is obtained from (16) by replacing the ordinary differentiations with respect to x by partial differentiations. The general solution is again of the form (16'), a and g here being arbitrary functions of t . Finally, the time variable t may be chosen such that the line element takes the same simple form (17) as in the special case treated in Section 2, g in the general case being an arbitrary function of t . The equations (18)—(29) and (63) are seen, therefore, to hold also in the general case.

In order to find the transformation (5) by which the expression (7) for interval is transformed into (17) and by which, conversely, the gravitational field in k is transformed away, we may proceed exactly as in Section 2. First, we solve the equations (24) and (25) for the motion of a free particle initially at rest. After that, the proper time τ of the particle and the initial values x_0, y_0, z_0 of the space coordinates in k are identified with the time and space coordinates T, X, Y, Z in K . The solution of the equations (24) and (25) is only somewhat more complicated than in the case of constant g considered in Section 2. Since g may be regarded as a known function of t , it is convenient to use t as parameter in (24) instead of τ . The elimination of τ is easily performed by means of (25) and, finally, applying elementary methods, a complete solution of the problem is possible.

We shall here give the results, only. For the transformations connecting the space-time variables of the systems K and k , we get

$$\left. \begin{aligned} X &= x \cosh \theta + \int_0^t \sinh \theta \, dt, & Y &= y, & Z &= z \\ T &= x \sinh \theta + \int_0^t \cosh \theta \, dt \\ \text{with} & & & & & \\ \Theta(t) &= \int_0^t g(t) \, dt. \end{aligned} \right\} \quad (64)$$

It is easily verified by means of direct calculation that the line element (7) is really brought into the form (17) by the transformation (64). Further, we see that the equations (64) in the case of constant g reduce to the equations (52) which are equi-

valent to (35). On the other hand, if g is assumed to be finite and constant for $-\tau' < t < \tau'$ and zero for all other times, (64) leads to the transformations (52) and (53) used in the discussions of Section 3.

When g is given as a function of t , the transformation (64) and, consequently, also the motion of the origin of k with respect to K is completely determined. Conversely, the function g and the transformation (64) are uniquely determined by the motion of the origin of k . Differentiating (64) by constant x , we get

$$\left. \begin{aligned} dX &= \sinh \theta \cdot (1 + gx) dt \\ dT &= \cosh \theta \cdot (1 + gx) dt. \end{aligned} \right\} (65)$$

The velocity $U = \left(\frac{dX}{dT}\right)_x$ of a fixed point in k with respect to K is, thus,

$$U = \operatorname{tgh} \theta, \quad (66)$$

an equation which may be regarded as a generalization of (54). Moreover, we get from (66), (65), and from the definition of θ in (64)

$$\frac{d}{dT} \left(\frac{U}{\sqrt{1-U^2}} \right) = \frac{g}{1+gx} \quad (67)$$

which shows that the motion of a point $x = \text{constant}$ is the same as that of a particle of mass m attacked by a force $\frac{mg}{1+gx}$, just like in the case of constant g (cf. p. 13).

When the motion of the origin $x = 0$ is given by the equation

$$X = \psi(T), \quad (68)$$

the corresponding g is obtained as a function of t by elimination of the variable T from the equations

$$\left. \begin{aligned} g &= \frac{d}{dT} \left(\frac{\psi'}{\sqrt{1-(\psi')^2}} \right) \\ t &= \int_0^T \sqrt{1-(\psi')^2} dT \end{aligned} \right\} (69)$$

which are easily derived from (67), (68), (66), and (65).

By means of the general equations (64), it is now easy to treat the clock paradox for an arbitrary motion of the clock C_2 during the interval of acceleration. Since, however, the treatment of this general case does not exhibit any essentially new features as compared with the treatment of the special case discussed in Section 3, we shall confine ourselves to the general remarks already made in this section.

References.

- 1) A. EINSTEIN, Ann. d. Phys. **17**, 891 (1905).
 - 2) P. LANGEVIN, L'évolution de l'espace et du temps, Scientia **10**, 31 (1911).
 - 3) M. v. LAUE, Phys. Z. **13**, 118 (1912).
 - 4) H. A. LORENTZ, Das Relativitätsprinzip, 3 Haarlemer Vorlesungen, Leipzig (1914).
 - 5) A. EINSTEIN, Naturwiss. **6**, 697 (1918).
 - 6) R. C. TOLMAN, Relativity, Thermodynamics and Cosmology, Oxford (1934).
 - 7) H. DINGLE, Proc. Nat. Acad. **19**, 559 (1933).
 - 8) A. EINSTEIN, Ann. d. Phys. **49**, 769 (1916).
 - 9) M. BORN, Ann. d. Phys. **30**, 1 (1909).
- A. SOMMERFELD, Ann. d. Phys. **33**, 670 (1910).
-

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 20

ZUR KENNTNIS DER AZOLE

VON

K. A. JENSEN UND A. FRIEDIGER



KØBENHAVN

I KOMMISSION HOS EJNAR MUNKSGAARD

1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S.

1. Einleitung.

Unter Azolen versteht man nach HANTZSCH¹ die stickstoffhaltigen, heterocyclischen Fünfringe mit zwei Heteroatomen (Imidazol, Oxazol, Thiazol, Pyrazol, Isoxazol, und Isothiazol); dieser Name wurde später auch auf die stickstoffhaltigen Fünfringe mit vier Heteroatomen (Triazole, Thiodiazole, Oxdiazole, Tetrazol) übertragen². Diese Azole sind sämtlich durch ihren ausgesprochenen »aromatischen« Charakter ausgezeichnet; vom Benzol und den einfachsten heterocyclischen Fünfringen, Furan, Thiophen und Pyrrol, unterscheiden sie sich jedoch zum Teil erheblich durch ihren mehr »polaren« Charakter, der sich z. B. in hohen Siedepunkten und grosser Wasserlöslichkeit äussert. Auch im chemischen Verhalten weichen sie in mehreren Punkten stark vom Benzol ab, und im übrigen bestehen im Verhalten der einzelnen Azole oft auffallende Unterschiede, die auf klassischer Grundlage ganz unverständlich sind. Ein tieferes Verständnis der Eigentümlichkeiten der Azole versprechen die neueren Vorstellungen über den aromatischen Bindungszustand zu geben. Um diese Vorstellungen auswerten zu können, bedarf es jedoch noch der Kenntnis einer Reihe physikalischer Konstanten der Azole. Ziel der vorliegenden Untersuchung war es, durch systematische Messungen der Dipolmomente von Azolen und Azolderivaten einen Beitrag zum Verständnis der Bindungsverhältnisse in den Azolen zu liefern³.

¹ A. HANTZSCH, Liebigs Ann. d. Chem. **249** (1888), 1.

² RICHTER-ANSCHÜTZ, Chemie der Kohlenstoffverbindungen, 11. Aufl., Bd. II (Bonn 1913), S. 765; 12. Aufl., Bd. III (Leipzig 1931), S. 87.

³ Nach Inangriffnahme dieser Untersuchung wurde uns eine kurz vorher erschienene Abhandlung von W. HÜCKEL, J. DATOW und E. SIMMERSBACH (Zeitschr. phys. Chem. Abt. A **186** (1940), 129) bekannt, in der unter teilweiser Verfolgung anderer Ziele auch die Dipolmomente von Imidazol, 4- und 1-Methylimidazol, Pyrazol und 1,2,4-Triazol bestimmt wurden.

2. Resonanzerscheinungen bei den Azolen.

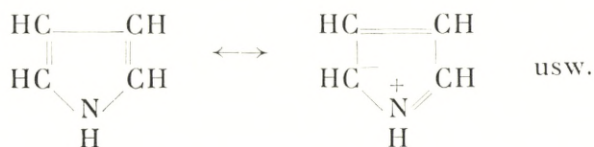
Die Azole gehören zu den sogenannten »benzoiden Systemen«, d. h. sie besitzen »aromatischen« Charakter und sind ebenso wie Benzol relativ gesättigte, sehr stabile Ringsysteme. Die Konstitution des Benzols ist bekanntlich mit Hilfe quantenmechanischer Methoden weitgehend aufgeklärt worden¹. Nach diesen Ergebnissen enthält der Benzolring 6 unlokalisierte Elektronen (π -Elektronen, »Elektronen zweiter Art« nach E. HÜCKEL), die eine abgeschlossene Gruppe von $2 + 4$ Elektronen bilden. Dass bei den Atomen ein Oktett ($2 + 6$ Elektronen), bei den (planen) Ringen dagegen ein Sextett ($2 + 4$ Elektronen) als abgeschlossene Gruppe auftritt, beruht darauf, dass im ersten Fall Kugelsymmetrie vorliegt, im zweiten aber nur eine Symmetrieachse vorhanden ist. Die Ladungsverteilung im Benzol ist völlig symmetrisch, die Bindung zwischen zwei benachbarten Kohlenstoffatomen ist deshalb überall gleich stark. Dies entspricht am ehesten der Benzolformel von THIELE². Nach dem sogenannten »ersten Näherungsverfahren« (PAULING) kann man aber bei der Diskussion der Konstitution des Benzols auch folgendermassen vorgehen: Indem man die 6 π -Elektronen in verschiedener Weise paart, gelangt man u. a. zu den beiden Kekulé-Strukturen und den drei Dewar-Strukturen (durch diese fünf von einander unabhängigen, sogenannten »kanonischen« Strukturen sind alle anderen formulierbaren, ebenen Valenzstrukturen des Benzols gegeben). Diese Valenzstrukturen entsprechen aber nicht je einem bestimmten Molekülzustand, sondern der ebenen Anordnung von 6 CH-Gruppen entspricht nur ein bestimmter Molekülzustand, für welchen keine bestimmte Valenzstruktur angegeben werden kann. Der wahre Zustand liegt zwischen allen denkbaren, formulierbaren »Grenzzuständen«. Die zu dem bestimmten Molekülzustand gehörige Eigenfunktion lässt sich aber durch Überlagerung der zu den fünf kanonischen Strukturen gehörigen (fiktiven) Eigenfunktionen bilden. Diese Überlagerung bezeichnet man nach PAULING als Resonanz oder nach INGOLD, sowie nach E. HÜCKEL, als Mesomerie. Die Konstitution des Benzols

¹ E. HÜCKEL, Grundzüge der Theorie ungesättigter und aromatischer Verbindungen (Berlin 1938). L. PAULING, The Nature of the Chemical Bond (New York 1939). Dasselbst weitere Literatur.

² J. THIELE, Liebigs Ann. d. Chem. **306** (1899), 127.

lässt sich also nicht durch eine bestimmte Valenzformel beschreiben, wohl aber durch eine Gesamtheit mehrerer; man kann auch sagen, im Benzolmolekül bestehe Resonanz zwischen den Kekulé- und den Dewar-Formeln. Der Gesichtspunkt des »ersten Näherungsverfahrens« hat sich für chemische Zwecke als besonders anschaulich und fruchtbar erwiesen. Die rein formale Bedeutung der »Resonanzstrukturen« oder »mesomeren Grenzzustände« sei jedoch ausdrücklich betont¹.

Die heterocyklischen Fünfringe enthalten nun ebenfalls eine Sechs-Elektronengruppe, indem die CH-Gruppen und N-Atome je ein Elektron, die O- und S-Atome und die NH-Gruppe dagegen je zwei π -Elektronen liefern. Der aromatische Charakter dieser Verbindungen erklärt sich somit dadurch, dass das zentrale Elektronensystem dem des Benzols entspricht. Wenn man aber die Konstitution der heterocyklischen Verbindungen vom Standpunkt der Resonanztheorie betrachtet, unterscheiden diese Verbindungen sich vom Benzol dadurch, dass für ihre Konstitution polare Formeln von wesentlicher Bedeutung sind. Für Pyrrol gibt es z. B. nur eine Formel, in der jedes Kohlenstoffatom mit vier Bindungen an Nachbaratome gebunden ist. Um den aromatischen Charakter des Pyrrols zu erklären, müssen jedoch auch andere Strukturen zu dem Zustand des Moleküls beitragen. Diese müssen dann polar sein:



PAULING und SHERMAN² haben versucht, auf Grundlage thermochemischer Daten die Resonanzenergie des Pyrrols und einiger anderer heterocyklischer Verbindungen zu berechnen. Wenn

¹ Es sind von verschiedenen Seiten Einwendungen gegen die Verwendung des Resonanzbegriffs erhoben worden. So schreibt A. J. STAVERMANN (Chemisch Weekblad **39** (1942), 344): »Door invoering van het elastische begrip resonantie bereikt Pauling inderdaad het gestelde doel, maar om het gevaer te bezweren, dat zich in plaats van werkelijk begrip van de zaak en soort jargon ontwikkelt sonder dieper betekenis«. — Es muss zugegeben werden, dass eine derartige Gefahr tatsächlich besteht; aber ein jeder, der die Entwicklung der organischen Chemie in den letzten 15 Jahren aufmerksam verfolgt hat, wird den sehr grossen heuristischen Wert des Resonanzbegriffs nicht bestreiten können.

² L. PAULING u. J. SHERMAN, Journ. Chem. Physics **1** (1933), 606.

auch ihre Berechnungen zu Bedenken Anlass gegeben haben, ist an dem rein qualitativen Ergebnis, dass Resonanzstrukturen vom angegebenen Typus für die Konstitution des Pyrrols von Bedeutung sind, kaum zu zweifeln.

Eine Stütze für die Annahme von polaren Resonanzformeln kann auch aus den Dipolmomenten der heterocyklischen Ringe abgeleitet werden, wie zuerst von SUTTON¹ nachgewiesen wurde. Betrachtet man die Dipolmomente² der heterocyklischen Ringe mit einem Heteroatom (Messungen von DE VRIES ROBLES³):

	μ		μ		μ
Furan	0,67	Dihydrofuran ..	1,53	Tetrahydrofuran	1,68
Thiophen....	0,54			Tetrahydrothiophen	1,87
Selenophen..	0,41			Tetrahydroselenophen ..	1,79
Pyrrol.....	1,80	Pyrrolin.....	1,42	Pyrrolidin	1,57

so sieht man, dass die Dipolmomente des Furans, Thiophens und Selenophens wesentlich kleiner als die der Tetrahydroderivate sind. Die letzteren sind etwas grösser als die der Äther ($\infty 1,2$), Thioäther ($\infty 1,6$), und Selenoäther ($\infty 1,5$), die ersteren dagegen erheblich kleiner. Dies lässt sich dadurch erklären, dass die Bindungsmomente O-C, S-C und Se-C (O, S und Se negativ) durch die Resonanz kleiner werden, weil O, S und Se in den polaren Resonanzformeln positiv geladen sind. Das Dipolmoment des Dihydrofurans zeigt, dass dieser Effekt bereits bei Aufhebung der einen Doppelbindung praktisch verschwunden ist. In Tetrahydropyrrol (Pyrrolidin) ist das Stickstoffatom offenbar positiv, da das Dipolmoment des Pyrrols grösser als das des Pyrrolidins ist. Bei Pyridin ($\mu = 2,11$) kommt man dagegen bei Berücksichtigung der Resonanzformeln zu dem Schluss, dass das Stickstoffatom negativ ist. Wenn man die Dipolmomente von Pyrrol- und Pyridinderivaten vergleichen will, ist also zu beachten, dass sie entgegengesetzt gerichtet sind.

Die Anschauungen von SUTTON erklären somit sehr anschaulich die für die heterocyklischen Ringe mit einem Heteroatom gefundenen Momente. Es ist jedoch zu beachten, dass ein Unterschied zwischen den gefundenen und den aus den Bindungsmo-

¹ L. E. SUTTON, Trans. Faraday Soc. **30** (1934), 789.

² Alle Dipolmomente sind in dieser Arbeit in Debye-Einheiten ($1 \text{ D} = 10^{-18}$ e. s. E.) angegeben.

³ H. DE VRIES ROBLES, Rec. trav. chim. Pays-Bas **58** (1939), 111.

menten berechneten Werten nicht notwendigerweise auf Resonanz (Mesomerie) zurückzuführen ist, sondern auch auf einer intramolekularen, elektrostatischen Induktion beruhen kann¹. Der letztere Effekt kann jedoch nur geringfügigere Unterschiede erklären.

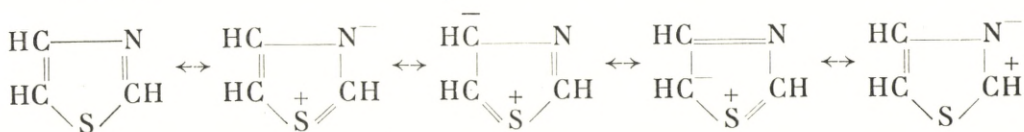
Um den Vorhandensein polarer Resonanzsysteme bei den Azolen zu untersuchen, haben wir die Dipolmomente folgender Verbindungen bestimmt:

	μ		μ
Thiazol	1,64	1,2,3-Triazol	1,77
Isoxazol	2,81	1,2,4-Triazol	3,17
Pyrazol	1,57	Tetrazol	5,11
3-Methylpyrazol	1,43	1-Methyltetrazol	5,38
1-Methylpyrazol	2,28		

Das Oxazol, das bisher nicht dargestellt worden ist, würde, nach dem Dipolmoment des Benzoxazols zu urteilen, ein Dipolmoment von derselben Grössenordnung wie Thiazol haben. Das Dipolmoment des Imidazols wurde von HÜCKEL, DATOW und SIMMERSBACH zu etwa 4,6 D in Dioxan und 6,2 D in Benzol gefunden (vgl. indessen unten).

Man könnte erwarten, dass Heteroatome leichter aufgeladen werden als Kohlenstoffatome, und dass die Dipolmomente deshalb mit der Anzahl der Heteroatome steigen würden. Dies ist wohl auch im grossen und ganzen der Fall; auffallend ist jedoch das kleine Moment von 1,2,3-Triazol und das ungemäin grosse Moment von Imidazol. Eigentümlich ist auch, dass das Dipolmoment von Isoxazol wesentlich grösser als das des Thiazols (Oxazols) ist, während bei den Diazolen die »Iso-Verbindung« (Pyrazol) bei weitem das kleinste Dipolmoment hat.

Für das Thiazol sind folgende Resonanzstrukturen zu erwarten:

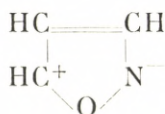


Man könnte erwarten, dass das Thiazol wegen der ersten der obigen Resonanzstrukturen stärker polar sei als Pyridin. Dies

¹ C. G. LE FÈVRE u. R. J. W. LE FÈVRE, Journ. Chem. Soc. 1937, 196, 1088.

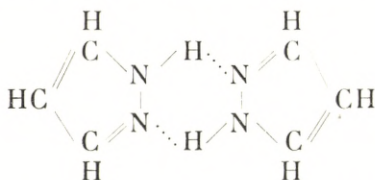
ist aber offenbar nicht der Fall; polare Zustände spielen beim Thiazol eher eine kleinere Rolle als beim Pyridin.

Wenn das Isoxazol ein ziemlich grosses Dipolmoment besitzt, kann dies nicht auf dem Moment der N-O-Bindung beruhen, denn dieses ist, nach den Dipolmomenten der Oxim-O-äther zu urteilen¹, nur klein. Das grosse Moment des Isoxazols muss vielmehr darauf beruhen, dass eine Resonanzstruktur vom Typus:



von Bedeutung ist (vgl. das Dipolmoment von Anthranil, S. 20).

Das von uns gefundene Dipolmoment des Pyrazols ist praktisch identisch mit dem von HÜCKEL u. Mitarb. (l. c.) gefundenen. In einer kürzlich erschienen Abhandlung² hat HÜCKEL jedoch die Hypothese aufgestellt, dass Pyrazol in Wirklichkeit ein grosses Dipolmoment (4-6 D) besitze und dass der kleine gefundene Wert darauf beruhe, dass das Pyrazol weitgehend zu Doppelmolekülen mit dem Dipolmoment Null assoziiert sei:



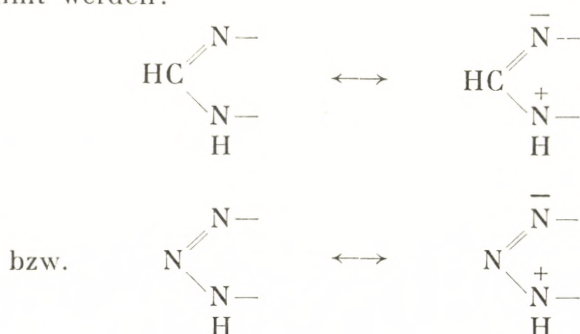
Wenn dies der Fall wäre, müsste das 1-(N-)Methylpyrazol jedoch ein grosses Dipolmoment besitzen, da es nicht unter Bildung von Wasserstoffbrücken assoziieren kann. Tatsächlich ist das Dipolmoment des 1-Methylpyrazols nur wenig grösser als das des Pyrazols. Das Dipolmoment des Pyrazolringesystems kann deshalb nur klein sein; der für Pyrazol gefundene Wert gibt seine Grössenordnung richtig wieder, wenn auch der Wert rein numerisch etwas durch die Assoziation beeinträchtigt wird. Eine genaue Messung des Dipolmoments bei extrem kleinen Konzentrationen konnte nicht durchgeführt werden; diese Messungen zeigen jedoch auch, dass eine wesentliche Erhöhung des Dipolmoments bei der Dissoziation nicht stattfindet.

¹ T. W. J. TAYLOR u. L. E. SUTTON, Journ. Chem. Soc. **1933**, 63.

² W. HÜCKEL u. W. JAHNENTZ, Ber. dtsh. chem. Ges. **75** (1942), 1438.

dere gebunden zu sein, es bildet also eine »Wasserstoffbrücke«¹ die zwei Imidazolmoleküle zusammenhält. HÜCKEL u. Mitarb. haben die Assoziation als Erklärungsmöglichkeit für das grosse Dipolmoment ausser Betracht gelassen, weil bei ihren Messungen das Dipolmoment ziemlich unabhängig vom Assoziationsgrad war. Das von ihnen untersuchte Konzentrationsintervall war jedoch nur klein, und wir haben deshalb das Dipolmoment des Imidazols bei ganz kleinen Konzentrationen gemessen. Dabei zeigte sich, dass das Dipolmoment mit abnehmender Konzentration stark abnimmt und bei unendlicher Verdünnung nur etwa 3,8 D beträgt. Die von HÜCKEL u. Mitarb. gefundenen, grossen Werte beruhen also auf Assoziation, und das monomere Imidazol besitzt in Wirklichkeit ein Dipolmoment von derselben Grösse wie das des 1-(N-)Methylimidazols. In Übereinstimmung mit dieser Erklärung ist das 1-Methylimidazol nicht assoziiert.

Das Dipolmoment des Imidazols ist aber doch noch wesentlich grösser als das des Pyrazols, und das Dipolmoment des 1-Methylimidazols ist ebenfalls grösser als das des 1-Methylpyrazols. Die letzteren Dipolmomente können als charakteristisch für die beiden Ringsysteme betrachtet werden, da sie nicht durch Assoziationseinflüsse kompliziert werden. Zur Erklärung des Unterschiedes zwischen diesen beiden Dipolmomenten glauben wir, dass die von W. HÜCKEL erwähnten Umstände massgebend sind, d. h. dass eine Aufladung vorzugsweise an nicht benachbarten Stickstoffatomen stattfindet. Wir meinen deshalb, dass die grossen Dipolmomente des 1-Methylimidazols, des 1,2,4-Triazols und des Tetrazols hauptsächlich durch folgende Resonanzformeln bestimmt werden:



¹ Über die Wasserstoffbrücke (H-Brücke, the hydrogen bond) vgl. L. PAULING, *The Nature of the Chemical Bond* (New York 1939), S. 264; B. EISTERT, *Tautomerie und Mesomerie* (Stuttgart 1938) S. 165.

Es ist zu beachten, dass sowohl für 1,2,4-Triazol als auch für Tetrazol zwei tautomere Formeln aufgestellt werden können, die aber beide die angeführte Atomgruppierung enthalten, Bei 1, 2, 3-Triazol dagegen unterscheiden die beiden tautomeren Formen:



sich wesentlich, indem die erste mit Imidazol, die zweite aber mit Pyrazol vergleichbar ist. Es wäre also für die erste ein grosses, für die zweite dagegen nur ein kleines Dipolmoment zu erwarten. Da das Dipolmoment des 1,2,3-Triazols tatsächlich nur wenig grösser als das des Pyrazols ist, scheint in dieser Verbindung vorwiegend die zweite tautomere Form vorzuliegen. Für die beiden tautomeren Formen des 1,2,4-Triazols:



wäre dagegen in beiden Fällen ein grosses Dipolmoment zu erwarten. Tatsächlich liegt es erheblich über dem Moment des isomeren 1,2,3-Triazols.

Eine Folgerung aus dieser Annahme ist, dass die Dipolmomente zweier isomerer N-Alkyl- oder N-Aryl-1,2,3-triazole sehr verschieden gross sein sollten; das der 1-substituierten Verbindung sollte gross, das der 2-substituierten Verbindung nur klein sein. Dagegen sollten die Dipolmomente zweier isomerer N-Alkyl(Aryl)-1,2,4-triazole beide gross sein. Diese Voraussage wurde durch Messung der Dipolmomente von vier isomeren Phenyltriazolen bestätigt (der Unterschied der Dipolmomente der beiden Derivate des 1,2,4-Triazols war jedoch grösser als erwartet):

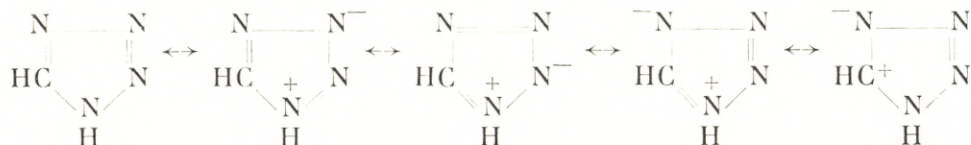
1-Phenyl-1,2,3-triazol	μ	4,08	1-Phenyl-1,2,4-triazol	μ	2,88
2-Phenyl-1,2,3-triazol	0,97	4-Phenyl-1,2,4-triazol	5,63		

Da das Dipolmoment von 4-Phenyl-1,2,4-triazol wesentlich grösser als das des 1,2,4-Triazols ist, scheint in letzterer Verbin-

dung hauptsächlich die erste der angeführten tautomeren Formen vorzuliegen.

Zur Beurteilung der Dipolmomente der C-Alkyl(Aryl)-derivate liegt kein grösseres Material vor. Es ist jedoch nicht zu erwarten, dass diese wesentlich verschieden von denen der Stammkörper sind. 3-Methylpyrazol hat nach unserer Messung fast dasselbe Dipolmoment wie Pyrazol, und 4-Methylimidazol hat nach HÜCKEL u. Mitarb. fast dasselbe Dipolmoment wie Imidazol.

Ein besonders grosses Dipolmoment hat das Tetrazol. Dass dieses Moment für das Ringsystem selbst charakteristisch und nicht etwa wie beim Imidazol auf andere Umstände zurückzuführen ist, geht daraus hervor, dass das 1-Methyltetrazol fast das gleiche Dipolmoment besitzt. Für Tetrazol kommen folgende Resonanzformeln in Betracht:



sowie entsprechende Formeln für die tautomere Form mit dem beweglichen Wasserstoffatom in Stellung 2.

Das Vorhandensein einer Sechsergruppe von Elektronen in den Azolen verleiht ihnen eine Reihe von Eigenschaften, die als typisch aromatisch gelten: Ihre Amine sind nur schwach basisch und können durch Diazotieren in kuppelungsfähige Diazoniumkörper übergeführt werden, ihre Hydroxylderivate haben Phenoleigenschaften: Löslichkeit in Alkali, Kuppelungsfähigkeit usw. Dagegen lassen sich nicht alle Azole nitrieren, halogenieren und sulfurieren, und sie sind nicht alle den Reaktionen von FRIEDEL-CRAFTS, CANNIZZARO, PERKIN, TIEMANN-REIMER und GATTERMANN zugänglich. Das vom Benzol abweichende Verhalten der heterocyclischen Verbindungen lässt sich auf der Grundlage der Resonanzformeln in grossen Zügen verständlich machen, wie dies von OCHIAI und Mitarbeiter¹ durch Untersuchungen an Pyridin, Pyrimidin u. a. gezeigt wurde. Aus den Resonanzformeln von Pyridin:

¹ E. OCHIAI, Journ. pharm. Soc. Japan 59 (1939), 2; E. OCHIAI u. F. NAGASAWA, *ibid.* 43; E. OCHIAI u. M. YANAI, *ibid.* 97; E. OCHIAI u. M. SIBATA, *ibid.* 256.



lässt sich z. B. schliessen, dass das Pyridin in α - und γ -Stellung gegenüber nucleophilen Reagenzien (Natriumamid, Lithiumalkyl) aktiv ist. Gegenüber elektrophilen Reagenzien ist es dagegen sehr beständig, bei energischem Nitrieren, Sulfurieren und Halogenieren findet jedoch Substitution statt, aber nur in β -Stellung.

In ähnlicher Weise lässt sich verständlich machen, dass Pyrazol nur in 4-Stellung und Imidazol nur in 4- oder 5-Stellung nitrierbar sind, während Thiazol, die Triazole und Tetrazol überhaupt nicht mehr dem direkten Nitrieren und Halogenieren zugänglich sind. Man könnte erwarten, dass sie dann gegenüber nucleophilen Reagenzien aktiv wären; das vorliegende experimentelle Material ist aber in dieser Hinsicht noch sehr dürftig, und es lag ausserhalb des Rahmens dieser Arbeit, es zu erweitern.

In ihren physikalischen Eigenschaften weichen die Azole noch mehr vom Benzol ab. Die Azole mit mehreren Stickstoffatomen haben ungewöhnlich hohe Schmelz- und Siedepunkte und sind in Wasser (und anderen hydroxylhaltigen Lösungsmitteln) spielend löslich, während sie schwerlöslich in Benzol und Äther und praktisch unlöslich in Ligroin sind. Die Schmelzpunkte und Siedepunkte der hier untersuchten Körper gehen mit den Dipolmomenten symbar:

	Schmp.	Siedep.	μ
Pyrazol	70°	185°	1,57
Imidazol	90°	256°	3,8
1,2,3-Triazol	23°	203°	1,77
1,2,4-Triazol	121°	260°	3,17
Tetrazol	155°	..	5,11

Demnach beruhen ihre hohen Schmelz- und Siedepunkte teilweise darauf, dass ihre Moleküle durch relativ starke elektrostatische Kräfte zusammengehalten werden. Dass es jedoch nicht genügt, die Dipolmomente heranzuziehen, um die hohen Siedepunkte zu erklären, geht daraus hervor, dass die Azole

mit nur einem Stickstoffatom, wie z. B. Thiazol und Isoxazol, wesentlich niedriger sieden. Das besondere Verhalten der Azole mit zwei, drei und vier Stickstoffatomen beruht vielmehr darauf, dass sie amphotere Verbindungen sind: Die positiv geladene NH-Gruppe verleiht ihnen ausgesprochenen Säurecharakter; wegen des einsamen Elektronenpaares an den anderen Stickstoffatomen haben sie aber gleichzeitig die Fähigkeit, Wasserstoffionen anzulagern. Somit kann durch Bildung von Wasserstoffbrücken eine intermolekulare Assoziation stattfinden, die wohl nicht immer so weitgehend, aber von ähnlicher Art wie beim Imidazol ist (dass die Assoziation beim Imidazol besonders ausgeprägt ist, beruht wahrscheinlich auf dem symmetrischen Bau des Imidazolmoleküls und vielleicht auch auf seiner relativ grossen Basenstärke). Beim N-Alkylieren verschwindet der Säurecharakter dieser Verbindungen und die so erhaltenen Derivate haben deshalb ähnliche Siedepunkte wie Derivate der »Monazole« mit demselben Dipolmoment. Beispielsweise ist 1-Methylpyrazol eine ölige Flüssigkeit von pyridinähnlichem Geruch mit fast demselben Siedepunkt (127°) wie 2-Methylthiazol. Die Bedeutung der N-Substitution für die Schmelz- und Siedepunkte wird weiterhin durch die folgende Tabelle veranschaulicht:

	N-Derivate		C-Derivate	
	Schmp.	Siedep.	Schmp.	Siedep.
1-Methylimidazol...	-6°	198°	2-Methylimidazol...	139° 268°
			4-Methylimidazol...	56° 263°
1-Methylpyrazol....	..	127°	3-Methylpyrazol....	.. 204°
			4-Methylpyrazol....	.. 205°
1-Methyl-1,2,4-triazol	20°	178°	3-Methyl-1,2,4-triazol	95° 265°
4-Methyl-1,2,4-triazol	90°	..		
1-Phenyl-1,2,4-triazol	47°	266°	3-Phenyl-1,2,4-triazol	118° ..
4-Phenyl-1,2,4-triazol	121°	..		
1-Phenyl-1,2,3-triazol	156°	..	5-Phenyl-1,2,3-triazol	145° ..
2-Phenyl-1,2,3-triazol	Öl	..		
1-Phenyltetrazol....	66°	..	5-Phenyltetrazol....	213° ..

In den N-Derivaten ist die Wirkung der NH-Gruppe eliminiert, und die Schmelz- und Siedepunkte werden deshalb hauptsächlich durch das Dipolmoment bestimmt. Dass 1-Methylimidazol höher als 1-Methylpyrazol siedet, ist also lediglich auf das grössere Dipolmoment des Imidazolringes zurückzuführen. Man sieht bei den vier Phenyltriazolen, für welche die Dipolmomente

oben angeführt wurden, dass die Schmelzpunkte und Dipolmomente symbar gehen. Beim Vergleich der Schmelz- und Siedepunkte der N- und C-Derivate muss daran erinnert werden, dass man durch N-Substitution Derivate erhalten kann, die ein wesentlich grösseres Dipolmoment als der Grundkörper besitzen (vgl. oben), und dass diese Wirkung eventuell die Wirkung der Eliminierung der NH-Gruppe neutralisieren könnte. So sieht man, dass die in 4-Stellung substituierten 1,2,4-Triazole (die ein wesentlich grösseres Dipolmoment als 1,2,4-Triazol selbst besitzen) fast dieselben Schmelzpunkte haben wie die 3-Derivate. Es handelt sich also bei der Wirkung des Dipolmoments und der Assoziation um Effekte von derselben Grössenordnung.

Die Löslichkeitsverhältnisse der Di-, Tri- und Tetrazole sind ebenfalls dadurch zu erklären, dass ihre Kristallgitter sowohl durch elektrostatische Kräfte wie durch Wasserstoffbrücken zusammengehalten werden. Ihre Schwerlöslichkeit in den typisch organischen Lösungsmitteln beruht darauf, dass diese Lösungsmittel praktisch nicht im Stande sind, einen derartigen Gitterverband zu lösen. Bei N-Alkylierung wird die eine dieser Kräfte ausser Spiel gesetzt, und wenn das Dipolmoment nur relativ klein ist, bekommt man Verbindungen, wie z. B. 1-Methylpyrazol, die in Äther und Ligroin leichtlöslich sind. Wenn das Dipolmoment gross ist, sind auch die N-Alkylderivate, wie z. B. 1-Methyltetrazol, stets wenig löslich in diesen Lösungsmitteln. Hydroxylhaltige Lösungsmittel besitzen dagegen gutes Lösungsvermögen, weil sich zwischen dem einsamen Elektronenpaar der Stickstoffatome und der Hydroxylgruppe eine Wasserstoffbrücke herstellen kann. Wasserlöslichkeit besteht also, sobald ein derartiges einsames Elektronenpaar vorhanden ist, weshalb auch Verbindungen wie Pyridin und Thiazol wasserlöslich sind. Da die hydroxylhaltigen Lösungsmittel amphoter sind, wird die Löslichkeit aber noch erhöht, wenn die zu lösende Verbindung eine saure NH-Gruppe enthält, die sich ebenfalls durch Herstellung einer Wasserstoffbrücke an der Bindung mit den Molekülen des Lösungsmittels beteiligen kann. Nach den polaren Resonanzformeln haben die Azole mit den grossen Dipolmomenten (Imidazol, die Triazole, Tetrazol) eine formale Ähnlichkeit mit Aminosäuren, und tatsächlich erinnern mehrere ihrer physikalischen Eigenschaften an diejenigen der Aminosäuren.

Die Basenstärke der Azole ist durchweg nur klein (eine Ausnahme bildet nur das Imidazol), und es scheint kein Zusammenhang zwischen Basenstärke und Dipolmoment zu bestehen. Dies steht im Widerspruch zu den polaren Resonanzstrukturen, wonach man bei erhöhter negativer Ladung der Stickstoffatome auch eine erhöhte Basenstärke erwarten könnte. Dass die Stickstoffatome in benzoiden heterocyclischen Ringen nur schwach basisch sind, beruht wahrscheinlich darauf, dass eine Wechselwirkung zwischen den π -Elektronen und den σ -Elektronen des einsamen Elektronenpaares stattfindet; es fehlt aber noch das theoretische Verständnis einer derartigen Wechselwirkung, und es lässt sich deshalb nichts darüber aussagen, von welchen Faktoren sie beeinflusst wird. Es ist z. B. bis jetzt ganz unerklärlich, warum Imidazol eine wesentlich stärkere Base ($p_B = 7$) als die anderen Azole ($p_B \approx 12$) ist. Dagegen ist die Acidität der Azole gut erklärbar, weil in den Resonanzformeln eine positiv geladene NH^+ -Gruppe auftritt. Falls die Basenstärke nur wenig vom Dipolmoment beeinflusst wird, ist deshalb auch ein Anstieg der Säurestärke mit dem Dipolmoment zu erwarten. Während Pyrazol eine ausserordentlich schwache Säure ist, haben 1,2,3-Triazol und 1,2,4-Triazol tatsächlich eine ähnliche Säurestärke wie Phenol, und Tetrazol lässt sich sogar mit Phenolphthalein als Indikator als einbasische Säure titrieren.

Wir haben ferner den Einfluss der Einführung von Hydroxyl- und Aminogruppen in die heterocyclischen Ringe untersucht (derartige Verbindungen können auch als Oxo- und Iminoverbindungen formuliert werden). Bei Gegenwart dieser Substituenten erhöht sich die Anzahl der möglichen Resonanzstrukturen, und es bestünde die Möglichkeit, dass polare Resonanzstrukturen bei derartigen Verbindungen von grösserer Bedeutung wären als bei den Stammkernen. Für 2-Aminothiazol wäre z. B. mit folgenden Formeln zu rechnen:

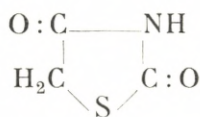


Die Messungen zeigen jedoch, dass eine derartige »extranucleare« Resonanz gewöhnlich nur geringfügig ist; die Dipolmo-

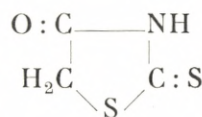
mente der Oxy- und Aminoderivate sind durchweg von derselben Grössenordnung wie die der Stammkerne:



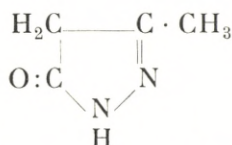
2-Aminothiazol
 $\mu = 1,75$



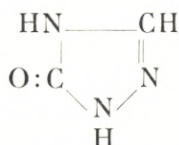
2,4-Dioxothiazolidin
(»Senfölessigsäure«)
 $\mu = 2,03$



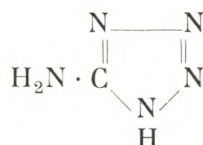
2,4-Thio-oxo-thiazolidin
(»Rhodanin«)
 $\mu = 2,20$



3-Methyl-pyrazolon(5)
 $\mu = 2,54$



1,2,4-Triazolon(5)
 $\mu = 3,30$



5-Aminotetrazol
 $\mu = 5,71$

In gewissen Fällen können Resonanzformeln von der angegebenen Art jedoch von ausschlaggebender Bedeutung sein, wie dies aus den grossen Dipolmomenten des Antipyrins und Thio-pyrins hervorgeht (vgl. S. 29).

Endlich wurden die Dipolmomente einiger kondensierter Ringsysteme vom Naphtalintypus bestimmt:

	μ		μ
Benzoxazol	1,47	Benzimidazol	3,93
Benz- α , β -isoxazol (Indoxazen)	3,03	Benzpyrazol (Indazol)	1,83
Benz- β , γ -isoxazol (Anthranil)	3,06	Benztriazol	4,07

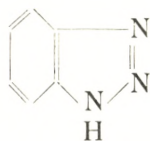
Von den fünfgliedrigen Ringen mit einem Heteroatom ist Indol von COWLEY und PARTINGTON¹ gemessen worden. Sein Dipolmoment wurde zu 2,05 gefunden, ein wenig grösser als das des Pyrrols (1,80 D).

Da im Naphtalinmolekül nur 10 π -Elektronen zur Verfügung stehen, wird der Bindungszustand im Naphtalin weniger aromatisch als der des Benzols sein. Wegen des unsymmetrischen Baus der Benzoderivate der Azole ist es jedoch nicht wahrscheinlich, dass der aromatische Charakter bei diesen Verbindungen in beiden Ringen gleichmässig herabgesetzt ist. Dass Verbindun-

¹ E. G. COWLEY u. J. R. PARTINGTON, Journ. Chem. Soc. **1936**, 47.

D. Kgl. Danske Vidensk. Selskab, Mat. fys. Medd. XX, 20.

gen wie Benzotriazol und Benzimidazol bei der Oxydation unter Zerstörung des Benzolringes Dicarbonsäuren der entsprechenden Azole geben, könnte darauf deuten, dass die Beständigkeit des Benzolringes stärker herabgesetzt ist als die des Heteroringes. Jedenfalls scheint der Bindungszustand in den kondensierten Azolderivaten nicht sehr verschieden von dem der einfachen Azole zu sein, und es wäre deshalb auch zu erwarten, dass ihre Dipolmomente von ähnlicher Grösse wie die der nichtkondensierten Verbindungen wären. Tatsächlich sind die Dipolmomente der meisten der Verbindungen von der erwarteten Grösse. Das Dipolmoment des Benzotriazols liegt jedoch erheblich über dem des 1,2,3-Triazols, und es scheint deshalb, dass die kondensierte Verbindung überwiegend in der tautomeren Form mit dem Wasserstoffatom in 1-Stellung vorliegt:



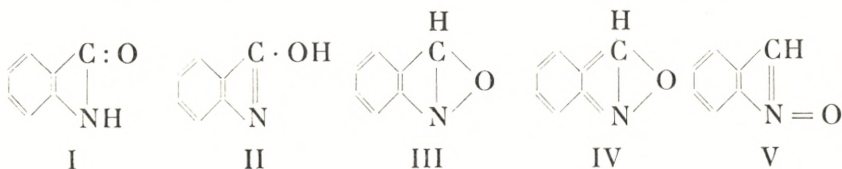
Von besonderem Interesse ist der Vergleich der Dipolmomente des Benzoxazols, des Indoxazens und des Anthranils. Die Konstitution des Anthranils ist bekanntlich eines der am heftigsten diskutierten Probleme der heterocyclischen Chemie. Die Entdecker des Anthranils, FRIEDLÄNDER und HENRIQUES¹, hielten es für ein inneres Anhydrid der Anthranilsäure, I (tautomer mit II), später kamen jedoch FRIEDLÄNDER und SCHREIBER² zu dem Schluss, dass die Auffassung der Verbindung als ein inneres Anhydrid des *o*-Hydroxylaminobenzaldehyds (III) wahrscheinlicher sei. Die letztere Auffassung, wonach Anthranil als Benz- β,γ -isoxazol zu bezeichnen ist, wurde von BAMBERGER durch sehr eingehende Untersuchungen gestützt³, andererseits wurde sie auch von einigen Chemikern abgelehnt, z. B. entspann sich über diese Frage eine sehr leidenschaftliche

¹ P. FRIEDLÄNDER u. R. HENRIQUES, Ber. dtsch. chem. Ges. **15** (1882), 2105.

² P. FRIEDLÄNDER u. W. SCHREIBER, Ber. dtsch. chem. Ges. **28** (1895), 1382.

³ E. BAMBERGER, Ber. dtsch. chem. Ges. **36** (1903), 819; E. BAMBERGER u. E. DEMUTH, *ibid.* **34** (1901), 4015; E. BAMBERGER u. F. ELGER, *ibid.* **36** (1903), 3645; E. BAMBERGER u. J. LUBLIN, *ibid.* **42** (1909), 1676; E. BAMBERGER, *ibid.* **42** (1909) 1647; Journ. prakt. Chem. **81** (1910), 254.

Diskussion zwischen BAMBERGER und G. HELLER, der für die Laktamformel (I) mehrere Argumente ins Feld führte¹.



Die refraktometrischen Konstanten des Anthranils glaubten ANSCHÜTZ und SCHMIDT² als Beweis für Formel I verwenden zu können, während BRÜHL³ sie zugunsten der Formel III auslegte. Dieser »von beiden Seiten mit ebensoviel Scharfsinn wie Gründlichkeit geführte Streit« (v. AUWERS), führte indessen auch nicht zu einer klaren Entscheidung. Später hat v. AUWERS⁴ jedoch das Anthranil vom spektrochemischen Standpunkte aus wiederum untersucht, und seine Ergebnisse sprechen entschieden gegen die Laktamformel und für die Benzisoxazol-Formel, wenn man nicht das Symbol III, sondern das Symbol IV zu Grunde legt. Auch das Ramanspektrum von Anthranil stimmt nach KOHLRAUSCH⁵ am besten mit dieser Formel überein.

Von STAUDINGER und MIESCHER wurde die »Nitronformel« (V) vorgeschlagen⁶; diese wurde aber von BAMBERGER⁷ wiederum abgelehnt.

Alles in allem spricht das grosse Material als Ganzes entschieden für die Benzisoxazol-Formel, wenn auch kein einzelnes Argument zwingend ist. Selbst HELLER schreibt in seiner zusammenfassenden Abhandlung über die Konstitution des Anthranils: »Will man auf Grund des vorliegenden Materials eine Formel für das Anthranil aufstellen, so könnte es nur eine Gleichgewichtsformel sein ...«.

¹ G. HELLER, Journ. prakt. Chem. **77** (1908), 145, 161.

G. HELLER, Über die Konstitution des Anthranils. Samml. chem. u. chem.-techn. Vorträge **23** (1917), 327.

² R. ANSCHÜTZ u. O. SCHMIDT, Ber. dtsch. chem. Ges. **35** (1902) 3470. O. SCHMIDT, Ber. dtsch. chem. Ges. **36** (1903), 2465. Zeitschr. phys. Chem. **58** (1907), 539.

³ J. W. BRÜHL, Ber. dtsch. chem. Ges. **36** (1903), 3637, 4294. Zeitschr. phys. Chem. **59** (1907), 507.

⁴ K. v. AUWERS, Liebigs Ann. d. Chem. **437** (1924), 63.

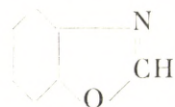
⁵ K. W. F. KOHLRAUSCH u. R. SEKA, Ber. dtsch. chem. Ges. **71** (1938), 1564.

⁶ H. STAUDINGER u. K. MIESCHER, Helv. chim. acta **2** (1919), 555. K. MIESCHER, Nitrene und Nitrene. Dissertation, Zürich 1918.

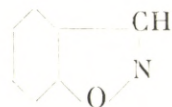
⁷ E. BAMBERGER, Helv. chim. acta **7** (1924), 814.

Ein Gleichgewicht zwischen Molekülen der Formel I (bzw. II) und III (bzw. IV) wird man jedoch nicht ohne besonders zwingende Gründe annehmen können, und man wird versuchen, ob die experimentellen Tatsachen sich nicht mit der Benzisoxazol-Formel in Übereinstimmung bringen lassen. Dass viele Anthranilderivate — z. B. C-Methylantranil, C-Phenylantranil und Anthroxansäure (Anthranil-C-carbonsäure) — nur von der Benzisoxazol-Formel abgeleitet werden können, spricht wohl auch zu Gunsten der Isoxazolformel des Grundkörpers, wenn auch dieses Argument von HELLER nicht als zwingend angesehen wurde.

Die Dipolmessung am Anthranil liefert nun einen Beitrag zur Beleuchtung dieses Problems. Das Dipolmoment des Anthranils wurde zu 3,06 gefunden, ein ähnlicher Wert wie der der isomeren Verbindung Indoxazen, während für Benzoxazol ein wesentlich kleinerer Wert gefunden wurde:



Benzoxazol
 $\mu = 1,47$



Indoxazen
 $\mu = 3,03$

Da das einfache Isoxazol das Dipolmoment 2,83 besitzt, deuten diese Werte darauf hin, dass Anthranil ebenso wie Indoxazen einen Isoxazol-Ring enthält. Es fragt sich aber, warum das Dipolmoment des Anthranils wesentlich grösser als das des Benzoxazols ist. Die beiden bisher aufgestellten Formeln III und IV geben keine Erklärung hierfür. Es sei hier sogleich festgestellt, dass diese beiden Formeln nur mesomere Grenzzustände desselben Moleküls darstellen (LEUCHS¹ hat versucht die Anthroxansäure in optisch aktive Komponenten zu spalten, um eine Entscheidung zwischen diesen beiden Formeln zu treffen; diese Aufgabe ist aber nach unserem jetzigen Wissen gegenstandslos, da die Formeln sich nur in der Verteilung der Elektronen unterscheiden). Ausser diesen beiden lassen sich aber noch mehrere andere mesomere Grenzzustände des Moleküls formulieren, von denen namentlich dem von der auf Seite 8 angeführten polaren Formel des Isoxazols abgeleiteten eine ge-

¹ H. LEUCHS, Ber. dtsh. chem. Ges. 58 (1925), 1452.

wisse Bedeutung zukommen dürfte. Ein Benz- β,γ -isoxazol würde demnach folgendermassen zu formulieren sein:



Von diesen drei Formeln erklären die beiden ersten die Mehrzahl der Eigenschaften des Anthranils, die dritte erklärt aber das grosse Dipolmoment und diejenigen Reaktionen — z. B. die Bildung der sogenannten Acylantranile (die wahrscheinlich Metoxazine sind) —, bei denen das Stickstoffatom der Angriffspunkt ist, und die von HELLER als Beweis für die Laktamformel angesehen wurden. Da die Benzisoxazolformel somit dem gesamten Verhalten des Anthranils gerecht wird, während andererseits mehrere Eigenschaften des Anthranils — z. B. die enge Beziehung zu *o*-Hydroxylaminobenzaldehyd und die Ähnlichkeit in den chemischen und physikalischen Eigenschaften zwischen Anthranil und *C*-Methylantranil — durch die Laktamformel schwer erklärbar sind, darf man wohl jetzt annehmen, dass Anthranil tatsächlich Benz- β,γ -isoxazol ist.

Es muss hervorgehoben werden, dass die Dipolmessung an sich die Laktamformel nicht ausschliesst, denn diese könnte auch ein Dipolmoment in der Nähe von 3 D ergeben (Dipolmoment des Benzamids 3,5 D). Dagegen wird bereits durch die Dipolmessung die Nitronformel V ausgeschlossen; eine derartige Verbindung müsste als ein Aminoxyd ein wesentlich grösseres Moment als das gefundene besitzen (vgl. das Dipolmoment des Dimethylfuroxans, S. 30).

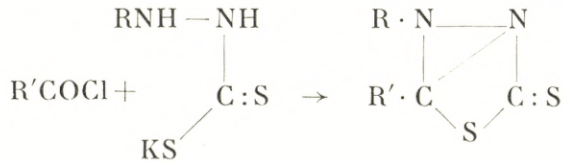
3. Über die Konstitution der Endo-Verbindungen von Busch.

Im Laufe der umfassenden »Untersuchungen in der Biazolreihe« von M. BUSCH und Mitarbeitern² wurden einige sehr in-

¹ Nichtnachbaratome werden am besten nicht mit Bindungsstrichen verbunden; vgl. E. HÜCKEL, Grundzüge der Theorie ungesättigter und aromatischer Verbindungen (Berlin 1938), S. 23.

² M. BUSCH u. Mitarb., Journ. prakt. Chem. **60** (1899), 187; **67** (1903), 201, 216, 243, 257, 263. M. BUSCH, Journ. prakt. Chem. **124** (1930), 301. M. BUSCH u. G. MEHRTENS, Ber. dtsch. chem. Ges. **38** (1905), 4049.

teressante Verbindungen, die »Isodithiobiazolone« entdeckt. Diese Verbindungen wurden zuerst bei einem eigenartigen Zerfall gewisser Thiazolindisulfide beobachtet und dann durch Einwirkung von Carbonsäurechloriden auf phenyldithiocarbazinsäures Kalium synthetisiert. Die letztere Reaktion wurde von Busch anfangs folgendermassen formuliert:

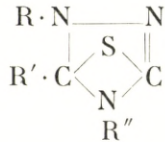


Später wurden diese Verbindungen jedoch — u. a. weil sie sich bei der Einwirkung von HgO nicht entschwefeln lassen — als Endothio-thiodiazoline formuliert:

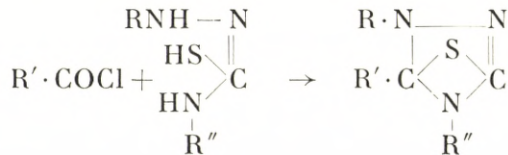


Die genannte Reaktion hat sich als sehr verallgemeinerungsfähig erwiesen.

Durch Einwirkung primärer aromatischer Amine auf diese Verbindungen wird ein Schwefelatom gegen NR'' ausgetauscht, und man kommt dann zu Verbindungen, die von Busch als Endothiotriazoline (= Endiminothiodiazoline) formuliert werden:



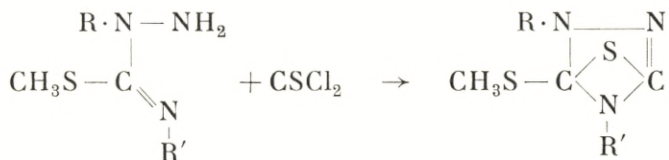
Sie lassen sich auch in direkter Synthese aus 1,4-Diarylthiosemicarbaziden und Carbonsäurechloriden gewinnen:



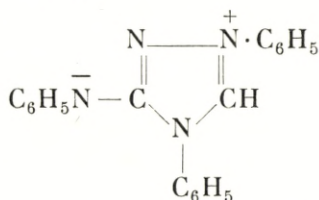
Die Tendenz zur Bildung dieser Verbindungen ist sogar so gross, dass sie sich durch Einwirkung von Benzaldehyd auf 1,4-Dialkylthiosemicarbazide bei höherer Temperatur bilden (der dabei entstehende Wasserstoff wird durch Nebenreaktionen verbraucht).

In ähnlicher Weise wie die Endothiotriazoline können Endoxytriazoline und Endiminotriazoline durch Wasserabspaltung aus Acyl-1,4-diarylsemicarbaziden und Acyl-1,4-diarylaminoguanidinen bereitet werden. Das bekannte Salpetersäurereagens »Nitron« (Diphenyl-endoanilotriazolin) gehört zur letzteren Kategorie.

Endoxy- und Endothio-Verbindungen entstehen ferner nach BUSCH durch Einwirkung von Phosgen und Thiophosgen auf S-Methyl-2,4-diarylisothiosemicarbazide, z. B.:



Die endo-Verbindungen von BUSCH nehmen eine Sonderstellung ein; die für diese Verbindungen postulierte Verknüpfung eines viergliedrigen und eines fünfgliedrigen Ringes findet man bei keiner anderen endo-Verbindung. WITTIG hat auf Grund der BREDT'schen Regel und der Spannungsverhältnisse die Richtigkeit der von BUSCH aufgestellten Formel des Nitrons in Zweifel gestellt¹. Ebenfalls auf Grund der BREDT'schen Regel² (»In Brückenringssystemen ist eine Doppelbindung am Brückenkopf unmöglich«), die sich in der Terpenchemie gut bewährt hat, verwirft A. SCHÖNBERG die Formeln von BUSCH und stellt eine neue Formel auf, wonach die Verbindungen als Amphoionen (Zwitterionen) betrachtet werden³. Beispielsweise wird das »Nitron« folgendermassen formuliert:

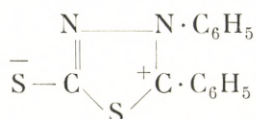


¹ G. WITTIG, Stereochemie (Leipzig 1930), S. 179 (Fussnote 3).

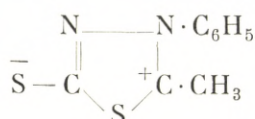
² J. BREDT, Liebigs Ann. d. Chem. **437** (1924), 1.

³ A. SCHÖNBERG, Journ. Chem. Soc. **1938**, 824.

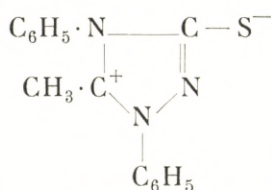
F. L. WARREN¹ hat auf die Aufforderung SCHÖNBERG's das Dipolmoment von Nitron bestimmt und den Wert 7,2, gefunden. Dieser grosse Wert rechtfertigt die Formel von SCHÖNBERG. Wegen der weittragenden Bedeutung der Annahme von SCHÖNBERG schien es angemessen, sie noch an einigen anderen endo-Verbindungen zu prüfen. Einer genaueren Bestimmung der Dipolmomente dieser Verbindungen steht allerdings ihre Schwerlöslichkeit in dipolfreien Lösungsmitteln im Wege. Grössenordnungsmässig lassen sich jedoch die Dipolmomente bestimmen. Es wurden folgende Verbindungen untersucht:



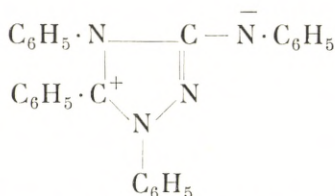
»2,3-Diphenyl-2,5-endothio-
1,3,4-thiodiazolin«
 $\mu = 8,8$ (Benzol)
 $\mu = 9,1$ (Dioxan)



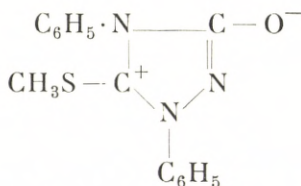
»2-Methyl-3-phenyl-2,5-endothio-
1,3,4-thiodiazolin«
 $\mu = 8,8$ (Benzol)



»1,4-Diphenyl-5-methyl-3,5-
endothio-1,2,4-triazolin«
 $\mu = 8,4$ (Chloroform)



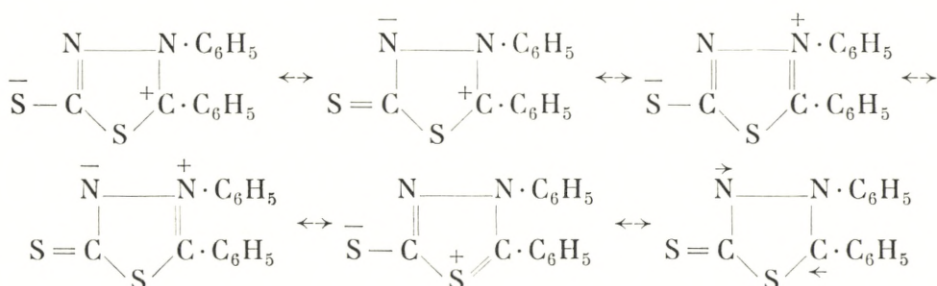
»1,4,5-Triphenyl-3,5-endoanilo-
1,2,4-triazolin«
 $\mu = 9,1$ (Benzol)
 $\mu = 8,8$ (Dioxan)



»1,4-Diphenyl-5-methylmercapto-
3,5-endoxy-1,2,4-triazolin«
 $\mu = 7,7$ (Dioxan)

¹ F. L. WARREN, Journ. Chem. Soc. 1938, 1100.

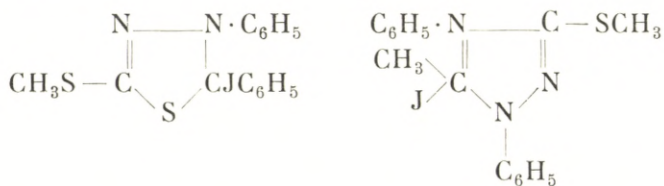
Die Dipolmomente sind also sehr gross; sie sind von derselben Grössenordnung wie die der Alkalihalogenide. Die Formulierung dieser Verbindungen als endo-Verbindungen ist deshalb nicht zutreffend, sondern die Verbindungen sind als Amphoionen zu formulieren. Jedoch lässt ihre Konstitution sich nicht durch eine einzige Formel wiedergeben, sondern es liegt Resonanz zwischen mehreren Formeln vor, was auch bereits von SCHÖNBERG vermutet wurde. Dies geht daraus hervor, dass die Dipolmomente dieser Verbindungen, obwohl sehr gross, nicht die Werte erreichen, die sich für die einzelnen Amphoionenstrukturen berechnen lassen. Den angeführten Formeln entsprechen Dipolmomente von der Grössenordnung 20 D; mesomere Formeln mit den Ladungen an Nachbaratomen würden Dipolmomente von der gefundenen Grössenordnung ergeben, derartige Formeln würden jedoch (wenn sie nicht gleichzeitig biradikaloid sind) die Ladungen an zwei Stickstoffatomen tragen und kommen deshalb nach dem chemischen Verhalten der Verbindungen jedenfalls als alleinige Möglichkeiten nicht in Betracht. Es ist deshalb wahrscheinlich, dass für die Konstitution dieser Verbindungen auch unpolare, biradikaloide Formeln von Bedeutung sind, so dass z. B. die Konstitution von »Diphenyl-endothio-thiodiazolin« durch folgendes mesomere System wiedergegeben werden kann:



Die letzte dieser Formeln ist äquivalent mit der ersten von BUSCH aufgestellten Formel, die einen Dreiring enthielt.

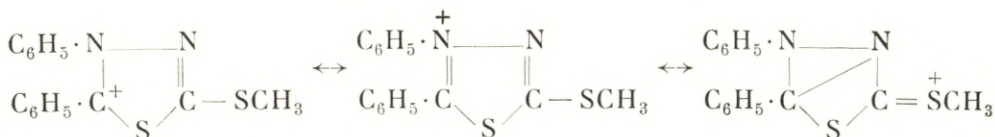
Die Eigenschaften der »endo-Verbindungen« sind in bester Übereinstimmung mit diesen Formeln. Es sind sehr beständige, hochschmelzende Verbindungen, die in den gewöhnlichen Solventien sehr schwerlöslich sind. Sehr charakteristisch ist ihre Reaktion mit Methyljodid, womit sie sehr glatt unter Selbst-

erwärmung Jodmethylate bilden. Die Reaktionsprodukte aus »Diphenyl-endothio-thiodiazolin« und »Diphenyl-methyl-endothio-triazolin« mit Methyljodid wurden von BUSCH folgendermassen formuliert:



Nach unserer Erfahrung sind diese Verbindungen jedoch wahre Salze. Ihre Dipolmomente in Chloroformlösung betragen 11 und 13 D, was einem Abstand von 2,3—2,7 Å zwischen den Elementarladungen entspricht. Ihrer Leitfähigkeit in wässriger bzw. alkoholischer Lösung gemäss verhalten sie sich wie binäre starke Elektrolyte.

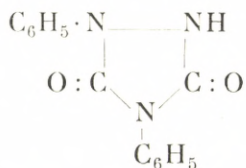
Diese Verbindungen sind also als Onium-Verbindungen zu formulieren. Ihre Eigenschaften — hohe Schmelzpunkte, Löslichkeit in Wasser und Alkohol, nicht aber in Benzol und Äther — sind in Übereinstimmung hiermit. Die Ladung des Kations kann jedoch nicht lokalisiert werden, sondern dieses Ion muss als mesomeres System formuliert werden, z.B.:



Das Kation ist also gleichzeitig ein Carbonium-, Ammonium-, und Sulfonium-ion; man darf aber wohl annehmen, dass die mittlere Formel die grösste Bedeutung hat, und kann diese benutzen, wenn man die Konstitution der Verbindungen dieses Typus mit einer einzigen Formel wiedergeben will.

Die entsprechenden Hydroxyde sind sehr unbeständig und konnten von BUSCH nicht in wohldefiniertem Zustand erhalten werden. Dagegen sind die Alkoholate wohldefinierte Verbindungen, die niedrige Schmelzpunkte besitzen und löslich in Äther, Benzol und Ligroin sind. Das aus dem ersten der oben angeführten Jodide gebildete Äthylat hat das Dipolmoment 1,83 D. Demnach

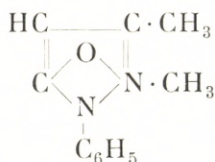
Das Dipolmoment, das nur 2,6 D beträgt beweist aber, dass die Verbindung einfach Diphenyl-dioxo-triazolidin (Diphenyl-urazol) ist:



oder richtiger: nach der klassischen Formulierung repräsentieren die zwei angeführten Formeln zwei verschiedene Verbindungen, nach den neuen Formeln der endo-Verbindungen sind sie aber nur mesomere Grenzzustände desselben Moleküls¹, und das Dipolmoment besagt, dass die erste Formel ohne Bedeutung für die Konstitution dieser Verbindung ist².

Diese Verbindung unterscheidet sich aber auch von den übrigen endo-Verbindungen von BUSCH dadurch, dass sie auf klassischer Grundlage monocyclisch formuliert werden konnte.

In diesem Zusammenhang ist die Konstitution von Antipyrin von Interesse. Antipyrin wurde von MICHAELIS bicyclisch formuliert³:



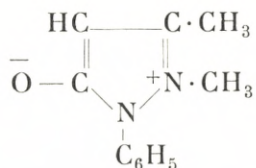
Ähnlich wurden auch Thiopyrin (Thioantipyrin) und ähnliche Verbindungen formuliert.

Diese Formel kommt natürlich an sich jetzt nicht mehr in Frage schon deshalb, weil sie ein fünfwertiges Stickstoffatom enthält; sie ist aber äquivalent mit folgender Phenolbetainformel:

¹ S. L. JANNIAH u. P. C. GUHA (Journ. Amer. Chem. Soc. **52** (1930), 4860) haben zwei Modifikationen (Schmp. 234 und 244) von 2-Thiol-5-imino-thio-diazolin erhalten und formulieren die eine als »Endoimino-thiol-thiodiazolin«. Diese Formulierung ist also nicht zutreffend. Wir haben bei der Nachprüfung der Darstellung dieser Verbindung nur eine Modifikation erhalten können.

² Dasselbe gilt unzweifelhaft für die entsprechenden Thiourazole, die BUSCH auch als endo-Verbindungen formulierte (Ber. dtsch. chem. Ges. **42** (1909), 4763).

³ A. MICHAELIS, Liebigs Ann. d. Chem. **320** (1901), 45.



Nach den obigen Ergebnissen mit Diphenylurazol war es nicht sehr wahrscheinlich, dass eine derartige Formulierung zutreffend ist. Auf der anderen Seite sprechen die Eigenschaften des Antipyryns und Thiopyryns sehr für eine solche: die Verbindungen sind leichtlöslich in Wasser und reagieren analog den endo-Verbindungen von BUSCH mit Methyljodid unter Bildung von Jodmethylenen, welche wahre Salze sind; sehr auffallend ist ferner, dass Thiopyrin durch Einwirkung von HgO nicht entschweifelt wird.

HALBEDEL¹ hat die Dielektrizitätskonstante von wässrigen Lösungen von Antipyryn gemessen und schliesst aus der Tatsache, dass kein Zuwachs der DK der Lösung gegen das Lösungsmittel zu verzeichnen war, dass eine Zwitterionenformel nicht in Frage kommt. Nach den Ergebnissen von HÜCKEL und JAHNENTZ² erhöht jedoch auch Imidazol trotz seines grossen Dipolmoments nicht die DK des Wassers, und es scheint deshalb, dass die durch Untersuchung der Aminosäuren gewonnenen Ergebnisse nicht auf Verbindungen mit verhältnismässig geringem Abstand zwischen den Ladungen ausgedehnt werden können.

Wir haben das Dipolmoment von Antipyryn in Benzollösung gemessen und den grossen Wert 5,48 gefunden. Für Thiopyryn wurde sogar das Dipolmoment zu 7,33 gefunden. Für diese Verbindungen spielen also polare Formeln der angegebenen Art eine nicht zu vernachlässigende Rolle. Für Antipyryn berechnet sich nach der polaren Formel das Dipolmoment zu etwa 18 D und für Thiopyryn zu etwa 20 D³; der Beitrag der polaren Formeln zur Konstitution dieser Verbindungen beträgt also 30—35 %.

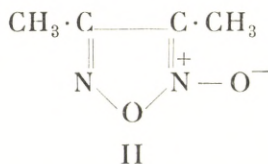
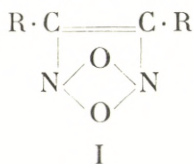
Der Unterschied zwischen Antipyryn und den endo-Verbindungen von BUSCH ist also nur ein gradueller, und die Übereinstimmung in ihrem chemischen Verhalten ist somit verständlich.

¹ G. HALBEDEL, Zeitschr. phys. Chem. Abt. B, **33** (1936), 83.

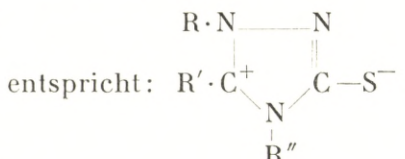
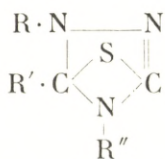
² W. HÜCKEL u. W. JAHNENTZ, Ber. dtsch. chem. Ges. **74** (1941), 652.

³ Atomabstände nach L. PAULING, The Nature of the Chemical Bond, S. 154; Valenzwinkel 110°.

Schliesslich haben wir das Dipolmoment von Dimethylfuroxan bestimmt. Die Furoxane sind auch als endo-Verbindungen (I) formuliert worden¹. Das Dipolmoment des Dimethylfuroxan ist recht gross, 4,4 D, jedoch nicht so gross wie die Dipolmomente der endo-Verbindungen von BUSCH. Die Grösse des Dipolmoments zeigt, dass die Verbindung ein Aminoxyd (II) ist; die einfachen Aminoxyde haben Dipolmomente von derselben Grössenordnung².



Nach den jetzt bewiesenen Formeln der endo-Verbindungen von BUSCH entsprechen jeder der Formeln von BUSCH zwei verschiedene Verbindungen, z. B.:



Endothiothiazolin oder
Endoiminothiodiazolin

III

IV

SCHÖNBERG hat vermutet, diese beiden Verbindungen könnten leicht in einander umgewandelt werden, und versucht in dieser Weise zu erklären, dass einige dieser Verbindungen in zwei verschiedenen Modifikationen mit etwas verschiedenen Schmelzpunkten erhalten worden sind. Es hat jedoch den Anschein, dass in diesen Modifikationen eher polymorphe Formen vorliegen. Verbindungen mit den zwei Formeln III und IV sollten ausgesprochene chemische Unterschiede aufweisen. Wahrscheinlicher ist, dass Verbindungen der Formel IV unbeständig sind. Bei der Einwirkung von aliphatischen Aminen auf die Jodmethylate der »Endothiothiodiazoline« bilden sich unter Entwicklung von Methylmercaptan Produkte, die nicht — wie zu erwarten — nach Formel IV, sondern nach Formel III konstituiert sind, wie

¹ A. G. GREEN, u. F. M. ROWE, Journ. Chem. Soc. **103** (1913), 898. M. O. FORSTER u. M. F. BARKER, *ibid.* **103** (1913), 1919.

² E. P. LINTON, Journ. Amer. Chem. Soc. **62** (1940), 1945.

aus ihrer Fähigkeit, Methyljodid wiederum unter Bildung von S-Methylderivaten zu addieren, eindeutig hervorgeht. In diesem Zusammenhang sei auch auf die labilen Phenylimino-thiodiazolidine hingewiesen, die sich beim Schmelzen in stabile Triazolinderivate (Thiourazole) umwandeln¹.

Experimenteller Teil.

1. Darstellung der untersuchten Verbindungen.

Die untersuchten Verbindungen sind sämtlich bekannt und zum Teil bereits sehr eingehend untersucht worden. Es genügt deshalb, als Kriterium auf Reinheit ihre Schmelz- und Siedepunkte heranzuziehen. Die angegebenen Temperaturen sind sämtlich korrigiert.

Für Thiazol wurde eine neue und für Chlorthiazol eine verbesserte Darstellungsmethode ausgearbeitet.

Da für die endo-Verbindungen eine rationelle Nomenklatur fehlt, werden sie nach der »Beilstein«-Nomenklatur benannt.

Pyrazol. Dargestellt aus Vinylbromid und Diazomethan². Umkristallisiert aus Ligroin. Schmp. 70°.

1-Methylpyrazol. Dargestellt nach DEDICHEN³ aus Pyrazol und Methyljodid. Kp. 127°—127,5°.

3-Methylpyrazol. Darstellung nach KNORR⁴ aus Natriumacetessigaldehyd und Hydrazinhydrat. Kp. 205°.

Imidazol. Durch Decarboxylierung der Imidazol-4,5-dicarbon-säure dargestellt⁵. Aus Benzol umkristallisiert. Schmp. 87,5°.

1,2,3-Triazol. Die Verbindung wurde zunächst durch Decarboxylierung der Dicarbonsäure⁶ dargestellt, aber nur in sehr geringer Ausbeute und nicht ganz rein erhalten. Die für die Dipolmessung benutzte Verbindung wurde dann aus Stickstoffwasserstoffsäure und Acetylen dargestellt. DIMROTH⁷ hat die Bildung von 1,2,3-Triazol bei dieser Umsetzung nachgewiesen, die Ver-

¹ M. BUSCH u. O. LIMPACH, Ber. dtsch. chem. Ges. **44** (1911), 560, 566. M. BUSCH, *ibid.* **42** (1909), 4763.

² E. OLIVERI-MANDALÁ, Gazz. chim. ital. **40**, I (1910), 120.

³ G. DEDICHEN, Ber. dtsch. chem. Ges. **39** (1906), 1845.

⁴ L. KNORR u. J. MACDONALD, Liebigs Ann. d. Chem. **279** (1894), 217.

⁵ R. G. FARGHER u. F. L. PYMAN, Journ. Chem. Soc. **115** (1919), 228.

⁶ J. A. BLADIN, Ber. dtsch. chem. Ges. **26** (1893), 2737.

⁷ O. DIMROTH u. G. FESTER, Ber. dtsch. chem. Ges. **43** (1910), 2222.

bindung wurde jedoch nur als 1-Methyltriazol isoliert. Nach unseren Erfahrungen lässt sich die Reaktion zwischen N_3H und C_2H_2 auch präparativ auswerten, obwohl die Ausbeute nur gering ist: Aus 10 g Natriumazid wurde die Stickstoffwasserstoffsäure freigemacht¹ und (zusammen mit etwas Wasser) in ein Gemisch von 100 ccm Aceton und 50 ccm Alkohol eindestilliert. Die Lösung wurde mit Acetylen gesättigt und in einer Druckflasche insgesamt 72 Stunden auf 100° erhitzt, jedoch so, dass die Lösung jede 18 Stunden gekühlt und wieder mit Acetylen gesättigt wurde. Der grösste Teil des Acetons und des Alkohols wurde im Vakuum mit einer Kolonne abdestilliert. Das Destillat enthielt grosse Mengen nicht umgesetzter Stickstoffwasserstoffsäure. Der Rückstand wurde wiederholt mit Äther ausgezogen, die ätherische Lösung wurde mit Na_2SO_4 getrocknet und destilliert. Das über 200° übergehende Produkt wurde mit Hilfe einer kleinen Kolonne fraktioniert. $Kp_{-765} : 205^\circ$. Schmp. $21-22^\circ$. Ausbeute 1,1 g = 10% im Verhältnis zur benutzten Menge Natriumazid.

1-Phenyl-1,2,3-Triazol. Wurde nach DIMROTH und FESTER² aus Phenylazid und Acetylen dargestellt. Umkristallisiert aus Wasser. Schmp. 56° .

2-Phenyl-1,2,3-Triazol. Wurde nach v. PECHMANN³ aus Glyoxalosazon dargestellt. $Kp.$ $227-228^\circ$ bei 767 mm.

1,2,4-Triazol. Dargestellt aus Formamid und Formhydrazid⁴. Umkristallisiert aus Chloroform. Schmp. 121° .

1-Phenyl-1,2,4-Triazol. Dargestellt aus Formylphenylhydrazin und Formamid nach PELLIZZARI⁵. Umkristallisiert aus Petroläther. Schmp. $46,8-47^\circ$.

4-Phenyl-1,2,4-Triazol. Dargestellt aus Formylhydrazin und Formanilid nach PELLIZZARI⁶. Umkristallisiert aus Benzol. Schmp. $121-122^\circ$.

Tetrazol. Zur Darstellung von Tetrazol sind in der Literatur mehrere Methoden angegeben worden. Es wurden folgende Methoden nachgeprüft: Diazotierung von 5-Aminotetrazol und Re-

¹ Inorganic Syntheses, herausg. v. H. S. BOOTH (New York 1939), Bd. I, S. 78.

² O. DIMROTH u. G. FESTER, Ber. dtsh. chem. Ges. 43 (1910), 2222.

³ H. v. PECHMANN, Liebigs Ann. d. Chem. 262 (1891), 290, und Ber. dtsh. chem. Ges. 21 (1888), 2768.

⁴ G. PELLIZZARI, Gazz. chim. ital. 24, II (1894), 225.

⁵ G. PELLIZZARI, Gazz. chim. ital. 24, II (1894), 227.

⁶ G. PELLIZZARI, Atti Rend. Reale Accad. Lincei, Roma. [5] 10 (1901), 366.

duktion der Diazoverbindung mit Alkohol (THIELE¹), oder mit Natriumhypophosphit (STOLLÉ²), Umwandlung der Diazoverbindung in 5-Jodtetrazol und Reduktion desselben mittels schwefliger Säure (STOLLÉ³) und ferner die Reaktion zwischen Azimid und Cyanwasserstoff (DIMROTH⁴). Es wurden aber nach sämtlichen Methoden sehr schlechte Ausbeuten erhalten. THIELE gibt keine Ausbeute an, und die von STOLLÉ und DIMROTH angegebenen Ausbeuten konnten bei weitem nicht erreicht werden. Es wurden mehrere Versuche angestellt um die Methoden von THIELE (z. B. durch Diazotierung mittels Amylnitrit in Alkohol) und STOLLÉ zu verbessern, aber ohne Erfolg.

Das für die Messung verwendete Tetrazol war nach THIELE dargestellt worden. Die Diazotierung geschah jedoch nach BÜLOW⁵. Ausbeute 21 0/0. Reinigung durch Umkristallisation aus Dioxan und Sublimation im Vakuum (1 mm 125—150°). Schmp. 156°.

1-Methyltetrazol. Darstellung von 1-Methyl-5-mercaptotetrazol aus Methylsenföl und Natriumazid und Oxydation des Mercaptotetrazols mittels Wasserstoffperoxyd⁶. Umkristallisation aus Alkohol. Schmp. 36°.

Isoxazol. Wurde aus Propargylaldehyd und Hydroxylaminhydrochlorid dargestellt⁷, und über das Cadmiumchloriddoppelsalz gereinigt. Kp. ₇₆₀: 95—96°.

Thiazol. Zur Darstellung von Thiazol liegt in der Literatur keine befriedigende Methode vor. Nachdem wir eine verbesserte Methode zur Darstellung von 2-Chlorthiazol ausgearbeitet hatten, wurde das Thiazol mit vorzüglicher Ausbeute durch Reduktion des 2-Chlorthiazols dargestellt: 31 g 2-Chlorthiazol werden in 100 ccm 50 0/0 Essigsäure gelöst; unter Umschütteln und Kühlung werden 25 g Zinkstaub portionsweise zugesetzt. Das Gemisch wird 4-5-Stunden bei Zimmertemperatur geschüttelt und dann 5 Minuten auf 100° erhitzt. Die Lösung wird mit Natronlauge (40 0/0) schwach alkalisch gemacht und mit Wasserdampf destilliert. Ausbeute 19,5 g = 90 0/0. Das Thiazol wurde durch De-

¹ J. THIELE u. H. INGLE, Liebigs Ann. d. Chem. **287** (1895), 244.

² R. STOLLÉ, Ber. dtsch. chem. Ges. **62** (1929), 1122.

³ R. STOLLÉ, Ber. dtsch. chem. Ges. **62** (1929), 1124.

⁴ O. DIMROTH u. G. FESTER, Ber. dtsch. chem. Ges. **43** (1910), 2223.

⁵ C. BÜLOW, Ber. dtsch. chem. Ges. **42** (1909), 4436.

⁶ R. STOLLÉ u. F. HENCKE-STARK, Journ. prakt. Chem. [2] **124** (1930), 281.

⁷ W. PELTZ, Beiträge zur Kenntnis des Propargylaldehyds und des aus ihm gewonnenen Isoxazols. Dissertation Kiel 1902. S. 23.

stillation durch eine Widmer-Kolonne (Trennungsvermögen etwa 2 theoretische Böden) gereinigt. Kp. ₇₄₉: 119,5—120°.

2-Chlorthiazol. Die Darstellung von 2-Chlorthiazol geschieht nach SCHATZMANN¹ durch Diazotierung von 2-Aminothiazol in stark salzsaurer Lösung. Ohne Zusatz von Katalysatoren findet dann eine Umwandlung des sehr unbeständigen Diazoniumsalzes in 2-Chlorthiazol statt. Gleichzeitig findet aber auch eine weitergehende Zersetzung statt, so dass die Ausbeute nur etwa 10% beträgt. Es wurde versucht, durch den bei der Darstellung von Halogenbenzolen üblichen Zusatz von Kupfer (I)-chlorid die Ausbeute zu verbessern, aber ohne Erfolg. Dagegen wurde durch Zusatz von Kupferbronze (»Naturkupfer C«) eine wesentliche Verbesserung der Ausbeute erreicht. Beim Zusatz des Kupferpulvers findet unter beträchtlicher Wärmeentwicklung eine lebhaft Stickstoffentwicklung statt. Wenn energisch gekühlt wurde, konnten Ausbeuten bis zu 40% erreicht werden. Eine weitere Erhöhung der Ausbeute war allerdings nicht möglich.

Es wurde in ähnlicher Weise versucht, die Ausbeute an 2-Bromthiazol — das bereits nach den üblichen Methoden in wesentlich besserer Ausbeute als das Chlorthiazol, nämlich zu etwa 40%, erhalten wird — zu verbessern. Das Diazoniumbromid ist jedoch weit unbeständiger als das Chlorid und zersetzt sich bereits während der Diazotierung, so dass nachheriger Zusatz von Naturkupfer ohne Einfluss ist. Selbst wenn bei —30° diazotiert wurde, war der Zusatz von Naturkupfer ohne Wirkung. Auch durch Zusatz von Naturkupfer während der Diazotierung konnte keine Verbesserung der Ausbeute erreicht werden.

Die Herstellung des 2-Chlorthiazols wurde folgendermassen vorgenommen: 100 g 2-Aminothiazol werden in 500 ccm konz. Salzsäure gelöst; unter mechanischer Rührung und Kühlung mittels eines Kältegemisches wird eine Lösung von 70 g Natriumnitrit in 70 ccm Wasser zugetropft. Die Temperatur wird bei —5° bis 0° gehalten; der Zusatz dauert dann etwa eine Stunde. Jetzt wird vorsichtig Naturkupfer in kleinen Portionen zugesetzt. Es findet heftige Stickstoffentwicklung und Temperaturerhöhung statt, man sorgt durch Einwerfen von Eisstücken dafür, dass die Temperatur nicht über 10° steigt (äussere Kühlung genügt

¹ P. SCHATZMANN, Liebigs Ann. d. Chem. **261** (1891), 10.

nicht). Der Zusatz von Naturkupfer wird fortgesetzt bis keine weitere Temperaturerhöhung stattfindet und die Stickstoffentwicklung aufgehört hat. Insgesamt werden etwa 4 g zugesetzt. Es wird jetzt 15 Minuten bei Zimmertemperatur gerührt, dann schwach alkalisch gemacht und mit Wasserdampf destilliert. Das Destillat wird mit Kaliumcarbonat versetzt und mit Äther ausgezogen. Die ätherische Lösung wird mit Kaliumcarbonat getrocknet und unter Anwendung einer Widmer-Kolonne fraktioniert. Ausbeute 20—25 g mit Kp. 145—47° = 33—41%.

Benzpyrazol (Indazol). Dargestellt aus N-Nitroso-benz-o-to-luidid¹. Umkristallisiert aus Wasser. Schmp. 147—148°.

Benzimidazol. Dargestellt nach WUNDT² aus o-Phenylendiamin und Ameisensäure. Umkristallisation aus Wasser. Schmp. 170—171°.

Benztriazol. Dargestellt durch Diazotierung von o-Phenylendiamin³. Reinigung durch Vakuumdestillation und zweimalige Umkristallisation aus Benzol. Schmp. 99°.

Benzoxazol. Darstellung nach v. NIEMENTOWSKI⁴ aus o-Aminophenol und Formamid. Die Isolierung des Benzoxazols geschah jedoch durch Ausziehen mit Äther. Die ätherische Lösung wurde mit $\frac{n}{10}$ -Salzsäure und dann mit Wasser geschüttelt, über Kaliumcarbonat getrocknet und im Vakuum destilliert. Kp.₁₇: 76—77°. Schmp. 30,5°.

α,β -Benzisoxazol (Indoxazen). LINDEMANN und THIELE⁵ erhielten diese Verbindung beim Erhitzen von N-Acetylsalicylaldoxim. Nach der Vorschrift soll eine lebhafte Reaktion bei 130° einsetzen. Diese wurde bei unseren Versuchen nicht beobachtet; bei etwa 170° (Badtemperatur) und 20 mm wurde jedoch ein Destillat erhalten, das im wesentlichen aus Indoxazen bestand. Zur Reinigung wurde die Verbindung in Äther aufgenommen und mehrmals mit $\frac{n}{1}$ -NaOH ausgezogen. Nach dem Trocknen über Na₂SO₄ wurde die ätherische Lösung unter Anwendung einer Widmer-Kolonne destilliert. Kp.₂₄: 94°.

¹ P. JACOBSEN u. L. HUBER, Ber. dtsch. chem. Ges. **41** (1908), 662.

² E. WUNDT, Ber. dtsch. chem. Ges. **11** (1878), 826.

³ Organic Syntheses **20** (1940), 16.

⁴ St. v. NIEMENTOWSKI, Ber. dtsch. chem. Ges. **30** (1897), 3064.

⁵ H. LINDEMANN u. H. THIELE, Liebigs Ann. d. Chem. **449** (1926), 75.

β,γ -Benzisoxazol (Anthranil). Dargestellt aus o-Nitrobenzaldehyd durch Reduktion mit Zinkstaub in Eisessig¹. Reinigung durch Fällung der HgCl₂-Doppelverbindung; diese wurde mittels KCl zersetzt, das Anthranil wurde mit Wasserdampf übergetrieben, ausgeäthert und im Vakuum destilliert. Kp.₁₅: 100°.

3-Methyl-pyrazolon-(5). Darstellung aus Acetessigester und Hydrazinhydrat². Umkristallisation aus Wasser. Schmp. 219°.

1-Phenyl-2,3-dimethyl-pyrazolon-(5) (Antipyrin). Ein käufliches Präparat (Ph. dan.) wurde mehrmals aus Benzol umkristallisiert. Schmp. 111°.

1-Phenyl-2,3-dimethyl-pyrazolthion-(5) (Thiopyrin). Aus Antipyrin nach MICHAELIS³. Umkristallisation aus Wasser und dann aus Benzol. Schmp. 166°.

1,2,4-Triazolon-(5). Dargestellt aus Acetonsemicarbazon und Ameisensäure⁴. Umkristallisation aus Alkohol, Schmp. 234°.

Dimethylfuroxan. Darstellung aus Dimethylglyoxim und NO₂⁵ in trockenem Äther. Reinigung durch Destillation im Vakuum. Kp.₁₈: 110—111°.

2,4-Dioxothiazolidin (Senfölessigsäure). Dargestellt aus Chloressigsäure und Thioharnstoff⁶. Umkristallisiert aus Wasser. Schmp. 125°.

2,4-Thio-oxo-thiazolidin (Rhodanin). Darstellung aus Chlor-essigsäure und Ammoniumrhodanid⁷. Umkristallisiert aus Wasser. Schmp. 169—170°.

5-Aminotetrazol. Dargestellt nach THIELE⁸ (statt Aminoguanidinnitrat wurden jedoch die leichter zugänglichen Salze, Aminoguanidinsulfat⁹ und Aminoguanidinbicarbonat¹⁰ verwendet) oder nach STOLLÉ¹¹ aus Dicyandiamid und Azoimid. Das Monohydrat wurde aus absolutem Alkohol umkristallisiert und bei 150° getrocknet. Schmp. 203—204°.

¹ P. FRIEDLÄNDER u. R. HENRIQUES, Ber. dtsh. chem. Ges. **15** (1882), 2105.

² Th. CURTIUS u. R. JAY, Journ. prakt. Chem. [2] **39** (1889), 52.

³ A. MICHAELIS u. H. BINDEWALD, Liebigs Ann. d. Chem. **320** (1901), 4.

⁴ O. WIDMAN u. A. CLEVE, Ber. dtsh. chem. Ges. **31** (1898), 379.

⁵ R. SCHOLL, Ber. dtsh. chem. Ges. **23** (1890), 3499.

⁶ J. VOLHARD, Journ. prakt. Chem. [2] **9** (1874), 9.

⁷ M. NENCKI, Journ. prakt. Chem. [2] **16** (1877), 2.

⁸ J. THIELE, Liebigs Ann. d. Chem. **270** (1892), 46 und 55.

⁹ G. B. L. SMITH u. E. ANZELMI, Journ. Amer. Chem. Soc. **57** (1935), 2730.

¹⁰ P. FANTL u. H. SILBERMANN, Liebigs Ann. d. Chem. **467** (1928), 279.

¹¹ R. STOLLÉ u. E. SCHICK, Ber. dtsh. chem. Ges. **62** (1929), 1120. Dtsch. Reichs-Pat. 426343.

2-Aminothiazol. Darstellung aus Chloracetal und Thioharnstoff¹. Umkristallisiert aus Wasser und dann aus Äther. Schmp. 90°.

»**1,4-Diphenyl-5-methyl-3,5-endothio-1,2,4-triazolin**«. Beim Behandeln von 3-Phenyl-2-chlor-2-methyl-1,3,4-thiazolidon-(5)-anil mit alkoholischem Ammoniak². Umkristallisiert aus verdünntem Alkohol (75 %). Schmp. 247°.

3-Methylmercapto-5-methyl-1,4-diphenyl-1,2,4-triazoliniumjodid. Aus der vorhergehenden Verbindung und Methyljodid³. Umkristallisiert aus Methanol-Äther. Schmp. 250°. Molekulare Leitfähigkeit in Äthanol bei 25°: $\mathcal{A}_\infty = 34,5$. Zum Vergleich diene, dass der entsprechende Wert für NaJ 47,8 beträgt (Landolt-Börnstein). Die Verbindung verhält sich somit in Lösung wie ein binärer Elektrolyt.

»**1,4,5-Triphenyl-3,5-endoanilo-1,2,4-triazolin**«. Dargestellt aus Triphenylaminoguanidin und Benzoylchlorid⁴. Umkristallisiert aus Benzol. Schmp. 234°.

»**1,4-Diphenyl-5-methylmercapto-3,5-endooxy-1,2,4-triazolin**«. Dargestellt aus 2,4-Diphenyl-S-methylthiosemicarbazid und Phosgen⁵. Umkristallisiert aus Benzol. Schmp. 185°.

1,4-Diphenyl-3,5-dioxo-1,2,4-triazolidin. Dargestellt durch Hydrolyse der vorstehenden Verbindung.⁶ Umkristallisiert aus absolutem Alkohol. Schmp. 165°.

»**Diphenylisothiobiazolon**« = »**2,3-Diphenyl-2,5-endothio-1,3,4-thiodiazolin**«. Dargestellt nach BUSCH⁷ aus dem Kaliumsalz der Phenylidithiocarbazinsäure und Benzoylchlorid. Umkristallisiert aus Benzol. Schmp. 226—227°.

5-Methylmercapto-2,3-diphenyl-1,3,4-thiodiazoliniumjodid. Aus vorstehender Verbindung und Methyljodid⁸. Umkristallisiert aus Chloroform-Äther. Schmp. 188—189°. Molekulare Leitfähigkeit in Wasser bei 25°: $\mathcal{A}_\infty = 93,5$. Der entsprechende Wert für NaJ beträgt 127,0 (Landolt-Börnstein).

¹ Dän. Pat. 59382 (1941).

² M. BUSCH u. S. SCHNEIDER, Journ. prakt. Chem. [2] **67** (1903), 253.

³ M. BUSCH u. S. SCHNEIDER, Journ. prakt. Chem. [2] **67** (1903), 255.

⁴ M. BUSCH u. G. MEHRTENS, Ber. dtsch. chem. Ges. **38** (1905), 4064.

⁵ M. BUSCH u. H. HOLZMANN, Ber. dtsch. chem. Ges. **34** (1901), 340.

⁶ M. BUSCH u. H. HOLZMANN, Ber. dtsch. chem. Ges. **34** (1901), 341.

⁷ M. BUSCH, Journ. prakt. Chem. [2] **60** (1899), 218.

⁸ M. BUSCH, W. KAMPHAUSEN u. S. SCHNEIDER, Journ. prakt. Chem. [2] **67** (1903), 222.

2-Äthoxy-5-methylmercapto-2,3-diphenyl-1,3,4-thiodiazolin.

Aus dem Jodid und Natriumäthylat.¹ Umkristallisiert aus Petroläther. Schmp. 100°. Trotz wiederholter Umkristallisationen konnte der von BUSCH angegebene Schmp. 106° nicht erreicht werden. Nach einer Stickstoffbestimmung (Mikro-Dumas) war die Verbindung rein. Möglicherweise ist 106° ein Druckfehler für 100°. Gef. 8,71 und 8,69 % N; ber. 8,48 %.

»3-Phenyl-2-methyl-2,5-endothio-1,3,4-thiodiazolin«. Darstellung aus dem Kaliumsalz der Phenyldithiocarbaminsäure und Acetylchlorid². Umkristallisiert aus einem Gemisch von Chloroform und Alkohol (1:1). Schmp. 220°.

2. Lösungsmittel.

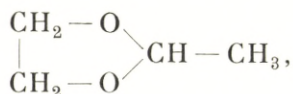
Benzol. »Benzol zur Analyse und Molekulargewichtsbestimmung, Kahlbaum« wurde über Natrium getrocknet. Für dieses Präparat wurden, übereinstimmend mit den besten Literaturwerten, folgende Konstanten gefunden:

Spezifisches Gewicht, $d_4^{25} = 0,8737$.

Brechungsexponenten: $n_D^{25} = 1,49801$ und $n_\infty^{25} = 1,4737$.

Die Dielektrizitätskonstante des Benzoles wurde nach HARTSHORN und OLIVER³ zu 2,2725 bei 25° gesetzt. Alle anderen Dielektrizitätskonstanten sind im Verhältnis zu der des Benzols bestimmt.

Dioxan. Die Dioxanpräparate des Handels sind alle stark verunreinigt; hauptsächlich durch »Glykolacetal«



ausserdem kommen aber auch Peroxyde und andere Verbindungen vor. Die Verunreinigungen können zum Teil durch anhaltendes Kochen mit Natrium beseitigt werden; zur völligen Reinigung ist aber auch eine Behandlung mit Salzsäure notwendig⁴. Da aber eine sehr effektive Kolonne (25—30 theoretische Böden)

¹ M. BUSCH, W. KAMPHAUSEN u. S. SCHNEIDER, Journ. prakt. Chem. [2] **67** (1903), 224.

² M. BUSCH u. S. SCHNEIDER, Journ. prakt. Chem. [2] **67** (1903), 250.

³ L. HARTSHORN u. D. A. OLIVER, Proc. Roy. Soc. London (A), **123** (1929), 664.

⁴ K. HESS u. H. FRAHM, Ber. dtsch. chem. Ges. **71** (1938), 2627.

zu unserer Verfügung stand¹, wurde versucht das Dioxan durch Destillation zu reinigen.

1 kg Dioxan »Kahlbaum« wurde 24 Stunden über KOH, dann über Silicagel getrocknet und danach destilliert (Rückflussverhältnis 1 : 33, Destillationsgeschwindigkeit: 1 Tropfen jede 30 Sek.). Es wurden ein Vorlauf von 82,5—88,1° (75 ccm), eine Mittelfraktion (5 ccm) von 88,1—101,4° (764 mm) und die Hauptfraktion (650 ccm) bei 101,4° (764 mm) und 101,8° (771 mm) aufgesammelt. Für die verschiedenen Fraktionen wurden folgende Brechungsexponenten gefunden:

	n_D^{25}
Vorlauf 1	1,40677
« 2—4	1,40719
Mittelfraktion 1 (88,1—101,4°)...	1,41034
« 2 (101,4°)	1,41940
« 3 «	1,41970
Hauptfraktion 1—4 (101,4°)	1,41990
Nachlauf	1,41990

Sowohl der Vorlauf wie der Rückstand reagierten heftig mit Natrium. Der Vorlauf ist hauptsächlich Glykolacetal ($n_D = 1,3955$).

Der Brechungsexponent und die Dichte $d_4^{25} = 1,02785$ der Hauptfraktion stimmen mit den zuverlässigsten Werten der Literatur gut überein. Der Brechungsexponent ist jedoch kein genügendes Kriterium auf Reinheit, denn die Dielektrizitätskonstanten der Hauptfraktion und des Nachlaufes waren sehr verschieden (2,2137 und 2,2688), obwohl die Brechungsexponenten identisch waren. Auch die Dielektrizitätskonstanten der Hauptfraktion waren nicht ganz dieselben (2,2137—2,2148). Die Hauptfraktion wurde über Natrium destilliert und über Natriumdraht aufbewahrt. Die Dielektrizitätskonstante dieses Präparates (I) war 2,2178. Für einige der Dipolmessungen wurde jedoch ein anderes Präparat (II) angewandt, das intensiver mit Natrium behandelt war, und das die Dielektrizitätskonstante 2,2075 aufwies. Dieser Unterschied ist für die Dipolmessung belanglos, zeigt aber, dass die Darstellung eines völlig reinen Dioxan-Präparates ausserordentlich schwierig ist.

¹ Herrn Civilingenieur, Dr. phil. A. Klit sei für die Überlassung der Kolonne und für wertvolle Ratschläge bezüglich der Ausführung der Destillation bestens gedankt.

Die in der Literatur angegebenen Werte für die Dielektrizitätskonstante des Dioxans schwanken sehr stark (der richtige Wert liegt jedoch wahrscheinlich ein wenig unter 2,210):

	ϵ^{25}	d_4^{25}	n_D^{25}
FREY und GILBERT ¹	2,3051	1,0268	1,42025
SMYTH und WALLS ²	2,306	1,0311	1,42032
ULICH und NESPITAL ³	2,233		
CALDWELL und LE FÈVRE ⁴	2,2093	1,0292	
DOSTROVSKY und LE FÈVRE ⁵	2,2106	2,0277	
» » ⁵	2,2212	1,0281	
K. A. JENSEN und A. BERG ⁶	2,2198		
Diese Untersuchung.....	2,2178	1,02785	1,41990
» ».....	2,2075	1,02785	1,41990

Chloroform. Ein Ph. dan.-Präparat wurde mit Wasser und verd. KOH gewaschen, über CaCl_2 getrocknet und fraktioniert. $\epsilon^{25} = 4,6775$ (LE FÈVRE: 4,724⁷); $d_4^{25} = 1,47998$; $n_D^{25} = 1,44316$ (HUBBARD: 1,44309⁸).

3. Dipolmessungen.

Zur Messung der Dielektrizitätskonstanten wurde der Präzisionsapparat von KIPP und ZONEN⁹ benutzt. Dieser besteht im Prinzip aus einer quarzkristallgesteuerten Hochfrequenzpentode (Frequenz: $1,503 \cdot 10^6$ Hertz) mit variablem Anodenschwingungskreis. Im Anodenschwingungskreis befindet sich ein mit der zu untersuchenden Substanz gefüllter Messkondensator (Messzelle) und ein variabler justierter Präzisionskondensator mit einer Totkapazität von 5 cm im Parallelschaltung. Die Resonanzeinstellung geschieht visuell durch Messung des Anodenstromes; bei einem bestimmten, kritischen Wert der totalen Kapazität des Anodenkreises ist dieser mit dem Quarzkristall in Resonanz, und es findet eine plötzliche Änderung des Anodenstromes statt (Abb. 1.), die mittels eines Milliampereometers angezeigt wird. Wenn man das Milliampereometer auf einen Punkt zwischen

¹ P. R. FREY u. E. C. GILBERT, Journ. Amer. Chem. Soc. **59** (1937), 1344.

² C. P. SMYTH u. W. S. WALLS, Journ. Amer. Chem. Soc. **53** (1931), 2117.

³ H. ULICH u. W. NESPITAL, Zeitschr. phys. Chem. Abt. B **16** (1932), 229.

⁴ C. C. CALDWELL u. R. J. W. LE FÈVRE, Journ. Chem. Soc. **1939**, 1621.

⁵ I. DOSTROVSKY u. R. J. W. LE FÈVRE, Journ. Chem. Soc. **1939**, 535.

⁶ K. A. JENSEN u. A. BERG, Liebigs Ann. d. Chem. **548** (1941), 116.

⁷ R. J. W. LE FÈVRE, Dipole Moments (London 1938) Seite 2.

⁸ J. C. HUBBARD, Zeitschr. phys. Chem. Abt. A. **74** (1910), 222.

⁹ Vgl. P. COHEN-HENRIQUES, Rec. trav. chim. Pays-bas **54** (1935), 327.

A und B einstellt, entspricht dies einer genau festgelegten Kapazität im Kreise. Bei der Messung wird die Messzelle erst mit Benzol gefüllt und der Anodenstrom durch Einstellung eines verstellbaren Kondensators, der parallel zur Messzelle geschaltet ist (Kapazität ca. 250 cm), auf einen Punkt zwischen A und B festgelegt. Die Messzelle wird dann mit der zu untersuchenden Lösung gefüllt und der Präzisionskondensator (Zylinderkondensator) verstellt, bis das Milliampereometer wieder denselben Anodenstrom anzeigt. Die auf dem Präzisionskondensator abge-

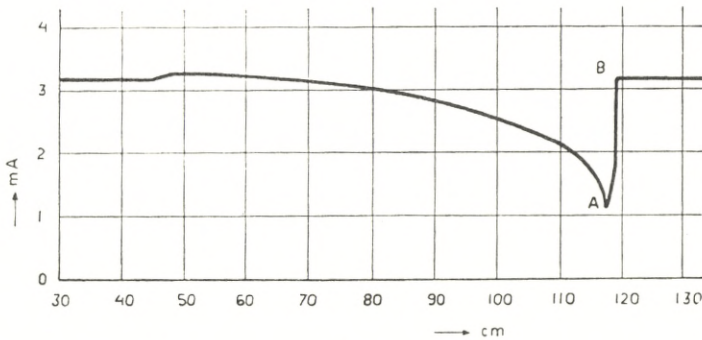


Fig. 1.

lesene Kapazitätsänderung sei ΔC und die Kapazität der Messzelle C ; $\Delta\epsilon$ ist dann gleich $\Delta C/C$, und die Dielektrizitätskonstante der Lösung ist somit bestimmt, da ϵ^{25} für Benzol gleich 2,2725 gesetzt wurde.

Der Präzisionskondensator ist in 0,001 cm eingeteilt und sollte nach den Angaben der Hersteller streng linear sein. Dies war jedoch nur für $\frac{2}{3}$ der Skala genau der Fall; so weit möglich wurde nur dieser Teil der Skala gebraucht, sonst wurden die abgelesenen Zahlen korrigiert.

Es wurden — je nach der Grösse von ΔC — zwei Messzellen verwendet, deren Kapazität zu 9,994 und 1,363 cm bestimmt worden war, und deren Inhalt etwa 2 ccm betrug. Die Messzelle befand sich in einem elektrisch geheizten Ölbad, dessen Temperatur mit Hilfe eines Kontaktthermometers auf $25,0^\circ \pm 0,1$ gehalten wurde.

Aus den so gemessenen Dielektrizitätskonstanten wurde in üblicher Weise die Molpolarisation der gelösten Substanz berechnet. Die meisten der Verbindungen wurden bei verschiedenen

Konzentrationen gemessen, und die entsprechenden Molpolarisationen wurden graphisch auf unendliche Verdünnung extrapoliert.

Die Löslichkeit der Endo-Verbindungen und einiger der anderen untersuchten Verbindungen ist so gering, dass auf der Mikrowage abgewogen werden musste, um eine genügend genaue Konzentrationsbestimmung zu erreichen.

Als Lösungsmittel wurde so weit möglich Benzol benutzt; viele der Verbindungen waren jedoch zu schwer löslich in Benzol, sie wurden in Dioxan und — in einzelnen Fällen — in Chloroform gemessen. Chloroform ist allerdings kein dipolfreies Lösungsmittel, es hat aber eine Polarisation, die praktisch unabhängig von Konzentration und Zustand ist¹, so dass sein Beitrag zur Polarisation der Lösung mit Hilfe der Mischungsregel berechnet werden kann. Nach den bekannten experimentellen Daten² führt die Messung von Dipolmomenten in Chloroform-Lösung gewöhnlich zu korrekten Vorstellungen über die Molekülstruktur, wenn auch die rein numerischen Werte — ebenso wie die in Dioxan gemessenen — von denen in benzolischer Lösung oder im Gaszustand abweichen können. Wegen der Möglichkeit der Bildung von Molekülverbindungen zwischen Chloroform und dem gelösten Stoff, sowie wegen der unbekannt polarisierenden Wirkung auf die gelösten Moleküle, ist dieses Lösungsmittel jedoch nur mit grosser Vorsicht zu benutzen.

Einige der Endo-Verbindungen wurden sowohl in Benzol als auch in Dioxan gemessen, und es wurden grössenordnungsmässig übereinstimmende Werte erhalten. Vollkommene Übereinstimmung ist nicht zu erwarten; denn es ist bekannt, dass in Dioxanlösung bestimmte Dipolmomente gelegentlich beträchtlich von den in benzolischer Lösung bestimmten abweichen können. (Vgl. HÜCKEL u. a.^{3,4}).

Die Dichten der Lösungen wurden anfangs mittels eines Pyknometers nach SPRENGEL-OSTWALD bestimmt; später wurde ein Flaschenpyknometer benutzt, das bequemer zu handhaben war und eine Genauigkeit von 0,0001 gewährleistete, die für unseren Zweck hinreichend war.

¹ C. G. LE FÈVRE u. R. J. W. LE FÈVRE, Journ. Chem. Soc. **1935**, 1747.

² R. J. W. LE FÈVRE u. G. J. RAYNES, Journ. Chem. Soc. **1938**, 1921.

³ W. HÜCKEL, J. DATOW u. E. SIMMERSBACH, Zeitschr. phys. Chem. Abt. A. **186** (1940), 166.

⁴ W. WASSILIEW u. J. SSYRKIN, Acta physico-chim. USSR **14** (1941), 414. (vgl. Chem. Zentr. 1941 II, 1500).

Zur Bestimmung von P_E wurden in einigen Fällen die Brechungsexponenten der Lösungen bestimmt. Für mehrere der Verbindungen lagen jedoch in der Literatur genaue Messungen der Molrefraktion für die α -, γ - und D-Linie vor; aus diesen Werten wurde dann $R_\infty = P_E$ berechnet (vgl. unten). Für die Endo-Verbindungen und für einige der anderen Verbindungen, deren Dipolmomente so gross sind, dass die P_E -Werte nur eine kleine Korrektur bedeuten, wurde P_E gleich R_D gesetzt und dieser Wert durch Addition der Atomrefraktionen berechnet, oder durch Messung der Brechungsexponenten einer Lösung für die D-Linie bestimmt.

Die Brechungsexponenten wurden mittels eines Refraktometers nach PULFRICH bestimmt. Die Erwärmung auf 25° geschah durch einen Wasserstrom aus einem Wasserthermostaten, dessen Temperatur auf $25,00 \pm 0,01^\circ$ gehalten wurde und der auch für die Bestimmung der spezifischen Gewichte benutzt wurde. Lichtquelle war ein Heliumrohr. Es wurden die Brechungsexponenten für die rote, gelbe und grüne Linie bestimmt und nach der vereinfachten Formel von CAUCHY $n_\lambda = n_\infty + a/\lambda^2$ auf unendlich grosse Wellenlänge extrapoliert.

Es bedeutet in den folgende Tabellen; x_1 : Molenbruch des gelösten Stoffes; ϵ^{25} : Dielektrizitätskonstante der Lösung bei 25° ; d_4^{25} : spezifisches Gewicht der Lösung bei 25° , bezogen auf Wasser von 4° C.; P : Molpolarisation des Gelösten; $P - P_E = P_{A+O}$: Atom- + Orientierungspolarisation des Gelösten; P_O^∞ : Orientierungspolarisation, auf unendliche Verdünnung extrapoliert; μ' : Dipolmoment ohne Korrektur für Atompolarisation; μ : Dipolmoment nach $P_O = P - P_E - 0,15 P_E$ indem die Atompolarisation nach K. L. WOLF¹ zu 15 0/0 der Elektronenpolarisation gesetzt wurde. Die Dipolmomente sind in Debye-Einheiten ($1 D = 10^{-18}$ e. s. E.) angegeben.

Fehlerberechnung: Durch partielle Differentiation der Gleichung:

$$\frac{\epsilon_{1,2} - 1}{\epsilon_{1,2} + 2} \cdot \frac{M_1 x_1 + M_2 (1 - x_1)}{d_{1,2}} = P_1 x_1 + P_2 (1 - x_1),$$

wo die Zahlen 1, 2 und 1,2 sich auf gelösten Stoff, Lösungsmittel und Lösung beziehen, bekommt man:

¹ K. L. WOLF, Phys. Ztschr. **31** (1930), 227.

$$\frac{\partial P_1}{P_1} = - \frac{P_{1,2}}{P_1} \cdot \frac{1}{x_1} \cdot \frac{\partial d_{1,2}}{d_{1,2}} \quad \text{und}$$

$$\frac{\partial P_1}{P_1} = \frac{3 \varepsilon_{1,2}}{\varepsilon_{1,2}^2 + \varepsilon_{1,2} - 2} \cdot \frac{P_{1,2}}{P_1} \cdot \frac{1}{x_1} \cdot \frac{\partial \varepsilon}{\varepsilon}.$$

Die Dielektrizitätskonstanten und die Dichten sind je mit einer Genauigkeit von etwa $1/4$ 0/0 bestimmt. Die Grösse von $\frac{\partial P_1}{P_1}$ variiert demnach bei unseren Versuchen zwischen $1/2$ 0/0 und 1 0/0, abhängig von der Konzentration und der Grösse von P_1 . Verbindungen mit den kleinsten P_1 -Werten sind bei relativ grosser Konzentration bestimmt, und der von der Bestimmung von ε und d herrührende Fehler ist deshalb nur klein (etwa 1 0/0); dafür wird die Genauigkeit durch die Unsicherheit der Korrektur für die Atompolarisation stark beeinträchtigt, so dass die Genauigkeit der Bestimmung von P_0 kaum grösser als 1 0/0 ist. Diese Korrektur spielt für die Verbindungen mit den grösseren Dipolmomenten keine Rolle; diese Verbindungen konnten aber wegen Schwerlöslichkeit nur in sehr kleinen Konzentrationen untersucht werden, so dass der von der Bestimmung von ε und d herrührende Fehler gross ist (bis zu 1 0/0), und ausserdem ist die Extrapolation auf P_1^∞ ziemlich unsicher. Da:

$$\mu = 0,001273 \sqrt{(P_1^\infty - P_E - P_A) \cdot T}, \quad \text{wird} \quad \frac{\partial \mu}{\mu} \approx 1/2 \frac{\partial P_1}{P_1}.$$

Bei den Dipolmomenten unter 4—5 kann deshalb mit einer Genauigkeit von $1/2$ —1 0/0, bei den grösseren Dipolmomenten mit einer Genauigkeit von 1—2 0/0 gerechnet werden.

Pyrazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P	
0,04433	2,4170	0,8798	64,85	$P^\infty = 72,0$
0,05400	2,4456	0,8816	63,42	$P_E = 18,0^1$
0,06586	2,4705	0,8833	60,27	$P_{O+\Lambda}^\infty = 54,0; \mu' = 1,61$
				$P_O^\infty = 51,3; \mu = 1,57$

¹ Berechnet nach den Molrefraktionen von K. v. AUWERS u. W. KOHLHAAS, Liebigs Ann. d. Chem. **437** (1924), 49.

1-Methylpyrazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,01059	2,3512	0,8752	131,0
0,01890	2,4130	0,8761	129,8
0,02666	2,4725	0,8768	129,6
0,03836	2,5610	0,8781	127,7
0,04411	2,6045	0,8797	126,1

$$P^\infty = 133,5$$

$$P_E = 22,1^1$$

$$P_{O+A}^\infty = 111,4$$

$$\mu' = 2,32$$

$$P_O^\infty = 108,1$$

$$\mu = 2,28$$

3-Methylpyrazol (in Benzol).

x_1	ε^{25}	n_∞	d_4^{25}	P	P_E	P_{O+A}
0,05266	2,4324	1,4740	0,8813	66,41	22,82	43,59
0,06933	2,4743	1,4742	0,8828	64,60	22,99	41,61
0,07953	2,5042	1,4744	0,8844	64,22	22,98	41,24

$$P_{O+A}^\infty = 45,6 \quad \mu' = 1,48$$

$$P_O^\infty = 42,2 \quad \mu = 1,43$$

Imidazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P	P_E^2	P_O	μ
0,000233	2,2777	0,8738	336,3	17,0	319,3	3,93
0,001140	2,3043	0,8741	420,6	17,0	403,6	4,42
0,005951	2,5557	0,8752	670,3	17,0	653,3	5,62

$$P_O^\infty = 305 \quad \mu = 3,84$$

1,2,3-Triazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,00865	2,3132	0,8757	85,16
0,01449	2,3538	0,8772	96,79
0,02291	2,4140	0,8787	104,17
0,04344	2,5953	0,8822	117,76

$$P^\infty = 82$$

$$P_{E+A} \sim R_D = 17^2$$

$$P_O^\infty = 65$$

$$\mu = 1,77$$

¹ Berechnet nach K. v. AUWERS und W. KOHLHAAS, Liebigs Ann. d. Chemie 437 (1924), 49.

² Berechnet aus den Atomrefraktionen.

1-Phenyl-1,2,3-Triazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,000933	2,2980	0,8746	376,0
0,00260	2,3412	0,8754	366,8
0,00682	2,4523	0,8773	359,8

$P^\infty = 380$
 $P_{E+A} \sim R_D = 34,6^1$
 $P_O^\infty = 345,4$
 $\mu = 4,08$

2-Phenyl-1,2,3-Triazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,01731	2,2901	0,8799	53,59
0,04287	2,3146	0,8898	52,42
0,04715	2,3195	0,8913	52,68

$P^\infty = 54,5$
 $P_{E+A} \sim R_D = 34,6^1$
 $P_O^\infty = 19,9$
 $\mu = 0,97$

1,2,4-Triazol (in Dioxan I).

x_1	ε^{25}	d_4^{25}	P
0,00730	2,3261	1,0291	223,6
0,01687	2,4801	1,0307	225,7

$P^\infty = 225$
 $P_{A+E} \sim R_D = 16,7^1$
 $P_O^\infty = 208$
 $\mu = 3,17$

1-Phenyl-1,2,4-Triazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,00383	2,3169	0,8754	205,4
0,01395	2,4343	0,8802	200,3
0,02550	2,5705	0,8849	198,0

$P^\infty = 206,6$
 $P_{E+A} \sim R_D = 34,6^1$
 $P_O^\infty = 172,0$
 $\mu = 2,88$

4-Phenyl-1,2,4-Triazol (in Benzol).

x_1	ε^{25}	d_4^{25}	P
0,00295	2,4032	0,8758	660,1
0,00421	2,4572	0,8761	651,0
0,00572	2,5205	0,8768	635,9

$P^\infty = 690$
 $P_{E+A} \sim R_D = 34,6^1$
 $P_O^\infty = 655,4$
 $\mu = 5,63$

¹ Berechnet aus den Atomrefraktionen.

Tetrazol (in Dioxan II).

x_1	ϵ^{25}	d_4^{25}	P
0,00265	2,3070	1,0286	547,7
0,00562	2,4190	1,0294	532,9
0,00940	2,5620	1,0300	518,1

$P_{E+A}^\infty \approx R_D = 17^1$
 $P_O^\infty = 560$
 $P_O^\infty = 543$
 $\mu = 5,11$

1-Methyltetrazol (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,00227	2,3636	0,8746	(594,8)
0,00525	2,4732	0,8754	611,7
0,00919	2,6140	0,8763	526,0

$P_{E+A}^\infty \approx R_D = 20^1$
 $P_O^\infty = 620$
 $P_O^\infty = 600$
 $\mu = 5,38$

Isoxazol (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,0233	2,5353	0,8772	173,9
0,0473	2,8110	0,8806	165,9
0,0734	3,1265	0,8847	158,6

$P_E = 182,0$
 $P_E = 16,0^2$
 $P_{O+A}^\infty = 166 \quad \mu' = 2,83$
 $P_O^\infty = 163,6; \mu = 2,81$

Thiazol (in Benzol)

x_1	ϵ^{25}	n_∞	d_4^{25}	P	P_E	P_{O+A}
0,04294	2,4507	1,4750	0,88548	78,73	21,05	56,87
0,04646	2,4624	1,4751	0,88644	77,66	21,07	56,78
0,05233	2,4858	1,4754	0,88798	77,18	21,19	55,70
0,05633	2,5107	1,4754	0,88854	80,44	22,35	58,28
0,06811	2,5612	1,4759	0,89185	78,65	21,36	56,27
0,07034	2,5636	1,4759	0,89324	76,73	21,00	55,60
0,08450	2,6223	1,4765	0,89679	75,99	21,19	54,70

$P_{O+A}^\infty = 59,40 \quad \mu' = 1,69$

$P_O^\infty = 56,2 \quad \mu = 1,64$

¹ Berechnet aus den Atomrefraktionen.

² Berechnet nach K. v. AUWERS, Ber. dtsh. chem. Ges. 57 (1924), 467.

Indazol (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P	
0,00570	2,2995	0,8761	97,16	$P_{E+A}^\infty \approx R_D = 28^1$
0,00886	2,3152	0,8773	98,41	$P_O^\infty = 69,6$
0,01112	2,3255	0,8782	97,28	$\mu = 1,83$

Benzimidazol (in Dioxan II).

x_1	ϵ^{25}	d_4^{25}	P	
0,00461	2,3215	1,0300	373,4	$P_{E+A}^\infty \approx R_D = 30^1$
0,00576	2,3478	1,0304	364,7	$P_O^\infty = 320$
0,00632	2,3698	1,0307	382,3	$\mu = 3,93$
0,01005	2,4717	1,0313	384,2	
0,01069	2,4985	1,0314	395,7	
0,01391	2,5370	1,0319	347,0	
0,01467	2,5585	1,0323	346,7	

Benztriazol (in Dioxan II).

x_1	ϵ^{25}	d_4^{25}	P	
0,00706	2,3817	1,0309	367,0	$P_{E+A}^\infty \approx R_D = 27^1$
0,00960	2,4465	1,0318	365,5	$P_O^\infty = 3,44$
0,01469	2,5789	1,0329	362,6	$\mu = 4,07$

Benzoxazol (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P	
0,04848	2,4074	0,8908	75,67	$P_{E+A}^\infty = 79,2$
0,07894	2,4983	0,9018	73,81	$P_E = 30^3$
0,12492	2,6173	0,9172	70,64	$P_{O+A}^\infty = 49,2$ $\mu' = 1,54$
				$P_O^\infty = 44,7$ $\mu = 1,47$

¹ Berechnet aus den Atomrefraktionen.² Die P-c-Kurve scheint ein Maximum zu haben; die Extrapolation ist deshalb ziemlich unsicher.³ Berechnet nach K. v. AUWERS u. E. WALTRAUT, Zeitschr. phys. Chem. **122** (1926), 247.

Indoxazen (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,00773	2,3722	0,8766	214,8
0,02275	2,5720	0,8814	211,7
0,03360	2,7120	0,8855	204,9

$P^\infty = 226$
 $P_E = 31^1$
 $P_{O+A}^\infty = 195 \quad \mu' = 3,07$
 $P_O^\infty = 190,4; \mu = 3,03$

Anthranil (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,02404	2,5865	0,8828	221,4
0,09597	3,5953	0,9077	190,5
0,11155	3,8301	0,9131	186,0

$P^\infty = 230$
 $P_E = 31^2$
 $P_{O+A}^\infty = 199 \quad \mu' = 3,10$
 $P_O^\infty = 194,4; \mu = 3,06$

3-Methylpyrazolon-(5) (in Dioxan II).

x_1	ϵ^{25}	d_4^{25}	P
0,001303	2,2199	1,0284	154,6
0,004548	2,2481	1,0301	145,1

$P^\infty = 158$
 $P_{E+A} \sim R_D = 24^3$
 $P_O^\infty = 134$
 $\mu = 2,54$

Antipyrin (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,002274	2,3698	0,8753	658,8
0,005223	2,4929	0,8771	635,9
0,006773	2,5613	0,8781	631,9

$P^\infty = 666$
 $P_{E+A} \sim R_D = 45^3$
 $P_O^\infty = 621$
 $\mu = 5,48$

Thiopyrin (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,000791	2,3370	0,8747	1214,3
0,001168	2,3626	0,8749	1150,9
0,001784	2,4107	0,8755	1144,1

$P^\infty = 1162$
 $P_{E+A} \sim R_D = 49^3$
 $P_O^\infty = 1113$
 $\mu = 7,33$

¹ Berechnet nach K. v. AUWERS, Ber. deutsch. Ges. 57 (1924), 467.

² Berechnet nach K. v. AUWERS, Liebigs. Ann. d. Chem. 437 (1924), 84.

³ Berechnet aus den Atomrefraktionen.

1,2,4-Triazolol-(5) (in Dioxan II).

x_1	ϵ^{25}	d_4^{25}	A	$P_{E+A} \infty R_D = 19^1$	$P^\infty = 243$
0,000758	2,2181	1,0278	227,4		$P_O^\infty = 224$
0,002231	2,2432	1,0291	243,4		$\mu = 3,30$

Dimethylfuroxan (in Benzol).

x_1	ϵ^{25}	d_4^{25}	n_∞	P	P_E	P_{O+A}
0,01112	2,5832	0,8772	1,4738	412,5	28,2	384,3
0,03058	3,1397	0,8846	1,4739	375,0	26,8	348,2
0,04372	3,5131	0,8892	1,4740	351,9	26,8	325,1
0,06785	4,2066	0,8985	1,4741	316,9	26,5	290,4

$$P_{O+A}^\infty = 404,2 \quad \mu' = 4,42$$

$$P_O^\infty = 400,1 \quad \mu = 4,40$$

2,4-Dioxothiazolidin (in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P	$P_{E+A} \infty R_D = 23^1$	$P^\infty = 108$
0,00143	2,2252	1,0279	106,6		$P_O^\infty = 85$
0,01247	2,2705	1,0341	81,1		$\mu = 2,03$
0,02905	2,3329	1,0422	76,6		

Rhodanin (in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P	$P_{E+A} \infty R_D = 21,5^2$	$P^\infty = 122$
0,00388	2,2468	1,0302	130,3		$P_O^\infty = 101$
0,01337	2,3109	1,0356	121,8		$\mu = 2,20$
0,02337	2,3839	1,0412	122,4		

5-Aminotetrazol (in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P	$P_{E+A} \infty R_D = 20^1$	$P^\infty = 676,2$
0,001988	2,3124	1,0279	696,2		$\mu = 5,71$

¹ Berechnet aus den Atomrefraktionen.

² n_D der einen Lösung ist gemessen worden.

2-Aminothiazol (in Benzol).

x_1	ϵ^{25}	d_4^{25}	n_∞	P	P_E	P_{O+A}
0,00704	2,3050	0,8764	1,4742	90,27	24,81	65,46
0,01000	2,3175	0,8778	1,4747	87,16	24,99	62,17
0,01685	2,3459	0,8804	1,4749	84,99	24,69	60,30

$$P_{A+O}^\infty = 67,5 \quad \mu' = 1,81$$

$$P_O^\infty = 63,8 \quad \mu = 1,75$$

»1,4-Diphenyl-5-methyl-3,5-endothio-1,2,4-triazolin«
(in Chloroform).

x_1	ϵ^{25}	d_4^{25}	P
0,000418	4,7788	1,4800	1402
0,001202	4,9393	1,4799	1240
0,003946	5,3892	1,4792	996

$$P^\infty = 1525$$

$$P_{E+A} \sim R_D = 65^1$$

$$P_O^\infty = 1460$$

$$\mu = 8,4$$

3-Methylmercapto-5-methyl-1,4-diphenyl-1,2,4-triazoliniumjodid
(in Chloroform)

x_1	ϵ^{25}	d_4^{25}	P
0,000685	5,1177	1,4800	3488
0,001390	5,9247	1,4800	4268
0,002085	7,4655	1,4800	5297

$$P^\infty = 2600$$

$$P_{E+A} \sim R_D = 70^1$$

$$P_O^\infty = 2530$$

$$\mu = 11,0$$

»1,4,5-Triphenyl-3,5-endoanilo-1,2,4-triazolin« (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,000265	2,3018	0,8742	1690
0,000513	2,3246	0,8753	1519
0,000675	2,3367	0,8754	1439

$$P^\infty = 1825$$

$$P_{E+A} \sim R_D = 94^1$$

$$P_O^\infty = 1731$$

$$\mu = 9,1$$

»1,4,5-Triphenyl-3,5-endoanilo-1,2,4-triazolin« (in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P
0,000589	2,2836	1,0284	1678
0,001841	2,4192	1,0293	1607

$$P^\infty = 1713$$

$$P_{E+A} \sim R_D = 94^1$$

$$P_O^\infty = 1619$$

$$\mu = 8,8$$

¹ Berechnet aus den Atomrefraktionen.

»1,4-Diphenyl-5-methylmercapto-3,5-endooxy-1,2,4-triazolin«
(in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P
0,000349	2,2364	1,0282	1250
0,001044	2,2897	1,0291	1173

$$P_{E+A}^\infty \sim R_D = 1287$$

$$P_{E+A}^\infty \sim R_D = 63^1$$

$$P_O^\infty = 1224$$

$$\mu = 7,7$$

1,4-Diphenyl-3,5-dioxo-1,2,4-triazolidin (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,001116	2,2793	0,8752	137,7

$$P_{E+A} \sim R_D = 56^1$$

$$P^O = 81,7$$

$$\mu = 1,98$$

1,4-Diphenyl-3,5-dioxo-1,2,4-triazolidin (in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P
0,002075	2,2402	1,0304	197,6
0,004003	2,2600	1,0337	187,1

$$P_{E+A} \sim R_D = 207$$

$$P_{E+A} \sim R_D = 56^1$$

$$P_O^\infty = 151$$

$$\mu = 2,71$$

»2,3-Diphenyl-2,5-endothio-1,3,4-thiodiazolin« (in Benzol).

x_1	ϵ^{25}	d_4^{25}	P
0,0002331	2,2956	0,8738	(1534)
0,0003107	2,3069	0,8738	1674

$$P_{EO+A} = R_D = 63^1$$

$$P_O = 1611$$

$$\mu = 8,8$$

»2,3-Diphenyl-2,5-endothio-1,3,4-thiodiazolin« (in Dioxan).

x_1	ϵ^{25}	d_4^{25}	P
0,001218	2,3667	1,0286	1771

$$P_{E+A} \sim R_D = 63^1$$

$$P_O = 1708$$

$$\mu = 9,1$$

5-Methylmercapto-2,3-diphenyl-1,3,4-thiodiazoliniumjodid
(in Chloroform).

x_1	ϵ^{25}	d_4^{25}	P
0,000798	5,2366	1,4800	3671
0,001233	5,6151	1,4800	3784
0,001650	6,0575	1,4800	3931

$$P_{E+A}^\infty = 3605$$

$$P_{E+A}^\infty \sim R_D = 70^1$$

$$P_O^\infty = 3535$$

$$\mu = 13,1$$

¹ Berechnet aus den Atomrefraktionen.

2-Äthoxy-5-methylmercapto-2,3-diphenyl-1,3,4-thiodiazolin
(in Benzol).

x_1	ϵ^{25}	d_4^{25}	P	
0,003452	2,2871	0,8773	143,6	$P_{E+A}^\infty = 144$
0,005670	2,2960	0,8799	140,8	$\propto R_D = 75^1$
0,008280	2,3064	0,8826	140,1	$P_O^\infty = 69$
				$\mu = 1,83$

»3-Phenyl-2-methyl-2,5-endothio-1,3,4-thiodiazolin«
(in Dioxan I).

x_1	ϵ^{25}	d_4^{25}	P	
0,0002171	2,2318	1,0279	1667	$P_{E+A}^\infty = 1667$
0,0002602	2,2418	1,0279	1600	$\propto R_D = 45^1$
				$P_O^\infty = 1622$
				$\mu = 8,8$

Übersicht der in dieser Arbeit gemessenen Dipolmomente.

Temp.: 25 C.

μ : Dipolmoment.

L: Lösungsmittel.

P_O^∞ : Orientierungspolarisation des Gelösten.

		μ	L	P_O^∞
CH_2N_4	Tetrazol	5,11	Dioxan	543
CH_3N_5	5-Aminotetrazol	5,71	Dioxan	676,2
$C_2H_3N_3$	1,2,3-Triazol	1,77	Benzol	65
$C_2H_3N_3$	1,2,4-Triazol	3,17	Dioxan	208
$C_2H_4N_4$	1-Methyltetrazol	5,38	Benzol	600
$C_2H_3ON_3$	1,2,4-Triazol(5)	3,30	Dioxan	224
$C_3H_4N_2$	Pyrazol	1,57	Benzol	51,3
$C_3H_4N_2$	Imidazol	3,84	Benzol	305
C_3H_3ON	Isoxazol	2,81	Benzol	163,6
C_3H_3NS	Thiazol	1,64	Benzol	56,2
$C_3H_4N_2S$	2-Aminothiazol	1,75	Benzol	63,8
$C_3H_3ONS_2$	Rhodanin	2,20	Dioxan	101
$C_3H_3O_2NS$	2,4-Dioxothiazolidin	2,03	Dioxan	85
$C_4H_6N_2$	1-Methylpyrazol	2,28	Benzol	108,1
$C_4H_6N_2$	3-Methylpyrazol	1,43	Benzol	42,2

¹ Berechnet aus den Atomrefraktionen.

		μ	L	P_{O}^{∞}
$\text{C}_4\text{H}_6\text{ON}_2$	3-Methylpyrazolon-(5)	2,54	Dioxan	134
$\text{C}_4\text{H}_6\text{O}_2\text{N}_2$	Dimethylfuroxan	4,40	Benzol	400,1
$\text{C}_6\text{H}_5\text{N}_3$	Benzotriazol	4,07	Dioxan	344
$\text{C}_7\text{H}_6\text{N}_2$	Indazol	1,83	Benzol	69,6
$\text{C}_7\text{H}_6\text{N}_2$	Benzimidazol	3,93	Dioxan	320
$\text{C}_7\text{H}_5\text{ON}$	Benzoxazol	1,47	Benzol	44,7
$\text{C}_7\text{H}_5\text{ON}$	Indoxazen	3,03	Benzol	190,4
$\text{C}_7\text{H}_5\text{ON}$	Anthranil	3,06	Benzol	194,4
$\text{C}_8\text{H}_7\text{N}_3$	1-Phenyl-1,2,3-triazol	4,08	Benzol	345,4
$\text{C}_8\text{H}_7\text{N}_3$	2-Phenyl-1,2,3-triazol	0,97	Benzol	19,9
$\text{C}_8\text{H}_7\text{N}_3$	1-Phenyl-1,2,4-triazol	2,88	Benzol	172
$\text{C}_8\text{H}_7\text{N}_3$	4-Phenyl-1,2,4-triazol	5,63	Benzol	655,4
$\text{C}_9\text{H}_8\text{N}_2\text{S}_2$	»Phenyl-methyl-endothio- thiodiazolin«	8,8	Dioxan	1622
$\text{C}_{11}\text{H}_{12}\text{ON}_2$	Antipyrin	5,48	Benzol	621
$\text{C}_{11}\text{H}_{12}\text{N}_2\text{S}$	Thiopyrin	7,33	Benzol	1113
$\text{C}_{14}\text{H}_{10}\text{N}_2\text{S}_2$	»Diphenyl-endothio-thiodiazolin« ..	8,8	Benzol	1611
..	» » »	9,1	Dioxan	1708
$\text{C}_{14}\text{H}_{11}\text{O}_2\text{N}_3$	»Diphenyl-dioxo-triazolidin«	1,98	Benzol	81,7
..	» » »	2,71	Dioxan	151
$\text{C}_{15}\text{H}_{13}\text{ON}_3\text{S}$	»Diphenyl-methyl-endothio- triazolin«	8,4	CHCl_3	1460
$\text{C}_{15}\text{H}_{13}\text{ON}_3\text{S}$	»Diphenyl-methylmercapto-endoxy- triazolin«	7,7	Dioxan	1224
$\text{C}_{15}\text{H}_{13}\text{N}_2\text{S}_2\text{J}$	Methylmercapto-diphenyl- thiodiazoliniumjodid	13,1	CHCl_3	3535
$\text{C}_{16}\text{H}_{16}\text{N}_3\text{S}_2\text{J}$	Methylmercapto-methyl-diphenyl- triazoliniumjodid	11,0	CHCl_3	2530
$\text{C}_{17}\text{H}_{18}\text{ON}_2\text{S}_2$	Äthoxy-methylmercapto dipenyl- thiodiazolin	1,83	Benzol	69
$\text{C}_{26}\text{H}_{20}\text{N}_4$	»Triphenyl-endoanilo-triazolin« ..	8,8	Dioxan	1619
	» » »	9,1	Benzol	1731

Dem Direktor des Laboratoriums, Professor Dr. phil. EINAR BILMANN, danken wir für gute Arbeitsbedingungen und dem Carlsbergfond für finanzielle Unterstützung.

Universitetets kemiske Laboratorium.

Kopenhagen Mai 1943.

DET KGL. DANSKE VIDENSKABERNES SELSKAB
MATEMATISK-FYSISKE MEDDELELSER, BIND XX, NR. 21

DIE GEOMETRIE DER SCHWACHEN FIGUREN

VON

JOHANNES HJELMSLEV



KØBENHAVN
I KOMMISSION HOS EJNAR MUNKSGAARD
1943

Printed in Denmark.
Bianco Lunos Bogtrykkeri A/S

Einleitung.

1. Unter einer schwachen Figur soll eine variable Figur verstanden werden, deren Punkte sämtlich gegen ein und denselben Punkt, deren Geraden sämtlich gegen ein und dieselbe Gerade und deren Ebenen sämtlich gegen ein und dieselbe Ebene konvergieren. Alle Abstände, Winkel, Flächeninhalte, Volumina einer solchen Figur konvergieren also gegen Null, und die Geometrie der Figur handelt von den Relationen zwischen den Grenzverhältnissen dieser gegen Null strebenden Grössen. Im folgenden werden allgemeine Methoden zur Erforschung dieser Geometrie entwickelt.

Als den der Natur der Sache angepassten Arbeitsbereich wählen wir den projektiven Raum. Die Grenzverhältnisse zwischen den in einer Figur auftretenden Grössen derselben Art sind projektive Invarianten und daher unabhängig von der speziellen (euklidischen oder nichteuklidischen) Metrik, die man jederzeit als Hilfsmittel frei wählen kann. Die Geometrie der schwachen Figuren erweist sich somit als Kapitel der projektiven Geometrie.

2. Wir beginnen mit der Besprechung der infinitesimalprojektiven Gebilde.

Zwei eineindeutig und stetig aufeinander bezogene Punktreihen $O \cdots A \cdots$ und $O' \cdots A' \cdots$ auf zwei Geraden l und l' werden (infinitesimal) projektiv in O, O' genannt, wenn folgendes der Fall ist: Man wähle zwei von O bzw. O' verschiedene, feste Punkte (Bezugspunkte) U und U_1 auf l bzw. l' , ferner einen von O und U verschiedenen Punkt P auf l . Ist dann P'_1 der durch die Relation

$$OAPU \propto O'A'P'_1U_1$$

bestimmte Punkt von l' , so soll P'_1 gegen einen Punkt P_1 konvergieren, wenn A gegen O (und daher der entsprechende Punkt A' gegen O') strebt.

Die Punktreihen $O \cdot \cdot A \cdot \cdot \cdot$ und $O' \cdot \cdot A' \cdot \cdot \cdot$ werden ordinär projektiv oder singulär projektiv in O, O' genannt, je nachdem P_1 von O' und U_1 verschieden ist oder mit einem dieser Punkte zusammenfällt¹.

Wir bedienen uns auch der Redeweise, die infinitesimalen Vektoren OA und $O'A'$ seien projektiv und haben die Repräsentanten OPU und $O'P_1U_1$.

In ähnlichem Sinne kann man von in O, O' (infinitesimal) projektiven Punktreihen auf zwei Kurven sprechen. Man hat dann zwei variable Bezugspunkte zu benutzen, die gegen Grenzpunkte U, U_1 auf den Tangenten der Kurven in O und O' konvergieren, und P ist durch einen variablen Punkt zu ersetzen, der gegen einen Punkt P auf der Tangente der ersten Kurve konvergiert. Auch hier sprechen wir von projektiven infinitesimalen Vektoren OA und $O'A'$, welche durch die auf den Kurventangenten gelegenen Repräsentanten OPU und $O'P_1U_1$ dargestellt werden.

An die Stelle der beiden Punktreihen können in analoger Weise zwei beliebige (gleich- oder verschiedenartige) stetige eindimensionale Gebilde treten, die aufeinander bezogen sind, z. B. Geraden- oder Ebenenbüschel, allgemein (eindimensionale) Punktscharen, Geradenscharen u. s. w. Auch hier kann man von untereinander projektiven infinitesimalen Vektoren bzw. Winkeln sprechen.

3. Ein weiteres fundamentales Hilfsmittel bei diesen Untersuchungen sind die hyperprojektiven Gebilde.

Zwei eineindeutig und stetig aufeinander bezogene Punktreihen $O \cdot \cdot A \cdot \cdot B \cdot \cdot C \cdot \cdot \cdot$ und $O' \cdot \cdot A' \cdot \cdot B' \cdot \cdot C' \cdot \cdot \cdot$ auf zwei Geraden werden hyperprojektiv in den Punkten O, O' genannt, wenn jedem bestimmten Grenzverhältnis α zweier infinitesimaler Vektoren OA und OB der einen Punktreihe ein eindeutig bestimmtes Grenzverhältnis α' der beiden entsprechenden infinitesimalen Vektoren $O'A'$ und $O'B'$ der anderen Reihe entspricht. Wird die funktionale Abhängigkeit $\alpha' = \varphi(\alpha)$ als stetig vorausgesetzt, so folgt $\varphi(\alpha) = \alpha^r$, wo r eine reelle Zahl ist. Den Grenzverhältnissen

¹ Eine eingehendere Untersuchung dieser Begriffsbildung findet man in der Abhandlung des Verfassers: *Infinitesimale Elemente im projektiven Raum* (Neuvième congrès des mathém. scandinaves, Helsingfors 1938). Insbesondere sei auf die dort behandelten Beziehungen zu Krümmung und Torsion hingewiesen.

$$\frac{OA}{OB}, \frac{OB}{OC}, \frac{OA}{OC}$$

entsprechen nämlich bzw.

$$\frac{O'A'}{O'B'}, \frac{O'B'}{O'C'}, \frac{O'A'}{O'C'},$$

so dass für beliebige Werte von α und β

$$\varphi(\alpha) \varphi(\beta) = \varphi(\alpha\beta)$$

gelten muss, woraus eben $\varphi(\alpha) = \alpha^r$ folgt.

Die Zahl r wird als Exponent der Korrespondenz oder genauer als Exponent der Korrespondenz $A \rightarrow A'$ im Punktpaar O, O' bezeichnet. Ist die Korrespondenz ordinär infinitesimal projektiv, so ist $r = 1$. Der umgekehrte Schluss ist jedoch nicht statthaft.

4. Auch hier können an die Stelle der Punktreihen beliebige andere eindimensionale stetige Gebilde treten, beispielsweise eine Punktschar $O \cdot \cdot A \cdot \cdot B \cdot \cdot \cdot$ und eine Geradenschar $o \cdot \cdot a \cdot \cdot b \cdot \cdot \cdot$, die derart aufeinander bezogen sind, dass jedem bestimmten Grenzverhältnis α zweier infinitesimaler Vektoren OA und OB der Punktschar eindeutig ein bestimmtes Grenzverhältnis α' der zugeordneten infinitesimalen Winkel oa und ob der Geradenschar entspricht, wobei die Abhängigkeit $\alpha' = \varphi(\alpha)$ wieder als stetig vorausgesetzt wird. In allen diesen Fällen wird von hyperprojektiven Gebilden gesprochen, und jedesmal gibt es eine bestimmte reelle Zahl r , derart dass $\varphi(\alpha) = \alpha^r$ ist, und diese Zahl heisst der Exponent der Korrespondenz, genauer: die Korrespondenz $A \rightarrow a$ hat den Exponenten r , die Korrespondenz $a \rightarrow A$ den Exponenten $\frac{1}{r}$.

I. Unendlich kleine Grössen.

5. Cauchy hat in seinen Exercices de Mathématiques, 1826, p. 145—150, die Ordnung unendlich kleiner Grössen folgendermassen definiert:

Eine Funktion $f(x)$, die für $x \rightarrow 0$ gegen 0 strebt, heisst unendlich klein von der Ordnung a (in Bezug auf x), wenn

$$\frac{f(x)}{x^r} \rightarrow \begin{cases} 0 & \text{für } r < a \\ \infty & \text{für } r > a, \end{cases}$$

während an das Verhalten von $\frac{f(x)}{x^r}$ für $r = a$ keine Forderung gestellt wird. Als Beispiele führt Cauchy die Funktionen

$$e^x x^a, \quad e^x x^a \log x, \quad e^x x^a \frac{1}{\log x}$$

an, die nach Division durch x^a bzw. gegen

$$1, \quad \infty, \quad 0$$

konvergieren. Selbstverständlich ist hierbei vorauszusetzen, dass x durch positive Werte gegen 0 geht.

Die Definition scheint ganz in Vergessenheit geraten zu sein; denn sie wird in keinem der bekannten neueren Standardwerke über Analysis erwähnt¹.

Wir wollen hier Cauchy's Definition wieder aufnehmen und hinzufügen: Eine Funktion $f(x)$, die für $x \rightarrow 0$ gegen 0 strebt, heisse unendlich klein von unendlich hoher Ordnung, wenn $\frac{f(x)}{x^r} \rightarrow 0$ für jede positive Zahl r und unendlich klein von der Ordnung 0, wenn $f(x) \rightarrow 0$ für $x \rightarrow 0$, aber $\frac{f(x)}{x^r} \rightarrow \infty$ für jeden positiven Wert von r . Beispiele von Funktionen der beiden letzteren Arten sind

$$y = e^{-\frac{1}{x}}, \quad y = -\frac{1}{\log x}.$$

Durch diese Definitionen wird also jeder unendlich kleinen Grösse $f(x)$ eine Zahl zugeordnet, die die Ordnung von $f(x)$ angibt und entweder 0, positiv oder ∞ ist.

Wie man sieht, werden für $x \neq 0$ keinerlei die Stetigkeit von $f(x)$ betreffende Annahmen gemacht, und auch der Grenzübergang $x \rightarrow 0$ braucht nicht kontinuierlich vor sich zu gehen; x kann eine beliebige Fundamentalfolge mit dem Grenzwert 0

¹ In J. Hadamard's Cours d'Analyse I, p. V wird im Gegensatz dazu sogar gesagt, dass $x^n (\log x)^p$ keine bestimmte Ordnung beigelegt werden soll.

durchlaufen. Übrigens können auch mehrdeutige Funktionen zugelassen werden.

6. Wir werden uns jedoch besonders für den Fall interessieren, wo die Gleichung $y = f(x)$ für $x > 0$ und $x \rightarrow 0$ einen Kurvenbogen im ersten Quadranten darstellt, der in jedem Punkt eine eindeutige, mit dem Punkt stetig variierende Tangente besitzt, und für den die positive x -Achse Halbtangente im Nullpunkt ist, so dass $\frac{f(x)}{x} \rightarrow 0$. Die Ordnung von $f(x)$ ist dann mindestens 1. Dass sie genau 1 ist, bedeutet, dass die Kurve $y = f(x)$ in der Umgebung des Nullpunktes von allen Parabeln $y = px^r$, wo $r > 1$, $p > 0$, umschlossen wird. Ist die Ordnung $a > 1$, so wird die Kurve von allen Parabeln $y = px^r$, wo $r > a$, $p > 0$, umschlossen, während sie selbst die zu den Werten $1 < r < a$ gehörigen Parabeln umschliesst. Die Parabeln $y = px^a$, $p > 0$, sind dann diejenigen, die der Kurve in der Umgebung von O am nächsten kommen. Durch jeden Punkt P der Kurve geht genau eine dieser Parabeln. Lässt man P gegen O gehen, und konvergiert diese Parabel hierbei gegen eine Grenzlage, so wird die letztere als oskulierende Parabel der Kurve in O bezeichnet.

Dass die Ordnung a unendlich gross ist, bedeutet, dass die Kurve in der Umgebung von O alle Parabeln $y = px^r$ umschliesst.

7. In Übereinstimmung mit der üblichen Terminologie sagen wir, die Kurve $y = f(x)$ habe eine Berührungsordnung (mit ihrer Tangente), die gleich der Ordnung von $\frac{f(x)}{x}$, also gleich $a-1$ ist.

8. Ist $y = f(x)$ unendlich klein von einer Ordnung, die < 1 und ≥ 0 ist, so berührt die Kurve die y -Achse in O , und man kann dann nach Vertauschung von x und y die obigen Betrachtungen auf die inverse Funktion anwenden.

9. Ausser einer Reihe naheliegender allgemeiner Sätze über Summe, Produkt u. a. von unendlich kleinen Grössen führt Cauchy an der genannten Stelle einen Satz an, der in Beziehung zu den Untersuchungen steht, die wir im folgenden anstellen werden:

Ist $f(x)$ unendlich klein von der Ordnung a , so konvergiert die Grösse $\frac{xf'(x)}{f(x)}$ für $x \rightarrow 0$ gegen a . Für die Kurve $y = f(x)$

(Fig. 1), die hier als konvex in der Umgebung von O vorausgesetzt werden soll, bedeutet dies, dass die x -Achse von der Tangente in P in einem solchen Punkt Q geschnitten wird, dass $\frac{OR}{QR} \rightarrow a$ und folglich $\frac{QP}{OQ} \rightarrow \frac{1}{a-1}$. Das Grenzwertverhältnis $\frac{1}{a-1}$ von $\frac{QP}{OQ}$ nennen wir den Index der Kurve im Punkte O , so dass Cauchy's Bemerkung besagt, dass der Index im Nullpunkt gleich dem reziproken Wert der Berührungsordnung ist.

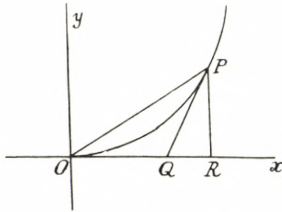


Fig. 1.

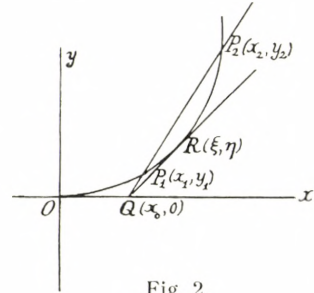


Fig. 2.

Im folgenden werden wir mit unseren Zielen vor Augen auf diese Eigenschaft zurückkommen.

10. Wir schränken nun unseren Kurvenbegriff $y = f(x)$ derart ein, dass der betrachtete Bogen OA in der Umgebung von O konvex (jedoch nicht notwendig frei von Ecken) ist. Wir nehmen ferner an, dass Abszisse x und Ordinate y des variablen Kurvenpunktes eine in O hyperprojektive Korrespondenz mit dem Exponenten $r+1$, wo $r > 0$ ist, bilden, dass also für zwei Punkte $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ der Kurve (Fig. 2), die beide gegen O konvergieren, die Grenzwertverhältnisse $\frac{y_2}{y_1}$ und $\frac{x_2}{x_1}$ in der Beziehung

$$\frac{y_2}{y_1} = \left(\frac{x_2}{x_1} \right)^{r+1}$$

stehen¹.

Ist das Grenzwertverhältnis $\frac{x_2}{x_1} = \alpha > 1$, so ist das Grenzwertverhältnis $\frac{y_2}{y_1}$ folglich $= \alpha^{r+1}$, und da sich die Abszisse x_0 des Schnittpunktes Q der x -Achse mit P_1P_2 aus

¹ Dabei ist $\frac{y_2}{y_1}$ statt $\lim \frac{y_2}{y_1}$ und $\frac{x_2}{x_1}$ statt $\lim \frac{x_2}{x_1}$ geschrieben. Diese kurze Schreibweise wird im folgenden überall verwendet.

$$x_0 = \frac{x_1 y_2 - x_2 y_1}{y_2 - y_1}$$

bestimmt, hat man die Grenzverhältnisse

$$\frac{x_0}{x_1} = \frac{\frac{y_2 - x_2}{y_1} \cdot x_1}{\frac{y_2 - 1}{y_1}} = \alpha \frac{\alpha^r - 1}{\alpha^{r+1} - 1}$$

und

$$\frac{x_0}{x_2} = \frac{\alpha^r - 1}{\alpha^{r+1} - 1}.$$

11. Die Tangente von Q aus an den Bogen $P_1 P_2$ berühre ihn im Punkte $R = (\xi, \eta)$. Das Grenzverhältnis $\frac{x_0}{\xi}$ ist dann an das Intervall $\left(\frac{x_0}{x_2}, \frac{x_0}{x_1}\right)$, also an $\left(\frac{\alpha^r - 1}{\alpha^{r+1} - 1}, \alpha \frac{\alpha^r - 1}{\alpha^{r+1} - 1}\right)$ gebunden.

Setzt man $\alpha = 1 + \varepsilon$, so erhält man

$$\frac{\alpha^r - 1}{\alpha^{r+1} - 1} = \frac{(1 + \varepsilon)^r - 1}{(1 + \varepsilon)^{r+1} - 1} = \frac{r}{r+1} \cdot \frac{1}{1 + \eta},$$

wo $\varepsilon > \eta > 0$; und das genannte Intervall kann

$$\left(\frac{r}{r+1} \cdot \frac{1}{1 + \eta}, \frac{r}{r+1} \cdot \frac{1 + \varepsilon}{1 + \eta}\right)$$

geschrieben werden, woraus ersichtlich ist, dass es im Intervall

$$\left(\frac{r}{r+1} \cdot \frac{1}{1 + \varepsilon}, \frac{r}{r+1} (1 + \varepsilon)\right)$$

enthalten ist.

12. Ist $R = (\xi, \eta)$ ein beliebiger von O verschiedener Punkt der Kurve, so kann man durch den Punkt $Q = (x_0, 0)$, in welchem die Tangente (oder eine Stützgerade) in R die x -Achse schneidet, stets eine Sekante $QP_1 P_2$ derart legen, dass für einen gegebenen, genügend kleinen positiven Wert von ε das Verhältnis $\frac{QP_2}{QP_1} = 1 + \varepsilon$ wird. Aus dem Obigen entnimmt man dann,

dass alle Häufungswerte von $\frac{x_0}{\xi}$ beim Grenzübergang $R \rightarrow O$ für jedes ε dem Intervall

$$\left(\frac{r}{r+1} \cdot \frac{1}{1+\varepsilon}, \frac{r}{r+1} (1+\varepsilon) \right)$$

angehören müssen, d. h. das Grenzwertverhältnis $\frac{x_0}{\xi}$ ist eindeutig bestimmt und gleich $\frac{r}{r+1}$. Hiermit gleichbedeutend ist, dass das Grenzwertverhältnis

$$\frac{OQ}{OR} = \frac{r}{r+1}$$

oder

$$\frac{QR}{OQ} = \frac{1}{r}$$

ist, also dass die Kurve den Index $\frac{1}{r}$ in O hat.

13. Dass die Korrespondenz $x \rightarrow y$ hyperprojektiv mit dem Exponenten $r+1$ ist, hat zur Folge, dass die Korrespondenz $x \rightarrow \frac{y}{x}$ hyperprojektiv mit dem Exponenten r ist, d. h.:

Die Korrespondenz, die dem laufenden Punkt P des Bogens OA seinen projizierenden Strahl OP zuordnet, ist hyperprojektiv mit dem Exponenten r im Grenzpunkt O und dessen zugeordneter Gerade, der Tangente in O .

Die Zahl r soll daher der Krümmungsexponent in O genannt werden. Von diesem Gesichtspunkte aus könnte man den bereits als Index bezeichneten reziproken Wert $\frac{1}{r}$ auch den Krümmungsgrad nennen.

Es liegt in der Natur der Sache, dass die Zahl r eine projektive Invariante ist.

Aus der Existenz des Krümmungsexponenten r als Exponent einer für den Grenzpunkt O und die Tangente in O hyperprojektiven Korrespondenz zwischen dem laufenden Kurvenpunkt P und seinem projizierenden Strahl OP kann man umgekehrt auf eine hyperprojektive Korrespondenz mit dem Exponenten $r+1$ zwischen den Koordinaten x, y des laufenden Punktes in Bezug

auf ein Koordinatensystem mit der Tangente als x -Achse und einer beliebigen anderen Geraden durch O als y -Achse schliessen.

Den Krümmungsexponenten r wollen wir der Kürze halber den Exponenten in O und die Zahl $r+1$ den Ordinatenexponenten in O nennen.

14. Wir wollen nun eingehender die Bedingung dafür untersuchen, dass die Relation $y = f(x)$ eine in $x = 0$ hyperprojektive Korrespondenz $x \rightarrow y$ mit dem Exponenten n (> 0) darstellt. Bei dieser Untersuchung soll die Funktion $f(x)$ vorläufig nur als positiv und stetig für alle $x > 0$ innerhalb eines gewissen Intervalls vorausgesetzt werden. Es handelt sich um die Frage, was man aus der für jedes $\mu < 1$ und ≥ 0 gegebenen Grenzbedingung

$$\frac{f(tx)}{f(x)} \rightarrow \mu^n \quad \text{für } x \rightarrow 0, \quad t \rightarrow \mu,$$

über das Verhalten der Funktion im Nullpunkt selbst schliessen kann.

Lässt man x eine »grenzgeometrische Progression«, d. h. eine Fundamentalfolge

$$x_0, x_1, x_2, \dots, x_i, \dots$$

durchlaufen, für welche

$$\frac{x_{i+1}}{x_i} \rightarrow \mu, \quad 0 \leq \mu < 1,$$

gilt, so bilden die entsprechenden Funktionswerte

$$f(x_0), f(x_1), f(x_2), \dots, f(x_i), \dots$$

wegen der gegebenen Grenzbedingung eine grenzgeometrische Progression mit dem Grenzquotienten μ^n und konvergieren daher ebenfalls gegen Null.

Hieraus kann man aber schliessen, dass $f(x)$ gegen Null konvergiert, wenn x eine ganz beliebige Nullfolge positiver Zahlen durchläuft.

Man nehme nämlich an, dass es eine Nullfolge positiver Zahlen $x_0, x_1, x_2, \dots, x_i, \dots$ derart gäbe, dass für die entsprechenden Funktionswerte

$$f(x_0), f(x_1), \dots, f(x_i), \dots \rightarrow k > 0$$

gilt. Man könnte dann stets eine solche Teilfolge

$$x_0, x_{i_1}, x_{i_2}, \dots$$

auswählen, dass

$$\frac{x_{i_{r+1}}}{x_{i_r}} < \frac{1}{2}.$$

Entweder ist dann diese Teilfolge selbst oder eine ihrer Teilfolgen eine grenzgeometrische Progression mit einem Grenzquotienten $\leq \frac{1}{2}$ (eventuell 0). Dann müssten aber die entsprechenden Funktionswerte gegen Null konvergieren, während sie nach Annahme gegen $k > 0$ konvergieren. Der Beweis gilt auch für $k = \infty$.

Es ergibt sich somit, dass für jede Nullfolge positiver Zahlen x_0, x_1, \dots die zugehörigen Funktionswerte $f(x_0), f(x_1), \dots$ ebenfalls eine Nullfolge bilden, dass also $f(x) \rightarrow 0$ für $x \rightarrow 0$, oder m. a. W. dass $f(x)$ eine unendlich kleine Grösse definiert.

15. Aus der obigen Untersuchung ist sofort ersichtlich, dass die Grenzbedingung

$$\frac{f(tx)}{f(x)} \rightarrow \mu^n \quad \text{für } x \rightarrow 0, \quad t \rightarrow \mu,$$

für einen negativen Wert n die Relation $f(x) \rightarrow \infty$ für $x \rightarrow 0$ zur Folge hat, so dass $f(x)$ eine unendlich grosse Grösse definiert.

16. Schliesslich nun der wichtige Satz: Hat die Kurve $y = f(x)$ im Nullpunkt O den Krümmungsexponenten r (und folglich den Ordinatenexponenten $r+1$), so ist ihre Berührungsordnung mit der Tangente in O gleich r , und $f(x)$ stellt eine unendlich kleine Grösse der Ordnung $r+1$ dar.

Setzt man nämlich

$$\varphi(x) = \frac{f(x)}{x^n}, \quad x > 0,$$

so hat man

$$\frac{\varphi(tx)}{\varphi(x)} = \frac{f(tx)}{f(x)} \cdot \frac{(tx)^n}{x^n} \rightarrow t^{r+1-n} \quad \text{für } x \rightarrow 0.$$

Für $n > r + 1$ ist $\varphi(x)$ also eine unendlich kleine und für $n > r + 1$ eine unendlich grosse Grösse, d. h. $f(x)$ ist unendlich klein von der Ordnung $r + 1$, $\frac{f(x)}{x}$ unendlich klein von der Ordnung r , und damit ist der Satz bewiesen.

17. Für $n = r + 1$ entsteht eine Funktion

$$\varphi(x) = \frac{f(x)}{x^{r+1}}, \quad x > 0,$$

über deren Verhalten für $x \rightarrow 0$ von vornherein nicht mehr als $\frac{\varphi(tx)}{\varphi(x)} \rightarrow 1$ ausgesagt werden kann. Nimmt man jedoch an, dass $\varphi(x)$ für $x \rightarrow 0$ einen endlichen oder unendlichen Grenzwert hat, so kommen folgende Möglichkeiten in Betracht:

1°. $\varphi(x) \rightarrow a \neq 0$ und ∞ für $x \rightarrow 0$.

In diesem Fall ist die Grenzbedingung stets erfüllt.

2°. $\varphi(x) \rightarrow 0$ für $x \rightarrow 0$.

Setzt man $\frac{\varphi(x)}{x} = \psi(x)$, so hat man

$$\frac{\psi(tx)}{\psi(x)} = \frac{\varphi(tx)}{tx} \cdot \frac{\varphi(x)}{x} \rightarrow \frac{1}{t},$$

woraus folgt, dass

$$\psi(x) = \frac{\varphi(x)}{x} \rightarrow \infty \quad \text{für } x \rightarrow 0,$$

also, dass die y -Achse Tangente der Kurve $y = \varphi(x)$ im Nullpunkt ist.

Für zwei variable Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ dieser Kurve (Fig. 3) soll nun gelten, dass $\frac{y_2}{y_1}$, wenn $\frac{x_2}{x_1}$ gegen einen Wert > 1 konvergiert, gegen 1 konvergiert. Wählt man nun die beiden variablen Punkte derart, dass $\frac{y_2}{y_1}$ gegen einen Wert > 1 konvergiert,

so muss $\frac{x_2}{x_1}$ ins Unendliche wachsen. Die inverse Korrespondenz $y \rightarrow x$ ist also hyperprojektiv mit unendlich hohem Exponenten.

(Die Korrespondenz $y \rightarrow x$ ist hierbei eindeutig gedacht; im übrigen spielt dies aber keine Rolle.)

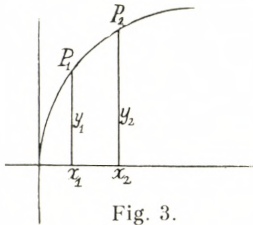


Fig. 3.

Als Resultat hat man daher, dass die fragliche Funktion φ in diesem Fall die inverse einer Funktion mit unendlich hohem Exponenten ist.

$$3^\circ. \varphi(x) \rightarrow \infty \quad \text{für } x \rightarrow 0.$$

Die reziproke Funktion $\frac{1}{\varphi(x)}$ konvergiert dann gegen 0, und dieser Fall kann daher sofort auf den vorigen mit dem Resultat zurückgeführt werden, dass $\varphi(x)$ hier die reziproke der inversen einer Funktion mit unendlich hohem Exponenten ist.

Die einfachsten Beispiele von Funktionen $\varphi(x)$ für die drei Fälle $1^\circ, 2^\circ, 3^\circ$ sind:

$$\varphi(x) = e^x, \quad \varphi(x) = -\frac{1}{\log x}, \quad \varphi(x) = -\log x.$$

Diesem entsprechend hat man die folgenden drei Beispiele von Kurven mit dem Exponenten r im Nullpunkt

$$y = x^{r+1}e^x, \quad y = -\frac{x^{r+1}}{\log x}, \quad y = -x^{r+1}\log x.$$

II. Schwache Figuren in der Ebene.

18. Ein schwaches Dreieck ABC ist ein variables Dreieck, dessen Ecken alle gegen einen Punkt O und dessen Seiten alle gegen eine Gerade l konvergieren (Fig. 4). Die infinitesimalen Seiten und Winkel des Dreiecks konvergieren sämtlich gegen Null. Die Geometrie des Dreiecks hat es mit den Relationen zwischen den Grenzverhältnissen der

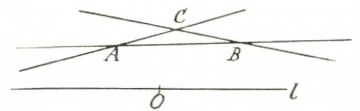


Fig. 4.

Seiten und den Grenzverhältnissen der Winkel zu tun. (Unter »Winkel« verstehen wir »infinitesimale Winkel«.) Die Grenzver-

hältnisse der Seiten bzw. Winkel seien durch die Verhältniszahlen a, b, c bzw. A, B, C gegeben. Man hat dann wegen der Konvergenz der Richtungen AC, CB, AB gegen dieselbe Richtung

$$c = a + b, \quad C = A + B,$$

$$(a, b, c) \sim (A, B, C),$$

wo die letzte Relation die Proportionalität von a, b, c und A, B, C zum Ausdruck bringt¹.

Hiermit ist die Trigonometrie des schwachen Dreiecks erschöpft.

19. Indem wir nun zur Betrachtung einer aus mehr als drei Punkten A_1, A_2, A_3, \dots bestehenden schwachen Figur mit dem Grenzpunkt O übergehen, wollen wir einschränkend annehmen, dass der Punkt O der Figur in dem Sinne angehört, dass die Geraden OA_i dieselbe Grenzlage wie die Geraden A_iA_k haben.

In der Figur (Fig. 5) ist angenommen, dass die Halbgerade x die gemeinsame Grenzlage der Halbgeraden $O(A_1), A_1(A_2), A_2(A_3), A_3(A_4), \dots$ u. s. w. ist. Das gleiche gilt dann für die Halbgeraden $O(A_2), O(A_3), O(A_4), A_1(A_3), A_1(A_4), A_2(A_4)$ u. s. w.

Wir nehmen an, dass die Abstände $OA_1, OA_2, OA_3, OA_4, \dots$ Grenzverhältnisse haben, die durch die positiven Verhältniszahlen $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \dots$ (wobei bis auf weiteres $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4 < \dots$ vorausgesetzt wird) angegeben werden, so dass also

$$(OA_1, OA_2, OA_3, OA_4, \dots) \sim (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \dots).$$

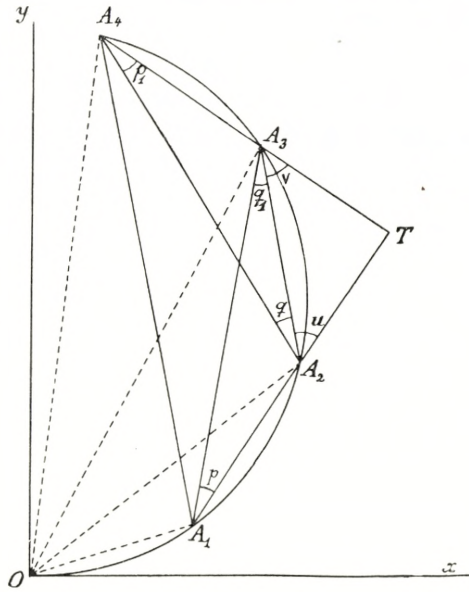


Fig. 5.

¹ Dieses Resultat ist zum ersten Mal von L'Hospital in seinem berühmten Werk: *Analyse des infiniments petits, pour l'intelligence des lignes courbes*, Paris 1696, sec. édit. 1716, p. 153. angegeben worden.

Ferner wird angenommen, dass die Halbgeraden $O(A_1)$, $O(A_2)$, $O(A_3)$, $O(A_4)$, \dots Winkel mit der gemeinsamen Grenzlage x bilden, deren Grenzverhältnisse durch positive Verhältniszahlen $\theta_1, \theta_2, \theta_3, \theta_4, \dots$ angegeben werden können.

Sämtliche Grenzverhältnisse zwischen Abständen und zwischen Winkeln der Figur lassen sich dann durch die Grössen α_i und θ_i ausdrücken.

In einem affinen Koordinatensystem x, y können die Koordinaten jedes der Punkte A_i durch $(\alpha_i, \theta_i \alpha_i)$ repräsentiert werden. Führt man die Grösse

$$[ikl] = \theta_i \alpha_i (\alpha_l - \alpha_k) + \theta_k \alpha_k (\alpha_i - \alpha_l) + \theta_l \alpha_l (\alpha_k - \alpha_i)$$

ein, die den doppelten Flächeninhalt des Dreiecks $A_i A_k A_l$ repräsentiert, so kann der infinitesimale Winkel $(ikl) = \angle A_i A_k A_l$ d. h. der infinitesimale Winkel zwischen den Geraden $A_i A_k$ und $A_k A_l$ durch

$$(ikl) = \frac{[ikl]}{(\alpha_k - \alpha_i)(\alpha_l - \alpha_k)}$$

ausgedrückt werden.

Führt man für die in der Figur hervorgehobenen infinitesimalen Winkel die dort angegebenen Bezeichnungen p, q, p_1, q_1, u, v ein, so erhält man das Grenzverhältnis

$$\frac{u}{v} = \frac{p + q_1}{q + p_1} = \frac{p}{p_1} \cdot \frac{1 + \frac{q_1}{p}}{\frac{q}{p_1} + 1} = \frac{[123]}{[234]} \cdot \frac{\alpha_4 - \alpha_3}{\alpha_2 - \alpha_1}.$$

Bezeichnet man den Schnittpunkt der Geraden $A_1 A_2$ und $A_3 A_4$ mit T , so hat man das Grenzverhältnis

$$\tau = \frac{A_2 T}{A_1 A_2} = \frac{[234]}{[134] - [234]}.$$

Die Strecken $A_i A_k$, $i < k$, werden durch $\alpha_k - \alpha_i$ und die Strecke OT durch

$$OT = OA_2 + A_2 T = \alpha_2 + (\alpha_2 - \alpha_1) \tau$$

repräsentiert.

Man sieht nun leicht, dass man für alle Abstände und Winkel der Figur durch die Grössen α_i und θ_i ausdrückbare, repräsentierende Zahlen finden kann.

20. Sind die Punkte O und A_i der Figur speziell an eine feste konvexe, die x -Achse in O berührende Kurve mit dem Exponenten $r(>0)$ in O gebunden, so hat man

$$\frac{u}{v} = \frac{\alpha_1^{r+1}(\alpha_3 - \alpha_2) + \alpha_2^{r+1}(\alpha_1 - \alpha_3) + \alpha_3^{r+1}(\alpha_2 - \alpha_1) \cdot \alpha_4 - \alpha_3}{\alpha_2^{r+1}(\alpha_4 - \alpha_3) + \alpha_3^{r+1}(\alpha_2 - \alpha_4) + \alpha_4^{r+1}(\alpha_3 - \alpha_2) \cdot \alpha_2 - \alpha_1}$$

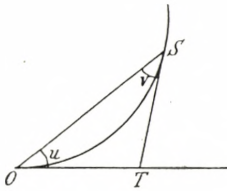


Fig. 6.

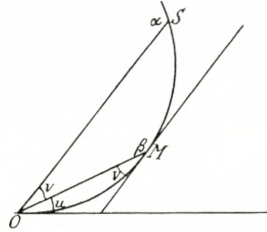


Fig. 7.

Für $r = 1$ ergibt sich insbesondere

$$\frac{u}{v} = \frac{\alpha_3 - \alpha_1}{\alpha_4 - \alpha_2}$$

Lässt man α_1 und α_2 gegen 0 und α_3 und α_4 gegen α konvergieren, so geht das Grenzwertverhältnis $\frac{u}{v}$ in

$$\frac{u}{v} = \frac{TS}{OT} = \frac{1}{r}$$

über (Fig. 6).

21. Von den speziellen, mit der Kurve verknüpften schwachen Figuren erwähnen wir noch diejenige, die entsteht, wenn man die zur infinitesimalen Sehne OS parallele Tangente zieht. Gehört zu S die Zahl α und zum Berührungspunkt M die Zahl β , so hat man mit den in der Figur (Fig. 7) angegebenen Bezeichnungen

$$\frac{u}{v} = \frac{1}{r},$$

$$\frac{u+v}{u} = 1+r = \left(\frac{\alpha}{\beta}\right)^r,$$

also das Grenzverhältnis

$$\frac{\alpha}{\beta} = \left(1+r\right)^{\frac{1}{r}}.$$

Wächst r ins Unendliche, so geht dieses Grenzverhältnis gegen 1; konvergiert r gegen 0, so geht es gegen e .

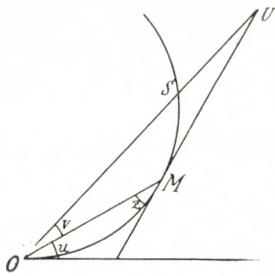


Fig. 8.

Es sei bemerkt, dass die Bedingung, dass die Tangente in M parallel OS ist, in diesem Zusammenhang in Wirklichkeit nur bedeutet, dass der Schnittpunkt der Tangente mit OS bei dem durchgeführten Grenzübergang keine Grenzlage in O hat.

22. Wir wollen noch die etwas weitergehende Aufgabe lösen, wo gefordert wird, dass der Schnittpunkt U der Tangente in M mit der Geraden OS beim Grenzübergang derart variiert, dass ein eindeutig bestimmtes Grenzverhältnis

$$\frac{OU}{OS} = \mu$$

existiert, das mit Vorzeichen, positiv oder negativ gerechnet wird, je nachdem OU und OS dieselbe oder entgegengesetzte Richtung haben.

In diesem Fall hat man mit den in Fig. 8 angegebenen Bezeichnungen

$$\frac{v}{z} = \frac{MU}{OU} = 1 - \frac{OM}{OU} = 1 - \frac{OM}{OS} \cdot \frac{1}{\mu},$$

$$\frac{z}{u} = r,$$

also

$$\frac{v}{u} = r - \frac{r}{\mu} \cdot \frac{OM}{OS},$$

während

$$\frac{v+u}{u} = \left(\frac{OS}{OM}\right)^r$$

ist, so dass man zur Bestimmung von $\xi = \frac{OS}{OM}$ die Gleichung

$$\xi^{r+1} - (r+1)\xi + \frac{r}{\mu} = 0$$

erhält. Hierin ist für $\mu = \infty$ die Lösung der vorigen Aufgabe (21.) enthalten.

23. Im übrigen verweisen wir auf die in 10. angegebenen Formeln, nach welchen die Grenzverhältnisse zwischen Abszissen und zwischen Ordinaten variabler Punkte (x, y) , die gegen O konvergieren, durch eine Gleichung der Form

$$\frac{y}{y_1} = \left(\frac{x}{x_1}\right)^{r+1}$$

der Behandlung zugänglich und die einschlägigen Fragen auf elementare analytisch-geometrische Betrachtungen reduziert werden. Die genannte Gleichung wird als Grenzgleichung der Kurve in O bezeichnet.

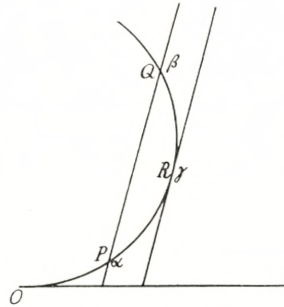


Fig. 9.

24. Als weiteres Anwendungsbeispiel sei noch die folgende Verallgemeinerung von 21. genannt: Auf dem Bogen PQ (Fig. 9), der gegen O konvergiert, werde ein Punkt R bestimmt, dessen Tangente der Sehne parallel ist. Die repräsentierende Zahl γ von R und die repräsentierenden Zahlen α und β von P bzw. Q sind dann durch die Gleichung

$$\gamma^r = \frac{1}{r+1} \cdot \frac{\alpha^{r+1} - \beta^{r+1}}{\alpha - \beta}$$

verknüpft.

Die repräsentierende Zahl des Schnittpunktes der Tangenten in P und Q ergibt sich leicht mit Hilfe einer früheren Betrachtung (19.).

25. Wir wenden uns nun der Untersuchung von zwei konvexen Kurven k und k_1 (Fig. 10) zu, die vom selben Punkt O ausgehen und dort dieselbe Halbtangente x und denselben Exponenten r haben. Die Kurven mögen ferner auf derselben Seite

der Tangente in O liegen, so dass sie eine Spitze zweiter Art bilden. Zur Behandlung der schwachen Figuren, die mit beiden Kurven verknüpft sind, müssen wir ausser den Grenzgleichungen der Kurven noch das Grenzverhältnis

$$i_1 = \frac{OA_1}{OA}$$

der beiden auf einer Sekante durch O ausgeschnittenen Sehnen heranziehen.

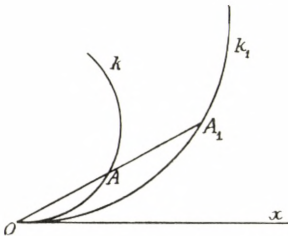


Fig. 10.

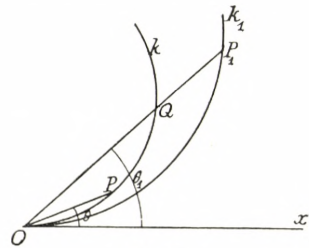


Fig. 11.

Von dieser Grösse, die als das Krümmungsverhältnis der Kurven (Krümmungsverhältnis von k und k_1) bezeichnet wird, wird vorausgesetzt, dass sie einen eindeutigen Wert hat. Mit ihrer Hilfe können beliebige Sehnen OP, OP_1 und ihre Neigungswinkel θ, θ_1 (Fig. 11) verglichen werden:

$$\frac{\theta}{\theta_1} = \left(\frac{OP}{OQ}\right)^r = \left(\frac{OP}{OP_1} \cdot \frac{OP_1}{OQ}\right)^r,$$

also

$$\frac{\theta}{\theta_1} = \left(i_1 \frac{OP}{OP_1}\right)^r.$$

26. Führt man (Fig. 12) ein affines Koordinatensystem x, y ein, so hat man für zwei Punkte (x_1, y) und (x_1, y_1) mit derselben Abszisse x_1

$$\frac{\theta}{\theta_1} = \left(\frac{x_1}{x_2}\right)^r = i_1^r, \quad \frac{y}{y_1} = \frac{\theta}{\theta_1},$$

also

$$\frac{y}{y_1} = i_1^r.$$

Von den Kurven k_1 der genannten Art, die ein bestimmtes endliches, von 0 verschiedenes Krümmungsverhältnis zu einer

gegebenen Kurve k haben, soll gesagt werden, dass sie den durch k bestimmten schwachen Bereich um O angeben; jeder Punkt (x, y) dieses Bereichs befriedigt eine Gleichung der Form

$$\frac{y}{y_1} = \mu \left(\frac{x}{x_1} \right)^{r+1},$$

wo μ ein Parameter ist.

Zwei Punkte $(x_1 \alpha_1, y_1 \mu_1 \alpha_1^{r+1})$ und $(x_1 \alpha_2, y_1 \mu_2 \alpha_2^{r+1})$, $\alpha_1 \neq \alpha_2$, des Bereiches bestimmen eine Gerade mit der Steigung

$$\omega = \frac{y_1 \cdot \mu_1 \alpha_1^{r+1} - \mu_2 \alpha_2^{r+1}}{x_1 \alpha_1 - \alpha_2}$$

und mit der Abszisse (das ist das auf der x -Achse abgeschnittene Stück)

$$x = x_1 \frac{\alpha_1 \alpha_2 (\mu_1 \alpha_1^r - \mu_2 \alpha_2^r)}{\mu_1 \alpha_1^{r+1} - \mu_2 \alpha_2^{r+1}}.$$

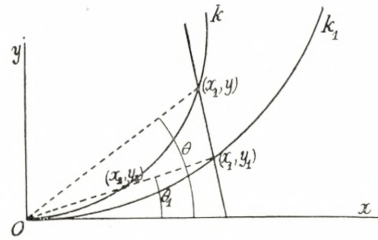


Fig. 12.

Setzt man $\frac{y_1}{x_1} = \omega_1$, so kann man sagen, dass die Gerade durch die Grenzverhältnisse

$$u = \frac{\omega}{\omega_1} = \frac{\mu_1 \alpha_1^{r+1} - \mu_2 \alpha_2^{r+1}}{\alpha_1 - \alpha_2},$$

$$\xi = \frac{x}{x_1} = \frac{\alpha_1 \alpha_2 (\mu_1 \alpha_1^r - \mu_2 \alpha_2^r)}{\mu_1 \alpha_1^{r+1} - \mu_2 \alpha_2^{r+1}}$$

bestimmt ist; u und ξ können somit als Koordinaten der Geraden verwendet werden.

Zwei Geraden (u_1, ξ_1) und (u_2, ξ_2) haben, wenn $u_1 \neq u_2$ ist, einen Schnittpunkt, der durch die Koordinaten

$$x = x_1 \frac{u_2 \xi_2 - u_1 \xi_1}{u_2 - u_1}, \quad y = y_1 u_1 u_2 \frac{\xi_2 - \xi_1}{u_2 - u_1}$$

bestimmt ist.

27. Die analytische Geometrie der aus zwei Kurven mit den Grenzgleichungen

$$\frac{y}{y_0} = \left(\frac{x}{x_0}\right)^{r+1}, \quad \frac{y}{y_1} = \left(\frac{x}{x_1}\right)^{r+1}$$

bestehenden Figur ist nun leicht zu übersehen.

Wir wählen (x_0, y_0) und (x_1, y_1) auf derselben Sekante durch O , so dass

$$\frac{x_1}{x_0} = \frac{y_1}{y_0} = i_1;$$

die Gleichungen der Kurven können dann formal als Gleichungen zweier Parabeln

$$y = px^{r+1}, \quad y = qx^{r+1}$$

geschrieben werden, wo

$$p = \frac{y_0}{x_0^{r+1}}, \quad q = \frac{y_1}{x_1^{r+1}}$$

und

$$\frac{p}{q} = \frac{y_0}{y_1} \cdot \left(\frac{x_1}{x_0}\right)^{r+1} = i_1^r.$$

Damit hat man die analytischen Grundlagen der schwachen Geometrie der beiden Kurven in O : Die diesbezüglichen Aufgaben können mittels gewöhnlicher algebraischer Rechnungen gelöst werden, indem man vorgeht, als ob es sich um die elementare analytische Geometrie der beiden Parabeln $y = px^{r+1}$, $y = qx^{r+1}$, wo $\frac{p}{q} = i_1^r$, $i_1 = \left(\frac{p}{q}\right)^{\frac{1}{r}}$, handelte.

28. Es ist jedoch zu beachten, dass die obigen Überlegungen in Wirklichkeit keinerlei Metrik voraussetzen. Die Grenzverhältnisse infinitesimaler Vektoren in Richtung der x -Achse sind rein projektiv bestimmt, und dasselbe gilt für infinitesimale Vektoren in der Richtung der y -Achse, allgemein für Vektoren einer bestimmten Richtung. Für zwei Vektoren verschiedener Richtung hat jedoch das Grenzverhältnis keinen bestimmten Sinn, solange man keine Metrik eingeführt hat. Jede der obengenannten Größen p und q , $p = \frac{y_0}{x_0^{r+1}}$, $q = \frac{y_1}{x_1^{r+1}}$, für sich hat daher nur rein formale, während $\frac{p}{q} = i_1^r$ reale Bedeutung hat. Führt man

aber irgend eine beliebig gewählte Metrik im infinitesimalen Bereiche um O ein, so erhält jede der genannten Grössen für sich eine geometrische Bedeutung; die Endresultate, die Geometrie der vorgelegten schwachen Figur, sind rein projektiver Natur und gelten unabhängig von der speziellen Metrik, die man während der Zwischenrechnung vor Augen gehabt hat. Auch die beiden Parabeln

$$y = px^{r+1}, \quad y = qx^{r+1}$$

haben nur die hierdurch fixierte mehr formale Bedeutung. Sie sollen die oskulierenden Parabeln der Kurven in O und p und q ihre Parameter genannt werden. Das Krümmungsverhältnis ist mit den Parametern, wie schon oben erwähnt, durch die Gleichung $\frac{p}{q} = i_1^r$ verknüpft.

III. Berührungsinvarianten.

29. Die Invariante i_1 ist die erste der sogenannten Berührungsinvarianten unserer beiden Kurven k und k_1 . Sie setzt voraus, dass die durch Zentralprojektion von O aus vermittelte Beziehung der Kurven infinitesimal projektiv in O ist, und gibt dann das Grenzverhältnis entsprechender infinitesimaler Vektoren OA_1 und OA an (vgl. Einl. 2.). Sie ist als Krümmungsverhältnis der Kurven k und k_1 in O bezeichnet worden. Im ordinären Fall, wo i_1 einen endlichen, von Null verschiedenen Wert hat, und den wir, wenn nichts anderes gesagt wird, stets vor Augen haben wollen, haben die Kurven denselben Exponenten r . Besitzen die Kurven gewöhnliche Krümmungen in O , so gibt i_1 deren Verhältnis an. Auch wenn beide Kurven die Krümmung 0 oder ∞ in O haben, das Grenzverhältnis $i_1 = \frac{OA_1}{OA}$ aber noch existiert, wollen wir, wie schon gesagt, an der Bezeichnung Krümmungsverhältnis festhalten.

In allen Fällen ist i_1 eine projektive Invariante. Wir werden nun im folgenden eine Reihe von ähnlichen projektiven Invarianten (Berührungsinvarianten) der beiden Kurven angeben, die sich indessen alle als durch i_1 ausdrückbar erweisen werden.

30. Die erste dieser Invarianten, i_2 , ist dual zu i_1 . Es seien t und t_1 die Tangenten von einem Punkt Q auf x an die Kur-

ven (Fig. 13), R und R_1 ihre Berührungspunkte und u und u_1 ihre Winkel mit x . Dann wird i_2 als Grenzverhältnis von u_1 und u für $Q \rightarrow O$ definiert. Da die Kurven denselben Index in O haben, hat man

$$i_2 = \frac{u_1}{u} = \frac{v_1}{v} = \left(\frac{OS}{OR}\right)^r = \left(\frac{OS}{OR_1}\right)^r,$$

also

$$i_2 = \frac{1}{i_1^r}.$$

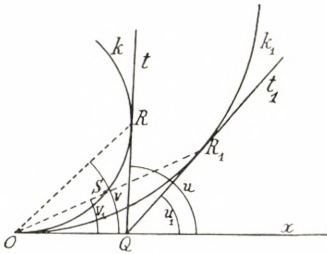


Fig. 13.

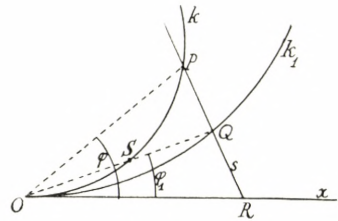


Fig. 14.

31. Die nächste anzuführende Invariante wird gebildet, indem man die Kurven und ihre gemeinsame Tangente in O mit einer Geraden s schneidet (Fig. 14), die gegen den Punkt O konvergiert, im übrigen aber nur der Bedingung unterworfen ist, dass sie keine Grenzlage auf x hat. Die variablen Schnittpunkte P , Q , R konvergieren dann alle gegen O . Die betreffende Invariante i_3 (Segresche Invariante) wird dann als das Grenzverhältnis

$$i_3 = \frac{RQ}{RP} = \frac{\varphi_1}{\varphi} = \left(\frac{OS}{OP}\right)^r = \left(\frac{OS}{OQ}\right)^r$$

definiert, so dass

$$i_3 = \frac{1}{i_1^r} = i_2$$

gilt.

32. Die zu i_3 duale Invariante sei i_4 . Sie wird gebildet, indem von einem Punkt T (Fig. 15), der gegen die zu x entgegengesetzte Halbgerade x_1 konvergiert, aber keine Grenzlage in O hat, die Tangenten t und t_1 an die Kurven legt. Diese Tangenten schneiden x in P und P_1 , und i_4 ist dann das Grenzverhältnis

$$i_4 = \frac{OP_1}{OP}.$$

Da die Winkel u_1 und u , und folglich auch φ_1 und φ das Grenzverhältnis 1 haben, sieht man jedoch unmittelbar, dass

$$i_4 = i_1$$

ist.

Welchen Grenzlagen U auf der Halbgeraden x_1 der Punkt T zustrebt, ist ohne Einfluss auf das Resultat.

Man sieht ferner, dass das Grenzverhältnis der infinitesimalen Winkel (st_1) und (st) , wo s die Gerade OT bezeichnet, denselben Wert i_4 hat.

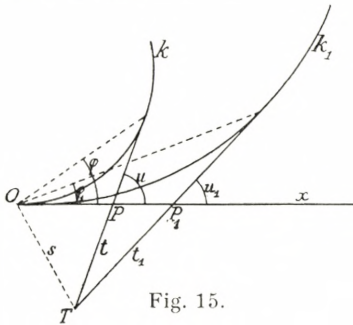


Fig. 15.

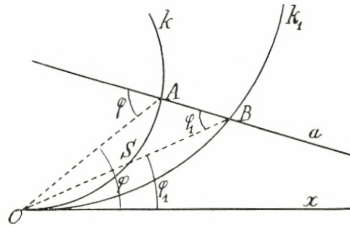


Fig. 16.

33. Endlich seien noch zwei zueinander duale Invarianten i_5 und i_6 genannt. Die erstere ist das Grenzverhältnis zweier infinitesimaler Vektoren OB und OA , wo A und B die Schnittpunkte der Kurven mit einer Geraden a sind (Fig. 16), die in der Weise gegen x konvergiert, dass O nicht Grenzlage von Schnittpunkten der Geraden a und x ist. Man erhält sofort

$$i_5 = \frac{OB}{OA} = \frac{\varphi}{\varphi_1}$$

und weiter

$$\frac{\varphi}{\varphi_1} = \left(\frac{OA}{OS}\right)^r = \left(\frac{OA}{OB} \cdot \frac{OB}{OS}\right)^r,$$

$$i_5 = \left(\frac{1}{i_5} \cdot i_1\right)^r$$

oder

$$i_5 = i_1^{\frac{r}{r+1}}.$$

Für $r = 1$ ergibt sich speziell $i_5 = \sqrt{i_1}$.

34. Die duale Invariante i_6 entsteht folgendermassen: Von einem Punkt T (Fig. 17), der derart gegen O konvergiert, dass keine Grenzlage der Geraden OT auf x fällt, werden die Tangenten t und t_1 an die Kurven gelegt. Es seien R und R_1 ihre Berührungspunkte und P und P_1 ihre Schnittpunkte mit x . Die Invariante i_6 ist dann das Grenzverhältnis

$$i_6 = \frac{u_1}{u}$$

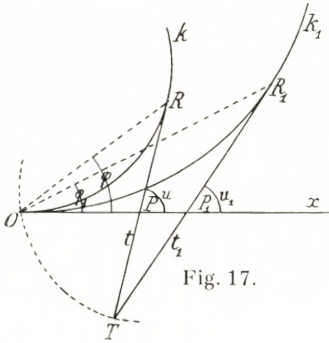


Fig. 17.

der Winkel $u_1 = (xt_1)$ und $u = (xt)$. Aus dem schwachen Dreieck TPP_1

entnimmt man, dass dieses Verhältnis auch durch $\frac{TP}{TP_1}$, also auch durch $\frac{OP}{OP_1}$ oder (da die Kurven denselben Index haben) durch $\frac{OR}{OR_1}$ dargestellt wird. In üblicher Weise erhält man daher

$$i_6 = \frac{u_1}{u} = \frac{\varphi_1}{\varphi} = \left(\frac{OS}{OR}\right)^r = \left(\frac{OS}{OR_1} \cdot \frac{OR_1}{OR}\right)^r$$

oder

$$i_6 = \left(\frac{1}{i_1} \cdot \frac{1}{i_6}\right)^r,$$

$$i_6 = \frac{1}{i_1^{r+1}} = \frac{1}{i_5}.$$

Ferner sei bemerkt, dass die drei Grenzverhältnisse

$$\frac{TP}{TP_1}, \quad \frac{OP}{OP_1}, \quad \frac{PR}{P_1R_1}$$

einander gleich sind und dass dem zufolge $\frac{TR}{TR_1}$ denselben Wert hat.

Die Invariante i_6 ist also auch gleich dem Grenzverhältnis $\frac{TR}{TR_1}$ der Längen der Tangenten von T an die Kurven.

35. Wir wollen noch die Relation zwischen den Schnittpunkten der beiden Kurven mit einer beliebigen, gegen x kon-

vergiierenden Sekante s angeben (Fig. 18). Die Sekante schneide die Kurve k in den Punkten P und Q , die derart gegen O konvergieren, dass das Grenzverhältnis $\frac{OQ}{OP} = \alpha (> 1)$ ist. Sie schneide die andre Kurve k_1 in den Punkten R und S , die durch die Grenzverhältnisse

$$\xi_1 = \frac{OR}{OP}, \quad \xi_2 = \frac{OS}{OP}$$

bestimmt seien. Eine gewöhnliche Rechnung zeigt dann, dass diese die Gleichung

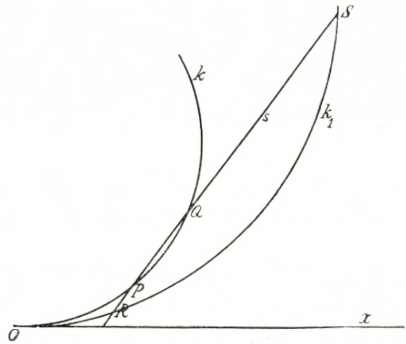


Fig. 18.

$$\xi^{r+1} - i_1^r \frac{\alpha^{r+1} - 1}{\alpha - 1} \xi + i_1^r \frac{\alpha^{r+1} - \alpha}{\alpha - 1} = 0$$

befriedigen, die genau zwei reelle Wurzeln hat.

Im Fall $i_1 > 1$ (der in der Figur wiedergegeben ist) ist die eine (dem Punkt R entsprechende) Wurzel kleiner als 1 und die andre (dem Punkt S entsprechende) grösser als 1. Für $i_1 = 1$ sind die Wurzeln $\xi_1 = 1, \xi_2 = \alpha$.

Ist s Tangente an k , fallen also P und Q zusammen (Fig. 19), so geht die Gleichung in

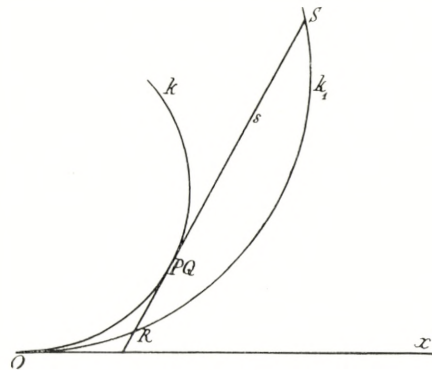


Fig. 19.

$$\xi^{r+1} - i_1^r (r + 1) \xi + i_1^r r = 0$$

über. Ist hier $i_1 = 1$, so ist sowohl ξ_1 als auch ξ_2 gleich 1.

36. Bisher haben wir von den beiden betrachteten konvexen Kurven nicht nur vorausgesetzt, dass sie denselben Anfangspunkt O , sondern auch, dass sie dort dieselbe Halbtangente x haben und auf derselben Seite derselben liegen. Nun kann man aber mittels einer harmonischen Homologie mit der Tangente als Achse und einem nicht auf dieser gelegenen Zentrum oder mit einer durch O gehenden Achse und

auf der Tangente gelegenen Zentrum oder schliesslich mit O als Zentrum und einer nicht durch O gehenden Achse die eine der Kurven in eine neue überführen, ohne dass die Grenzverhältnisse der mit der Kurve verknüpften infinitesimalen Vektoren mit dem Anfangspunkt O geändert werden. Hiernach ist klar, dass die Betrachtungen auf solche konvexen Kurven mit demselben Anfangspunkt O und derselben Tangente in O übertragen werden können, welche die zusätzliche Bedingung, dieselbe Halbtangente zu haben und auf derselben Seite zu liegen, nicht erfüllen. In

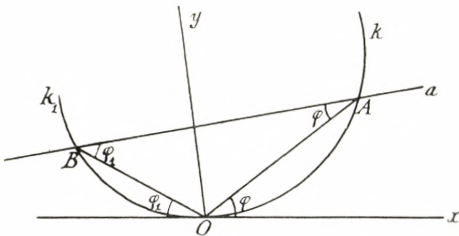


Fig. 20.

den alsdann möglichen Fällen, wo die beiden Kurven zusammen eine Kurve mit O als Konvexitätspunkt, mit O als Spitze erster Art oder mit O als Wendepunkt ausmachen, kann man wie früher von einer Invariante i_1 , dem Krümmungsverhältnis, sprechen. Man könnte ihm ein Vorzeichen beilegen, wozu wir hier jedoch keinen Anlass haben.

37. Wir betrachten zunächst den Fall, wo die Kurven k und k_1 zusammen einen konvexen Bogen ausmachen (Fig. 20). Eine variable Gerade a , die derart gegen x konvergiert, dass ihr Schnittpunkt mit x keine Grenzlage in O hat, schneidet k und k_1 in Punkten A und B , die beide gegen O konvergieren. Wir suchen das Grenzverhältnis

$$\frac{OB}{OA} = \frac{\varphi}{\varphi_1}$$

der Vektoren OB und OA . Unter Heranziehung der zu k_1 symmetrischen Kurve mit Bezug auf eine Gerade y erhält man wie in 33.

$$\frac{OB}{OA} = \frac{r}{i_1^{r+1}},$$

wo i_1 das Krümmungsverhältnis von k und k_1 ist.

Im Anschluss an 34. bemerken wir ferner, dass das Grenzverhältnis der Längen der Tangenten TR und TR_1 an die Kurven

von einem Punkt T aus, der derart gegen O konvergiert, dass keine Grenzlage von OT mit x zusammenfällt, sich zu

$$i_0 = \frac{1}{r} \frac{1}{i_1^{r+1}}$$

ergibt.

38. Konvergiert a dagegen in der Weise gegen x , dass der Schnittpunkt P mit x gegen O geht (Fig. 21), so hängt das

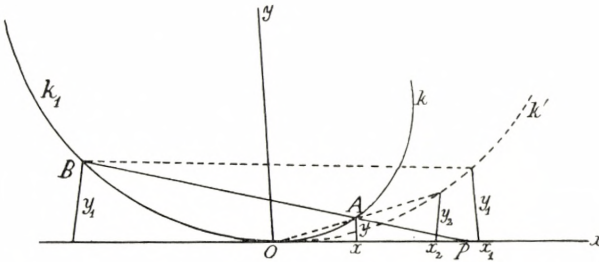


Fig. 21.

gesuchte Grenzverhältnis von $\frac{PB}{PA}$ ab. Ist das letztere Grenzverhältnis gleich μ , so erhält man unter Heranziehung der zu k_1 mit Bezug auf y symmetrischen Kurve k'

$$\frac{y_2}{y_1} = \left(\frac{x_2}{x_1}\right)^{r+1};$$

und wegen

$$\frac{y_2}{y} = \frac{x_2}{x} = i_1, \quad \frac{y_1}{y} = \mu$$

entnimmt man hieraus

$$\frac{i_1}{\mu} = \left(i_1 \frac{x}{x_1}\right)^{r+1}$$

oder

$$\frac{OB}{OA} = \frac{x_1}{x} = \mu^{\frac{1}{r+1}} i_1^{\frac{r}{r+1}}.$$

Das Resultat von 37. ist hierin für $\mu = 1$ enthalten.

Im einfachsten Fall $r = 1$, $i_1 = 1$, $\mu = 1$ sind die beiden Vektoren OB und OA einander gleich.

Ist $r = 1$, $i_1 = 1$, μ beliebig, so ergibt sich

$$\frac{OB}{OA} = \sqrt{\mu}.$$

In leicht verständlicher Schreibweise hat man dann die Gleichung

$$PO^2 = PA \cdot PB,$$

die als Ausdruck für »den infinitesimalen Satz über die Potenz eines Punktes« aufgefasst werden kann.

39. Wir nehmen nun an (Fig. 22), dass wir zwei Kurvenzweige k und k_1 haben, die einen Wendepunkt in O bilden und ein bestimmtes Krümmungsverhältnis

$$i_1 = \frac{OA_1}{OA}$$

haben, wobei die Strecken ohne Vorzeichen gerechnet werden. Von den beiden Zweigen wird angenommen, dass sie den Exponenten r in O haben. Von einem Punkt T der Sekante OAA_1 sind die Tangenten an die beiden Kurvenzweige gezogen; die

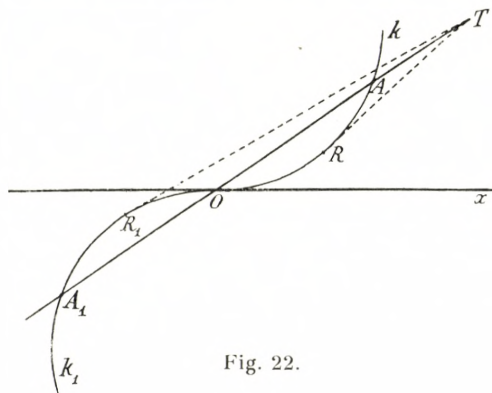


Fig. 22.

Berührungspunkte seien R und R_1 . Der Punkt T konvergiere gegen x , jedoch so, dass er keine Grenzlage in O hat. Die Punkte A und A_1 , R und R_1 konvergieren alle gegen O , so dass die

Grenzverhältnisse $\frac{OA}{OR}$ und $\frac{OA_1}{OR_1}$ nach 21. denselben Wert $(1+r)^{\frac{1}{r}}$ haben, woraus folgt, dass das Grenzverhältnis

$$\frac{OR_1}{OR} = \frac{OA_1}{OA} = i_1$$

ist.

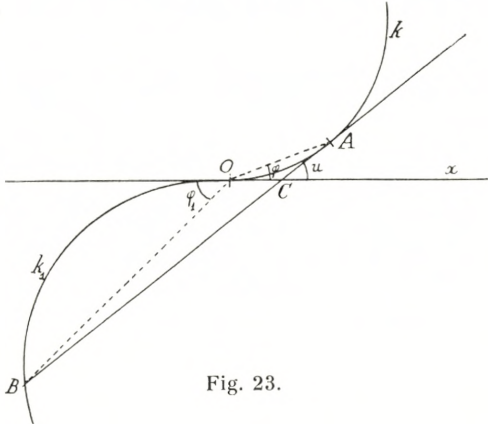


Fig. 23.

Falls T gegen O konvergiert, das Grenzverhältnis $\frac{OT}{OA}$ einen bestimmten vorgegebenen Wert hat und das Grenzverhältnis $\frac{OT}{OA_1}$ somit gleichfalls bekannt ist, kann man die Resultate von 22. heranziehen.

40. Die Tangente an k im Punkte A schneide nun k_1 in B und x in C (Fig. 23). Wir wollen das Grenzverhältnis $\xi = \frac{OB}{OA}$ bestimmen, wenn A , und damit auch B , gegen O konvergiert.

Man hat

$$\frac{OC}{OB} = \begin{cases} \frac{OC}{OA} \cdot \frac{OA}{OB} = \frac{r}{r+1} \cdot \frac{1}{\xi} \\ \frac{\varphi_1 - u}{u} = \frac{\varphi_1}{\varphi} \cdot \frac{\varphi}{u} - 1 = \left(\frac{1}{i_1 \xi}\right)^r \frac{1}{r+1} - 1, \end{cases}$$

woraus folgt

$$\xi^{r+1} - i_1^r (r+1) \xi - i_1^r r = 0.$$

Diese Gleichung hat genau eine positive Wurzel, und ξ ist also durch sie bestimmt.

Das Verhältnis, in dem das Stück der Tangente zwischen dem Berührungspunkt und dem Schnittpunkt von der Wendetangente geteilt wird, ist

$$\frac{BC}{CA} = (r+1)\xi + r.$$

Aus der obigen Gleichung folgt dann

$$\left(\frac{OB}{OA}\right)^{r+1} : \frac{BC}{CA} = i_1^r.$$

41. Wir betrachten wieder zwei Kurvenzweige k, k_1 (Fig. 24), die in O einen Wendepunkt mit der Tangente x bilden und dort den Krümmungsexponenten r und das Krümmungsverhältnis i_1 haben. Auf k werden zwei Punkte A und B gewählt, die in der Weise gegen O konvergieren, dass das Grenzverhältnis $\frac{OB}{OA} = \alpha (> 1)$ ist. Die Gerade $AB = s$ schneidet k_1 im Punkt C ,

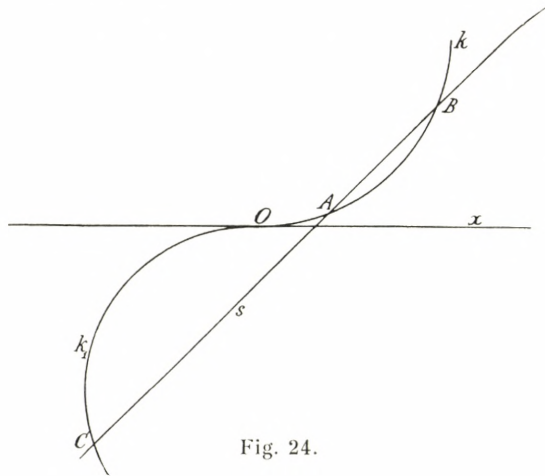


Fig. 24.

der dann gleichfalls gegen O konvergiert. Man sucht den Wert des Grenzverhältnisses $\xi = \frac{OC}{OA}$. Man erhält hierfür die Gleichung

$$\xi^{r+1} - i_1^r \frac{\alpha^{r+1} - 1}{\alpha - 1} \xi - i_1^r \frac{\alpha^{r+1} - \alpha}{\alpha - 1} = 0.$$

Hiermit hat man die Grenzrelation zwischen den drei Schnittpunkten A, B, C . Lässt man α gegen 1 gehen, so erhält man die vorige Gleichung.

Unterwirft man s der Bedingung, dass das Grenzverhältnis $\frac{CA}{AB}$ einen bestimmten Wert μ haben soll, so hat man die Relation

$$\xi = \frac{OC}{OA} = \frac{CA - OA}{OA} = \mu\alpha - (\mu + 1),$$

die zusammen mit der obigen Gleichung zur Bestimmung von ξ und α dient.

42. Analoge Aufgaben hat man, wenn eine Spitze erster Art vorliegt (Fig. 25). Die Tangente in einem Punkt A von k schneide k_1 in B . Man sucht das Grenzverhältnis $\xi = \frac{OB}{OA}$, wenn A , und damit B , gegen O konvergiert. Zur Bestimmung von ξ findet man die Gleichung

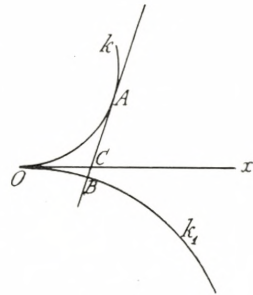


Fig. 25.

$$\xi^{r+1} + i_1^r (r + 1) \xi + i_1^r r = 0,$$

die genau eine positive Wurzel $\xi_1 \leq \frac{r}{r+1}$ besitzt. Das Teilverhältnis $\frac{BC}{CA}$ ergibt sich zu $r - (r + 1) \xi_1$. Man hat dann

$$\left(\frac{OB}{OA}\right)^{r+1} : \frac{BC}{CA} = i_1^r.$$

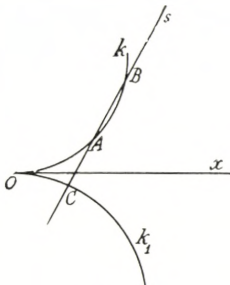


Fig. 26.

Man betrachte nun eine Sekante s (Fig. 26), die den einen Zweig in A und B , den anderen in C schneidet, so dass das Grenzverhältnis $\frac{OB}{OA} = \alpha$ ist. Zur Bestimmung des Grenzverhältnisses $\xi = \frac{OC}{OA}$ hat man dann

$$\xi^{r+1} + i_1^r \frac{\alpha^{r+1} - 1}{\alpha - 1} \xi - i_1^r \frac{\alpha^{r+1} - \alpha}{\alpha - 1} = 0.$$

43. Bei der Untersuchung schwacher Figuren ist es von Wichtigkeit, bestimmte Regeln für die Anwendung des

Dualitätsprinzips zu haben. Wir wollen daher einige Worte darüber hinzufügen. In einem ebenen Polarsystem entspricht einem schwachen Dreieck ein schwaches Dreieck, und da die beiden Dreiecke homolog sind, sind entsprechende Stücke proportional. Einem konvexen Bogen mit den Tangenten a, b in den Endpunkten A, B (Fig. 27) entspricht im Polarsystem ein konvexer Bogen mit den Tangenten a_1, b_1 in den Endpunkten A_1, B_1 (Fig. 28), derart dass a_1 und b_1 die Polaren von A und B und A_1 und B_1 die Pole von a und b sind. Konvergiert B, b

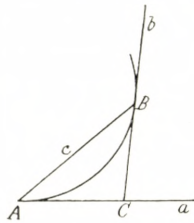


Fig. 27.

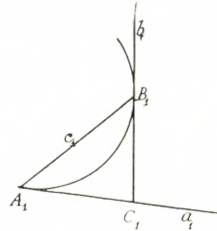


Fig. 28.

gegen A, a , so konvergiert auch B_1, b_1 gegen A_1, a_1 . Dem Grenzverhältnis $\frac{BC}{AC}$ der einen Figur entspricht das Grenzverhältnis $\frac{b_1c_1}{A_1C_1} = \frac{A_1C_1}{B_1C_1}$ der anderen, und hieraus folgt, dass, wenn der eine Bogen einen bestimmten Index in A hat, auch der andere einen Index in A_1 hat, und zwar den reziproken.

Ferner erkennt man mittels des genannten Grenzüberganges folgendes: falls der Vektor AB und der Winkel ab hyperprojektiv sind, so gilt dies auch für den Winkel a_1b_1 und den Vektor A_1B_1 , d. h. wenn der eine Bogen einen bestimmten Exponenten r in A hat, so hat der andere einen bestimmten Exponenten r_1 im entsprechenden Punkt A_1 , und da die beiden Indices reziprok sind, gilt dies auch für r und r_1 .

Mit Hilfe dieser Bemerkungen kann man aus den Resultaten von 24., 35., 39., 40., 41., 42. neue Relationen herleiten sowie Zusammenhänge zwischen früher abgeleiteten Relationen unmittelbar erkennen.

IV. Indexrelationen.

44. Im folgenden beschäftigen wir uns speziell mit Relationen, die an den Indexbegriff anknüpfen, ohne Rücksicht darauf, ob

in dem betreffenden Kurvenpunkt ein bestimmter Krümmungsexponent existiert oder nicht.

Die beiden vom Punkt A ausgehenden konvexen Kurven l und m (Fig. 29) mögen eine Spitze zweiter Art bilden, derart dass m von l umschlossen wird. Die Tangente b an m im Punkte B schneide die gemeinsame Tangente a in D sowie die Kurve l in C_1 und C_2 . Die Tangenten c_1 und c_2 an l in C_1 bzw. C_2 schneiden a in E bzw. F und einander in G . Konvergiert B gegen A , so konvergieren die Punkte C_1, C_2, D, E, F, G sämtlich gegen A und ihre Verbindungsgeraden gegen a . Wir suchen eine »Repräsentantenfigur« für die hierdurch bestimmte schwache Figur, wobei wir voraussetzen, dass die Kurven bestimmte ordinäre Indices λ und μ in A haben und dass bestimmte Grenzverhältnisse

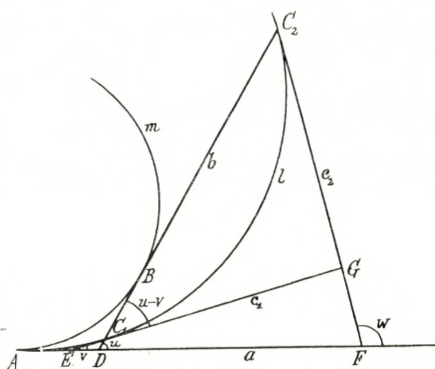


Fig. 29.

existieren. Die Indizes der Kurven geben direkt die beiden Grenzverhältnisse

$$\xi_1 = \frac{AC_1}{AB}, \quad \xi_2 = \frac{AC_2}{AB}$$

existieren. Die Indizes der Kurven geben direkt die beiden Grenzverhältnisse

$$\lambda = \frac{EC_1}{AE} \text{ oder } \frac{FC_2}{AF}, \quad \mu = \frac{DB}{AD}$$

an.

Ein Diagramm aller Grenzverhältnisse erhält man, indem man auf einer Zahlengeraden die Punkte A, B, C_1, C_2, D, E, F in der Weise markiert, dass die Verhältnisse der durch die Abszissendifferenzen bestimmten Intervalle die betreffenden Grenz-

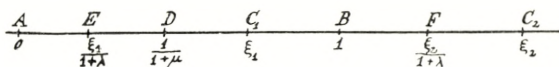


Fig. 30.

verhältnisse angeben. Auf diese Weise findet man die folgende Repräsentantenfigur (Fig. 30) mit den angegebenen Abszissen,

wobei den Punkten A und B die Abszissen 0 und 1 zugeschrieben sind.

Der Punkt G ist nicht angegeben. Seine Abszisse ergibt sich mit Hilfe des Satzes von Menelaos zu

$$AG = \frac{1 + \mu}{1 + \lambda} \xi_1 \xi_2,$$

womit man zugleich das Grenzverhältnis

$$\frac{AG}{AB} = \frac{1 + \mu}{1 + \lambda} \xi_1 \xi_2$$

hat.

Alle Grenzverhältnisse von Strecken und Winkeln der vorliegenden schwachen Figur können dann an der Repräsentantenfigur abgelesen werden. Man erhält so

$$\eta_1 = \frac{v}{u} = \frac{DC_1}{EC_1} = \frac{\lambda + 1}{\lambda} \left(1 - \frac{1}{1 + \mu} \cdot \frac{1}{\xi_1} \right),$$

$$\eta_2 = \frac{w}{u} = \frac{DC_2}{FC_2} = \frac{\lambda + 1}{\lambda} \left(1 - \frac{1}{1 + \mu} \cdot \frac{1}{\xi_2} \right),$$

$$\frac{u - v}{w - u} = \frac{GC_2}{C_1G} = \frac{1 - \eta_1}{\eta_2 - 1} = \frac{\frac{1}{\xi_1} - \frac{1 + \mu}{1 + \lambda}}{1 + \mu - \frac{1}{\xi_2}}.$$

Wegen $1 \geq \eta_1 \geq 0$ hat man

$$\frac{1 + \lambda}{1 + \mu} \geq \xi_1 \geq \frac{1}{1 + \mu}.$$

In ähnlicher Weise ergibt sich

$$\xi_2 \geq \frac{1 + \lambda}{1 + \mu}.$$

Wenn ein konvexer Bogen, der von A ausgeht und innerhalb der Spitze zwischen l und m verläuft, einen Index in A hat, so liegt dieser zwischen

$$\xi_1 (1 + \mu) - 1 \quad \text{und} \quad \xi_2 (1 + \mu) - 1.$$

Die Indizes λ und μ der Kurven l und m selbst gehören gleichfalls diesem Intervall an.

Haben die Kurven insbesondere denselben Exponenten r in A und eine Invariante $i_1 = 1$, so gilt $\xi_1 = \xi_2 = 1$, $\lambda = \mu = \frac{1}{r}$, und das Diagramm reduziert sich in der Weise, dass $E = D = F = \frac{r}{r+1}$, $C_1 = B = C_2 = G = 1$ wird.

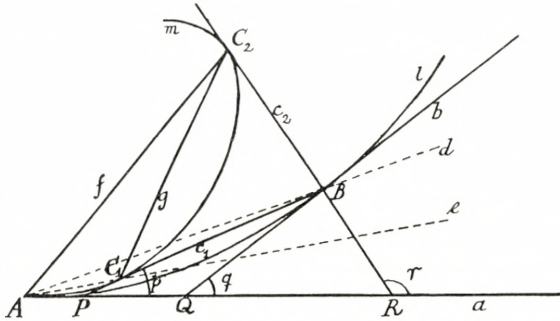


Fig. 31.

45. Wendet man das Dualitätsprinzip an, so erhält man die folgende Figur:

Von einem Punkt B auf l (Fig. 31) lege man an m die Tangenten c_1 und c_2 mit den Berührungspunkten C_1 und C_2 . Die Geraden a, b, c_1, c_2, d, e, f und g entsprechen den Punkten A, B, C_1, C_2, D, E, F und G der Fig. 29 dual.

Dem Grenzverhältnis ξ_1 der Fig. 29 entspricht hier

$$\frac{ac_1}{ab} = \frac{p}{q} = \frac{QB}{PB} = \frac{1}{\eta_2},$$

und ebenso entspricht ξ_2

$$\frac{ac_2}{ab} = \frac{r}{q} = \frac{QB}{RB} = \frac{1}{\eta_1}.$$

Umgekehrt entspricht η_1

$$\frac{AC_1}{AB} = \frac{1}{\xi_2}$$

und η_2

$$\frac{AC_2}{AB} = \frac{1}{\xi_1}.$$

Ferner hat man

$$\frac{C_1B}{C_2B} = \frac{gc_2}{c_1g} = \begin{cases} 1 - \frac{1}{\xi_2} \\ \frac{1}{\xi_1} - 1 \\ \frac{\mu(\lambda + 1)}{\lambda(\mu + 1)} \cdot \frac{1}{\eta_1\eta_2} \end{cases}$$

Beim Übergang zur dualen Figur sind folgende Vertauschungen vorzunehmen:

$$\lambda \rightarrow \frac{1}{\mu}, \quad \mu \rightarrow \frac{1}{\lambda}, \quad \xi_1 \rightarrow \frac{1}{\eta_2}, \quad \xi_2 \rightarrow \frac{1}{\eta_1}, \quad \eta_1 \rightarrow \frac{1}{\xi_2}, \quad \eta_2 \rightarrow \frac{1}{\xi_1}.$$

46. Ein schwaches Dreieck PQR , von dem zwei Ecken P und Q auf l liegen, sei der Kurve m umschrieben. Die Seiten

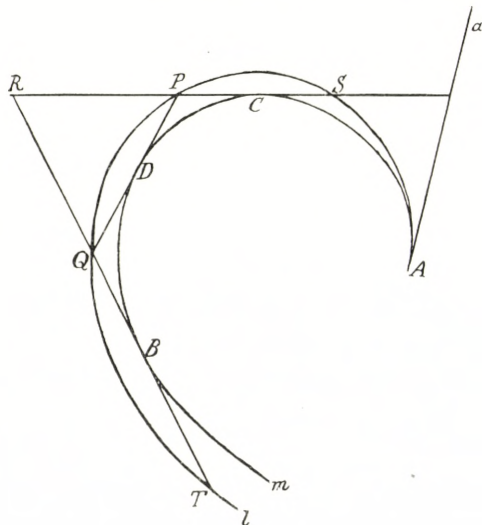


Fig. 32.

mögen m in den Punkten B, C, D (Fig. 32) berühren. Man kann dann alle Grenzverhältnisse in Abhängigkeit von den oben genannten Grössen ermitteln. Man hat

λ von l höchstens gleich dem Index μ von m ist, wenn m von l umschlossen wird.

Dieses wichtige Resultat führt sofort zu dem Satz, dass die Kurve AB in A keine ordinäre Krümmung haben kann, wenn der Index in A grösser als 1 ist. Jeder in A berührende Kegelschnitt muss dann nämlich die Kurve umschliessen. Ist der Index kleiner als 1, so wird hingegen jeder berührende Kegelschnitt von der Kurve umschlossen. Nur wenn der Index 1 ist, besteht die Möglichkeit ordinärer Krümmung. Es braucht jedoch keine bestimmte ordinäre Krümmung zu existieren, aber:

Hat die Kurve ordinäre Krümmung, so ist der Index gleich 1.

Allgemeiner: Hat die Kurve keine singuläre Krümmung (d. h. hat sie beschränkte Krümmung), so ist der Index 1.

Wenn die Kurve singuläre Krümmung hat, so braucht der Index nicht 1 zu sein. Beispiel: $y = e^{-\frac{1}{x}}$, wo der Index 0 ist.

48. Haben die beiden Kurven l und m denselben Index, so hat auch jede zwischen ihnen verlaufende konvexe Kurve diesen Index, vorausgesetzt, dass sie überhaupt einen bestimmten Index hat. Dass dieses nicht der Fall zu sein braucht, entnimmt man daraus, dass ein konvexer Streckenzug $PQR \dots$, der in l eingeschrieben und m umschrieben ist, keinen bestimmten Index hat, da die betreffenden Grenzwerte zwischen $\xi_1(1 + \mu) - 1$ und $\xi_2(1 + \mu) - 1$ oszillieren (vgl. 44.). Ist dagegen zugleich $\xi_1 = \xi_2$, d. h. $i_1 = 1$, so hat der genannte Streckenzug einen bestimmten Index, und zwar denselben wie die gegebenen Kurven. In diesem Fall ist es auch klar, dass alle konvexen Kurven, die innerhalb der Spitze verlaufen, den gleichen Index und das Krümmungsverhältnis 1 zu den gegebenen Kurven haben.

Hat eine konvexe Kurve k im Anfangspunkt A eine bestimmte oskulierende Parabel $y = ax^r$, die nicht entartet ist ($a > 0$, $r > 1$), so hat k im Punkte den Krümmungsexponenten $r - 1$ und den Index $\frac{1}{r - 1}$. Hieraus folgt speziell für $r = 2$, dass ein konvexer Bogen, der in einem Punkte eindeutige, nicht ausgeartete Krümmung besitzt, in diesem Punkt den Index 1 hat, woraus sich insbesondere ergibt, dass der Schmiegunskreis mit dem Krümmungskreis zusammenfällt.

Mit Hilfe der obigen Konstruktion findet man konvexe Streckenzüge als Beispiele von Kurven mit beliebig (auch 0 oder ∞) vorgeschriebenem Index.

V. Indextheorie der Raumkurven.

49. Die im folgenden zu betrachtenden Raumkurven werden stückweise monoton angenommen. Einleitend beschränken wir uns auf die Betrachtung eines monotonen Kurvenzweiges, der von dem Punkt A ausgeht, dessen Umgebung untersucht werden soll. Bei der daran anschließenden Untersuchung innerer Kurvenpunkte werden die beiden von dem betreffenden Punkt ausgehenden Zweige einzeln betrachtet.

Das Verhalten des zu betrachtenden Kurvenzweiges $A \cdots$ im Punkte A betreffend, gehen wir davon aus, dass die folgenden drei Indexzahlen λ , μ , ν existieren und von 0 und ∞ verschieden sind:

1°. Der Index λ der Raumkurve im Punkte A . Darunter wird das Grenzverhältnis für $B \rightarrow A$ der beiden infinitesimalen Winkel α und β verstanden, die die Sehne AB von A bis zu einem Nachbarpunkt B mit den Tangenten in diesen Punkten bildet. Die Zahl λ ist zugleich Index einer jeden Zentralprojektion von einem beliebigen Projektionszentrum O aus, das nicht der Schmiegebene in A angehört, oder, was auf dasselbe hinauskommt, Index des projizierenden Kegels in der Erzeugenden OA . Dies entnimmt man sofort daraus, dass das Grenzverhältnis der Winkel α und β gleich dem Grenzverhältnis ihrer Projektionen ist.

2°. Der Index μ des von A aus projizierenden Kegels in der Kurventangente a in A . Diese Zahl stimmt mit dem Index einer jeden ebenen Leitkurve des Kegels in dem A entsprechenden Punkt auf a überein.

3°. Der Index ν einer ebenen Leitkurve der Raumkurve (Schnitt der Tangentenfläche der Raumkurve mit einer nicht durch A gehenden Ebene) in dem A entsprechenden Punkt P auf a . Die Zahl ν erweist sich als unabhängig von der Wahl der ebenen Leitkurve. Es seien nämlich A und B zwei Nachbarpunkte mit den Tangenten a und b (Fig. 34). Die Tangenten-

fläche werde von zwei Ebenen, die, wenn sie a in zwei verschiedenen Punkten P, P_1 schneiden, in diesem Zusammenhang als parallel angesehen werden können, in zwei Bögen PQ und P_1Q_1 mit den paarweise parallelen Tangenten PT und QT, P_1T_1 und Q_1T_1 geschnitten, wo T und T_1 auf der Schnittgeraden t der Schmiegeebenen in A und B liegen. Die Tangenten a und b werden von t in zwei Punkten A_1 und B_1 geschnitten, die für

$B \rightarrow A$ gegen A konvergieren. Man hat dann die Grenzverhältnisgleichungen

$$\frac{TQ}{T_1Q_1} = \frac{B_1T}{B_1T_1}, \quad \frac{PT}{P_1T_1} = \frac{A_1T}{A_1T_1},$$

woraus

$$\frac{TQ}{PT} = \frac{T_1Q_1}{P_1T_1}$$

folgt, was eben zeigt, dass ν von der Lage der Schnittebene unabhängig ist, wenn diese nur nicht durch A geht.

Schneiden die Ebenen die Erzeugende a im selben Punkt

P , so zeigt eine direkte Betrachtung des Grenzüberganges, dass die beiden Schnittkurven denselben Index in diesem Punkt haben.

Insbesondere kann man (in einer metrischen oder auch nur affinen Darstellung) die Ebene der Leitkurve als Bezugsebene (unendlich ferne Ebene) verwenden, so dass ν bei dieser Interpretation als Index des Richtungskegels bezeichnet werden kann.

Von den drei Zahlen λ, μ, ν ausgehend, kann man nun die infinitesimalen Eigenschaften der Kurve in dem betreffenden Punkt studieren.

Bemerkung I: Die obigen Beweise gelten auch für nicht monotone Kurven, wenn nur A_1 und B_1 mit B gegen A konvergieren.

Bemerkung II: An die Stelle ebener Leitkurven können auch monotone räumliche Leitkurven treten, die a in von A verschiedenen Punkten schneiden und deren Schmiegeebene in einem

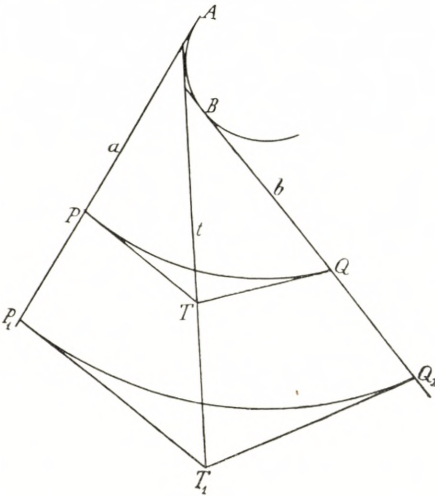


Fig. 34.

Schnittpunkt mit a von der Schmiegeebene der gegebenen Kurve in A verschieden ist. Alle derartigen Leitkurven haben in ihrem Schnittpunkt mit a denselben Index, nämlich ν .

Bemerkung III: Die Beschränkung unserer Untersuchungen auf monotone Kurven erweist sich somit als unnötig. Es wäre hinreichend gewesen voraussetzen, dass die Kurven Charakteristiken sind, d. h. Kurven mit überallgewöhnlicher, stetig variierender Tangente und mit der Eigenschaft, dass jede feste Ebene durch die Tangente a eines festen Kurvenpunktes A von der Tangente b eines variablen Kurvenpunktes B , der gegen A konvergiert, in einem Punkt mit der Grenzlage A geschnitten wird¹.

Aus konventionellen Gründen wollen wir hier jedoch daran festhalten, dass die zu betrachtenden Kurven monoton sind.

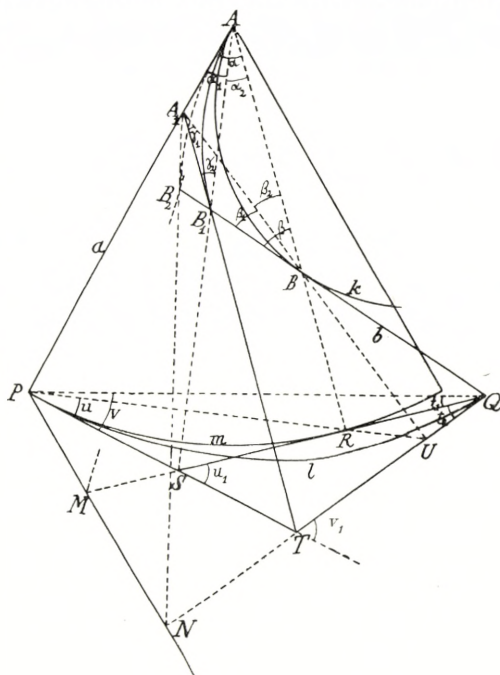


Fig. 35.

50. Fig. 35 zeigt die Raumkurve k mit der Punktschar $A \cdot \cdot B \cdot \cdot \cdot$ und der Tangentenschar $a \cdot \cdot b \cdot \cdot \cdot$, die Leitkurve l in der Ebene σ mit der Punktschar $P \cdot \cdot Q \cdot \cdot \cdot$ sowie den projizierenden Kegel $A(k)$ mit der Leitkurve $m = P \cdot \cdot R \cdot \cdot \cdot$ in der Ebene σ . Die Schmiegeebene in A hat die Spur PST in σ , und B_1 sei der Schnittpunkt der Tangente b mit dieser Schmiegeebene. Die Tangentenfläche von k schneidet die Schmiegeebene (des Punktes A) in der Kurve $k^0 = A \cdot \cdot B_1$, die in B_1 die Tangente A_1B_1T hat; TUQ ist die Spur der Schmiegeebene des Punktes B . Die beiden

¹ Vgl. hierzu die Abhandlung des Verfassers: Eksempler paa geometriske Undersøgelser over Integralkurver i Rummet. (Festskrift til J. F. Steffensen, København 1943).

Schmiegebenen schneiden einander in der Geraden A_1B_1 (welche Tangente an k^0 in B_1 ist). Das Viereck AA_1B_1B ist dann beim Grenzübergang $B \rightarrow A$ schwach. Seine Seiten streben der Richtung nach gegen a und der Grösse nach gegen 0, und gleichzeitig gehen seine infinitesimalen Winkel, für die wir die in der Figur angegebenen Bezeichnungen einführen, gegen 0. Alle Seitenebenen des Tetraeders AA_1B_1B konvergieren gegen die Schmiegeebene in A und alle infinitesimalen Flächenwinkel gegen 0. Die Kanten des Tetraeders werden von der Ebene σ in den Ecken P, Q, R, S, T, U eines schwachen vollständigen Vierseits geschnitten.

Alle Grenzverhältnisse des Vierecks AA_1B_1B und des soeben genannten Vierseits $PQRSTU$ können nun durch die Zahlen λ, μ, ν ausgedrückt werden.

Zur Konstruktion einer Repräsentantenfigur hat man die Verhältnisse

$$\frac{PR}{RQ} = \lambda, \quad \frac{SR}{PS} = \mu, \quad \frac{TQ}{PT} = \nu$$

zur Verfügung, und durch sie ist die ganze Figur bestimmt, da beispielsweise die Abszisse von U sich daraus ergibt, dass die Punktenpaare PQ, RT, SU nach dem Desarguesschen Involutionsatz über das vollständige Vierseit einer Involution auf der Geraden angehören. Das Diagramm ist hiernach durch die unter den Punkten (Fig. 36) angegebenen Abszissen bestimmt.

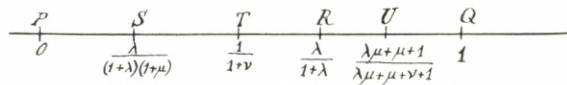


Fig. 36.

51. In dieser Figur können alle Grenzverhältnisse sowohl zwischen infinitesimalen Vektoren als auch zwischen infinitesimalen Winkeln des Vierseits abgelesen werden:

1°. Als Grenzverhältnisse der Seiten des Vierecks AA_1B_1B ergeben sich:

$$\frac{A_1B_1}{AA_1} = \frac{\alpha_1}{\gamma_2} = \frac{PS}{ST} = \frac{\lambda(\nu + 1)}{\lambda(\mu - \nu) + \mu + 1},$$

was zugleich den Index der Schnittkurve k° der Tangentenfläche mit der Schmiegeebene in A angibt. Ferner

$$\frac{A_1B_1}{B_1B} = \frac{\beta_1}{\gamma_1} = \frac{UQ}{TU} = \frac{\nu + 1}{\mu(\lambda + 1)}.$$

2°. Die Verhältnisse

$$\boxed{AA_1 : A_1B_1 : B_1B}$$

können also durch die drei Zahlen

$$\boxed{\frac{\mu + 1}{\lambda} + \mu - \nu, \quad \nu + 1, \quad \mu(\lambda + 1)}$$

ausgedrückt werden, und AB kan zugleich durch die Summe dieser Zahlen, also durch

$$\left(1 + \frac{1}{\lambda}\right)(\mu\lambda + \mu + 1)$$

repräsentiert werden.

Der Index der Schnittkurve k° der Schmiegeebene mit der Tangentenfläche ist wie schon erwähnt

$$\frac{A_1B_1}{AA_1} = \frac{\lambda(\nu + 1)}{\lambda(\mu - \nu) + \mu + 1}.$$

3°. Projiziert man die Raumkurve k derart auf eine Ebene, dass die Schmiegeebene nicht Projektionsebene ist, so hat die entstehende ebene Kurve denselben Index wie die Raumkurve.

4°. Ist die Schmiegeebene Projektionsebene, die Tangente aber nicht Projektionsstrahl, so ist der Index der Projektion der Raumkurve

$$\lambda_1 = \frac{B_1B}{AB_1} = \frac{\alpha_2}{\beta} = \frac{SR}{RQ},$$

also

$$\boxed{\lambda_1 = \frac{\lambda\mu}{\mu + 1}}.$$

Im Falle $\lambda = \mu = \nu = 1$ hat man

$$\lambda_1 = \frac{1}{2}, \quad \lambda_2 = 2, \quad \frac{PM}{MN} = \frac{1}{2}, \quad \frac{VW}{WZ} = 2,$$

alles wohlbekannte Resultate unter den gewöhnlichen elementaren Voraussetzungen über Raumkurven, die z. B. bei algebraischen Raumkurven 3. Ordnung, allgemeiner bei kubikalnen Kurven¹ (bei denen die Krümmungen selbst nicht zu existieren brauchen) erfüllt sind.

52. In einem ebenen Polarsystem entspricht, wie früher erwähnt, einem schwachen Dreieck wieder ein schwaches Dreieck, wobei entsprechende Stücke proportional sind. Hieraus folgt auch für den Raum ein einfaches Prinzip zur Bestimmung der Verhältnisse einer schwachen Figur, die durch Polarität aus einer gegebenen schwachen Figur entstanden ist.

Die einer gegebenen Raumkurve mit den Indexzahlen λ, μ, ν entsprechende Raumkurve hat danach die Indexzahlen

$$\lambda_1 = \frac{1}{\nu}, \quad \nu_1 = \frac{1}{\lambda}, \quad \mu_1 = \frac{\lambda(\mu - \nu) + \mu + 1}{\lambda(\nu + 1)},$$

wo die letzte zum Index der Schnittkurve der Tangentenfläche mit der Schmiegeebene reziprok ist, da diese Kurve dem projizierenden Kegel dual entspricht.

Selbstverständlich hat man auch

$$\mu = \frac{\lambda_1(\mu_1 - \nu_1) + \mu_1 + 1}{\lambda_1(\nu_1 + 1)}$$

in Übereinstimmung mit den bereits gefundenen Ausdrücken.

Sollen die Indizes einer Kurve beim Übergang zur dualen ungeändert bleiben, so muss

$$\lambda\nu = 1$$

sein. An μ werden keine Forderungen gestellt. Dieser Index wird sich aber in einer Hauptklasse von Fällen als gleich ν erweisen.

¹ Das sind Kurven mit der Eigenschaft, dass 6 beliebige ihrer Punkte stets eine nicht ausgeartete algebraische Raumkurve 3. Ordnung bestimmen und alle Grenzlagen solcher algebraischen Kurven gleichfalls nicht ausgeartet sind.

53. Die bisherigen Untersuchungen beruhen ausschliesslich auf der Existenz der Indexzahlen λ , μ , ν , und von Relationen zwischen diesen hat man nur eine, $\mu \geq \nu$, die daraus folgt, dass die Kurve m von der Kurve l umschlossen wird. Die gegenseitige Lage von k^0 und k ergibt nichts Neues.

Wir wollen nun die etwas weitergehende Voraussetzung machen, dass die Raumkurve eine im Punkte A hyperprojektive Verbindung zwischen der Punktchar der

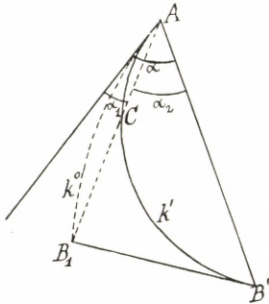


Fig. 38.

Kurve und dem von A ausgehenden projizierenden Kegel vermittelt, was damit gleichbedeutend ist, dass k durch eine Projektion, bei der die Schmiegeebene nicht Projektionsebene ist, in eine ebene Kurve übergeführt wird, die in dem A entsprechenden Punkt einen bestimmten Krümmungsexponenten r hat. Diese Zahl r soll nun auch als Krümmungsexponent der Raumkurve in A be-

zeichnet werden, und es ist klar, dass $r = \frac{1}{\lambda}$ ist.

54. Hiernach kann man zeigen, dass ein bestimmtes Krümmungsverhältnis zwischen k^0 (der Schnittkurve der Tangentenebene mit der Schmiegeebene in A) und k besteht, worunter das Krümmungsverhältnis ξ zwischen k^0 und der Projektion k' der Raumkurve auf die Schmiegeebene zu verstehen ist. (Vgl. Fig. 35, S. 43, und die in der nebenstehenden Fig. 38 dargestellte Projektion, wo B' die Projektion von B und C den Schnittpunkt von k' und AB_1 bezeichnen.) Man hat

$$\xi = \frac{AC}{AB_1} = \frac{AC}{AB'} \cdot \frac{AB'}{AB_1}$$

oder

$$\xi = \left(\frac{\alpha_1}{\alpha}\right)^\lambda \frac{\alpha_2 + \beta}{\beta}.$$

Aus Fig. 35, S. 43, entnimmt man nun

$$\xi = \left(\frac{PS}{PR}\right)^\lambda \frac{SQ}{RQ}$$

oder

$$\xi = \frac{\lambda\mu + \mu + 1}{(\mu + 1)^{\lambda + 1}}.$$

Da das Krümmungsverhältnis somit eine bestimmte, von 0 verschiedene Zahl ist, haben k' und k^0 , also auch k und k^0 , denselben Exponenten und daher auch denselben Index, d. h.

$$\frac{\lambda(\nu+1)}{\lambda(\mu-\nu)+\mu+1} = \lambda$$

oder

$$\mu = \nu,$$

folglich:

Wenn die Raumkurve k einen bestimmten Krümmungsexponenten r und damit einen bestimmten Index $\lambda = \frac{1}{r}$ hat, und wenn ausserdem die beiden anderen Indizes μ, ν existieren, so sind diese einander gleich, und k^0 und k haben ein bestimmtes Krümmungsverhältnis

$$\kappa = \frac{\lambda\mu + \mu + 1}{(\mu + 1)^{\lambda+1}}.$$

Der erste Teil des Satzes kann offenbar auch so formuliert werden:

Wenn eine von einem Punkt A ausgehende, monotone Raumkurve einen bestimmten Krümmungsexponenten r , also auch einen bestimmten Index $\lambda = \frac{1}{r}$, und ausserdem einen bestimmten Index μ in diesem Punkt besitzt, so existiert auch der dritte Index ν und ist gleich μ .

55. Mit Hilfe einer der obigen ähnlichen Rechnung kann man das Krümmungsverhältnis η der Kurven m und l in der Ebene σ finden, wo m die Projektion der Kurve k von A aus und l den Schnitt der Tangentenfläche mit σ bezeichnen: diese beiden Kurven haben unter den jetzigen Voraussetzungen denselben Exponenten μ . Es ergibt sich

$$\eta = \frac{1 + \lambda}{\lambda} \left(\frac{\lambda\mu + \mu}{\lambda\mu + \mu + 1} \right)^\mu.$$

56. Wir wollen nun ausserdem voraussetzen, dass die Beziehung zwischen der Punktschar der Kurve k und dem von der Tangente α ausgehenden projizierenden Ebenenbüschel hyper-

projektiv im Punkte A und in dessen zugeordneter Ebene, der Schmiegeebene α , ist. Der Exponent dieser Beziehung wird der Torsionsexponent der Kurve in A genannt. Sind B und C zwei Nachbarpunkte von A , und werden die Ebenen aB und aC mit β bzw. γ bezeichnet, so ist das Grenzverhältnis $\frac{AB}{AC}$ der beiden infinitesimalen Vektoren AB und AC mit dem Grenzverhältnis $\frac{\alpha\beta}{\alpha\gamma}$ der beiden infinitesimalen Winkel $\alpha\beta$ und $\alpha\gamma$ durch die Gleichung

$$\frac{\alpha\beta}{\alpha\gamma} = \left(\frac{AB}{AC}\right)^s$$

verbunden. Werden die Geraden AB und AC mit b bzw. c bezeichnet, so hat man ferner

$$\frac{ab}{ac} = \left(\frac{AB}{AC}\right)^r.$$

Aus den beiden Gleichungen folgt dann

$$\frac{\alpha\beta}{\alpha\gamma} = \left(\frac{ab}{ac}\right)^{\frac{s}{r}},$$

also, dass der Krümmungsexponent des projizierenden Kegels $\frac{s}{r}$ ist, also

$$\mu = \nu = \frac{r}{s}.$$

Man erkennt, dass die Existenz der Exponenten r und s die Existenz von λ , μ , ν und die Gleichung $\mu = \nu$ zur Folge hat.

57. Eine naheliegende Betrachtung lehrt, dass die Projektion der Raumkurve, wenn die Schmiegeebene Projektionsebene, die Tangente aber nicht Projektionsstrahl ist, den Exponenten $r + s$ hat im Einklang mit der früher gefundenen Formel

$$\lambda_1 = \frac{\lambda\mu}{\mu + 1} = \frac{1}{r + s}.$$

Ferner sieht man, dass eine Projektion in Richtung der Tangente von einem von A verschiedenen Punkt aus den Exponenten

$\frac{s}{r+1}$ hat im Einklang mit der früher gefundenen Formel

$$\lambda_2 = \mu(\lambda + 1) = \frac{r+1}{s}.$$

Wir führen noch die Krümmungsverhältnisse ξ und η an:

$$\xi = \left(1 + \frac{1}{r+s}\right) \left(\frac{s}{r+s}\right)^{\frac{1}{r}}, \quad \eta = (1+r) \left(\frac{r+s}{rs+r+s}\right)^{\frac{r}{s}}.$$

Schliesslich ist der Exponent der Schnittkurve der Tangentenfläche mit einer beliebigen, von der Schmiegeebene verschiedenen Ebene durch a

$$\frac{\mu+1}{\mu\lambda} = r+s$$

und der Exponent der Schnittkurve der Tangentenfläche mit einer beliebigen Ebene durch A , die a nicht enthält,

$$\frac{1}{\mu(\lambda+1)} = \frac{s}{r+1}.$$

Zwei Raumkurven, die einander in einem Polarsystem (einer Korrelation) entsprechen, haben Exponenten r, s und r', s' , die durch die Gleichungen

$$r' = \frac{r}{s}, \quad s' = \frac{1}{s}$$

verbunden sind.

VI. Schwache Figuren im Raume.

58. Eine vom Punkt A ausgehende, monotone Raumkurve mit der Halbtangente x , dem Krümmungsexponenten r und dem Torsionsexponenten s in diesem Punkt, sei auf ein Koordinatensystem x, y, z bezogen, dessen y -Achse in der Schmiegehalbebene und dessen z -Achse in dem Halbraum liegt, der die Kurve enthält. Ein beliebiger, längs der Kurve gegen A konvergierender Punkt wird dann durch drei infinitesimale Vektoren x, y, z dargestellt, deren Richtungen gegen die drei Halbachsen des Koordinatensystems konvergieren.

Eine schwache Figur $A_1A_2\cdots A_i\cdots$, die an die Kurve gebunden ist, ist bestimmt, wenn die zu den Punkten gehörigen Abszissen bestimmte Grenzverhältnisse haben. Wir geben diese mittels der Verhältniszahlen $x_1, x_2, \dots, x_i, \dots$ an. Die den anderen Koordinaten entsprechenden Grenzverhältnisse werden analog durch $y_1, y_2, \dots, y_i, \dots$ und $z_1, z_2, \dots, z_i, \dots$ angegeben. Man hat dann

$$\frac{y_i}{y_1} = \left(\frac{x_i}{x_1}\right)^{r+1}, \quad \frac{z_i}{z_1} = \left(\frac{x_i}{x_1}\right)^{r+s+1}$$

als »Grenzgleichungen« der Kurve. Mit diesen Gleichungen beherrscht man die schwache Figur vollständig in jedem Koordinatensystem der angegebenen Art, so dass jede sie betreffende Frage auf ganz elementare Weise, als ob es sich um eine Raumparabel mit den Gleichungen

$$y = px^{r+1}, \quad z = qx^{r+s+1}$$

in affinen Koordinaten handelte, gelöst werden kann.

59. Als Beispiel kann angeführt werden, dass die Schmiegeebene eines Punktes mit der Abszisse x ein Stück X von der x -Achse abschneidet, das durch das Verhältnis

$$\frac{X}{x} = \frac{r(r+s)}{(r+1)(r+s+1)}$$

bestimmt ist.

Dies entnimmt man jedoch in mehr direkter Weise aus der Fig. 35, S. 43:

$$\frac{X}{x} = \frac{AA_1}{AB} = \frac{\frac{\mu+1}{\lambda} + \mu - \nu}{\left(1 + \frac{1}{\lambda}\right)\left(\lambda\mu + \mu + 1\right)},$$

womit man zugleich ein allgemeineres Resultat hat. Setzt man hierin speziell $\lambda = \frac{1}{r}$, $\mu = \nu = \frac{r}{s}$, so erhält man den obigen Ausdruck.

60. Ein weiteres Beispiel, das das obige als Grenzfall umfasst, ist das folgende:

Drei gegen A konvergierende Punkte der Kurve, für die die Grenzverhältnisse der Abszissen (d. h. der entsprechenden infini-

tesimalen Vektoren) durch die Verhältniszahlen x_1, x_2, x_3 , gegeben sind, bestimmen eine Ebene, die von der Tangente (der x -Achse) ein Stück abschneidet, dessen zugehörige Verhältniszahl

$$x = \left| \begin{array}{ccc} x_1^{r+s+1} & x_1^r & x_1 \\ x_2^{r+s+1} & x_2^r & x_2 \\ x_3^{r+s+1} & x_3^r & x_3 \end{array} \right| : \left| \begin{array}{ccc} x_1^{r+s+1} & x_1^r & 1 \\ x_2^{r+s+1} & x_2^r & 1 \\ x_3^{r+s+1} & x_3^r & 1 \end{array} \right|$$

ist.

61. Eine beliebige schwache Figur $A_1 \cdot A_2 \cdot A_3 \cdot \dots$, die in eine monotone, vom Grenzpunkt A der Figur ausgehende Kurve eingeschrieben ist, kann unter den benutzten Voraussetzungen unmittelbar mit Hilfe eines in der obigen Weise angebrachten Koordinatensystems behandelt werden. Die Relationen zwischen den Verhältniszahlen der Winkel und denen der Vektoren können direkt aus der ebenen Figur, S. 15, entnommen werden, da die Projektion auf die xy -Ebene eine ebene schwache Figur ist, deren Verhältniszahlen mit den entsprechenden der räumlichen Figur übereinstimmen.

Die Verhältnisse der Flächenwinkel des Tetraeders $A_1 A_2 A_3 A_4$ werden durch die Grössen

$$(A_1 A_2), (A_2 A_3), (A_3 A_4), (A_1 A_3), (A_1 A_4), (A_2 A_4)$$

repräsentiert, wo

$$(A_1 A_2) = \frac{\alpha_2 - \alpha_1}{[123][124]}, \quad (A_2 A_3) = \frac{\alpha_3 - \alpha_2}{[234][231]}, \quad (A_3 A_4) = \frac{\alpha_4 - \alpha_3}{[341][342]}$$

und analog für die übrigen (mit der S. 16 angegebenen Bedeutung der Symbole in den Nennern). So erhält man z. B. für das Grenzverhältnis $(A_1 A_2) : (A_2 A_3)$

$$\frac{(A_1 A_2)}{(A_2 A_3)} = \frac{\alpha_2 - \alpha_1}{\alpha_3 - \alpha_2} \cdot \frac{[234]}{[124]}.$$

VII. Berührungsinvarianten von Raumkurven.

62. Zwei monotone Raumkurven k und k_1 , die vom selben Punkt A ausgehen, dort dieselbe Halbtangente a und dieselbe

Schmiegehalbebene α haben, auf derselben Seite der letzteren liegen und schliesslich in A die gleichen Exponenten r und s haben, besitzen zwei wesentliche Berührungsinvarianten, nämlich:

1°. Das Krümmungsverhältnis I_1 , d. h. das Krümmungsverhältnis der Projektionen der beiden Kurven auf die gemeinsame Schmiegeebene von einem beliebigen Punkt aus¹.

2°. Das Torsionsverhältnis I_2 , d. h. das Grenzverhältnis der beiden Vektoren (Sehnen), die in einer von a ausgehenden, gegen α konvergierenden Halbebene von den Kurven ausgeschnitten werden.

63. Bezüglich der Invariante I_1 wollen wir sofort zeigen: Wenn sie für ein Projektionszentrum P existiert, so existiert sie auch für jedes andere Projektionszentrum Q und hat den gleichen Wert. Dass sie für das Projektionszentrum P existiert und den Wert I_1 hat, bedeutet nämlich, dass eine gegen a konvergierende Ebene durch AP die Kurven in zwei Punkten K und K_1 schneidet, die derart gegen A konvergieren, dass das Grenzverhältnis

$$\frac{AK_1}{AK} = I_1$$

ist. Gleichzeitig ist das Grenzverhältnis der Winkel, die die Vektoren AK_1 und AK mit a bilden, gleich 1.

Legt man nun eine Ebene durch AQ und K , die k_1 in einem Punkt K'_1 schneidet, so haben die Winkel, die die Vektoren AK'_1 und AK während des Grenzüberganges mit a bilden, ebenfalls das Grenzverhältnis 1. Hieraus folgt aber, dass auch die Winkel, die AK_1 und AK'_1 mit a bilden, das Grenzverhältnis 1 haben, also dass dasselbe für die Vektoren AK_1 und AK'_1 gilt, d. h. dass das Grenzverhältnis

$$\frac{AK'_1}{AK} = I_1$$

ist, was zu beweisen war.

¹ Selbstverständlich darf dieser Punkt nicht in der Ebene liegen, auf die projiziert wird, und dies wollen wir daher hier und in ähnlichen Fällen stillschweigend annehmen.

Man kann nun auch sagen, dass die Kegel, die die Kurven von einem beliebigen Projektionszentrum P ausserhalb der Schmiegeebene aus projizieren, in der Erzeugenden PA das Krümmungsverhältnis I_1 haben.

Aus der obigen Untersuchung geht zugleich hervor, dass jede gegen a , aber nicht gegen α konvergierende Ebene durch A die Kurven in solchen Punkten K und K_1 schneidet, dass das Grenzverhältnis der Vektoren AK_1 und AK gleich I_1 ist.

64. Eine dritte Berührungsinvariante I_3 ist das Krümmungsverhältnis der die Kurven von A aus projizierenden Kegel. Legt man von a aus eine Halbebene, die die Kurven in M und M_1 schneidet, und lässt man diese Halbebene gegen α konvergieren, so ist das Grenzverhältnis $\frac{AM_1}{AM}$ gleich I_2 , und die Winkel θ und θ_1 , die die Vektoren AM und AM_1 mit der Halbtangente a bilden, haben das Grenzverhältnis

$$\frac{\theta_1}{\theta} = I_3.$$

Eine variable Ebene durch A und M , die gegen a , aber nicht gegen α konvergiert, schneidet die Kurve k_1 in einem neuen Punkt M_2 . Man hat dann

$$\frac{AM_2}{AM} = I_1, \quad \frac{AM_1}{AM} = I_2,$$

$$\left(\frac{AM_1}{AM_2}\right)^r = \frac{\theta_1}{\theta} = I_3,$$

also

$$I_3 = \left(\frac{I_2}{I_1}\right)^r.$$

Die Existenz von I_3 ist eine Folge der Existenz von I_1 und I_2 .

$I_3 = 1$ ist gleichbedeutend mit $I_1 = I_2$. Dieser Fall liegt insbesondere vor, wenn die Kurven in A denselben projizierenden Kegel haben. $I_1 < I_2$ ergibt $I_3 > 1$; im ersten Fall umschliesst der projizierende Kegel der zweiten Kurve den der ersten, im zweiten Fall umgekehrt.

65. Um einen Überblick über die verschiedenen invarianten Relationen zwischen den beiden Kurven zu erhalten, ist es nützlich, zwei beliebige infinitesimale, auf den beiden Kurven endende Vektoren AM und AM_1 zu betrachten. Die Winkel, die die Vektoren mit a bilden, seien θ und θ_1 .

Durch AM lege man eine variable Ebene, die beim Grenzübergang $M \rightarrow A$ gegen a , aber nicht gegen α konvergiert. Diese Ebene schneidet k_1 in einem solchen Punkt M_2 , dass der von AM_2 und a gebildete Winkel θ_2 das Grenzverhältnis 1 zu θ hat. Es gilt nun

$$\frac{AM_2}{AM} = I_1,$$

$$\left(\frac{AM_1}{AM_2}\right)^r = \frac{\theta_1}{\theta},$$

woraus

$$\frac{\theta}{\theta_1} = \left(\frac{AM_2}{AM} \cdot \frac{AM}{AM_1}\right)^r$$

oder

$$\frac{\theta}{\theta_1} = \left(I_1 \cdot \frac{AM}{AM_1}\right)^r \quad (I)$$

folgt.

Bezeichnet man nun die Winkel, die die Halbebenen aM und aM_1 mit a bilden, mit φ und φ_1 und führt (dem obigen analog) den Punkt M_2 ein, in dem die erstgenannte Halbebene k_1 schneidet, so erhält man in ähnlicher Weise

$$\frac{\varphi}{\varphi_1} = \left(I_2 \cdot \frac{AM}{AM_1}\right)^s. \quad (II)$$

66. Projiziert man die beiden Kurven k und k_1 von einem Punkt O aus, der in der gemeinsamen Schmiegeebene in A , aber nicht auf der gemeinsamen Tangente liegt, so erhält man zwei projizierende Kegel, deren Krümmungsverhältnis I_4 längs OA mit Hilfe der Formel in 33. für zwei ebene Kurven bestimmt werden kann. Man erhält so

$$\left(\frac{I_2}{I_1}\right)^r \frac{s}{r+s} = \left(\frac{1}{I_1} I_4\right)^r,$$

also

$$I_4 = I_1^{r+s} I_2^{\frac{s}{r+s}}.$$

67. Wählt man schliesslich ein Projektionszentrum O , das auf der gemeinsamen Tangente der Kurven in A liegt (aber von A verschieden angenommen wird, da der andere Fall bereits behandelt ist), so erhält man zwei projizierende Kegel, deren Krümmungsverhältnis (mit den in 64. angegebenen Bezeichnungen)

$$I_5 = \frac{AM_1}{AM} \cdot \frac{\theta_1}{\theta} = I_2 I_3 = \frac{I_2^{r+1}}{I_1^r}$$

ist.

68. Wir gehen nun zur Besprechung der zu den obigen dualen Invarianten, d. h. der Berührungsinvarianten der Schnittkurven der Tangentenflächen mit ein und derselben Ebene über.

1°. Geht die Schnittebene nicht durch A , so haben die Schnittkurven l und l_1 mit den beiden Tangentenflächen dasselbe Krümmungsverhältnis wie die Schnittkurven m und m_1 der Ebenen mit den von A aus projizierenden Kegeln. Dies folgt unmittelbar daraus, dass das Krümmungsverhältnis von l und m mit dem von l_1 und m_1 übereinstimmt (vgl. den S. 51 angegebenen Ausdruck für dieses Krümmungsverhältnis η). Hieraus folgt:

Das Krümmungsverhältnis der Schnittkurven der Tangentenflächen mit einer beliebigen nicht durch A gehenden Ebene ist gleich dem Krümmungsverhältnis der die Raumkurven von A aus projizierenden Kegel, also gleich

$$I_3 = \left(\frac{I_2}{I_1} \right)^r.$$

2°. Eine ähnliche Betrachtung lehrt:

Das Krümmungsverhältnis der Schnittkurven k^0 und k_1^0 der Tangentenflächen mit der gemeinsamen Schmiegeebene α in A stimmt mit dem der Raumkurven selbst überein, ist also gleich I_1 .

3°. Falls die Schnittebene die Tangente a enthält, aber von der Schmiegeebene verschieden ist, ist das Krümmungsverhältnis der Schnittkurven

$$I_1^{r+s} I_2^{r+s}$$

dieselbe Invariante wie für die Projektionen der Kurven in dem Fall, wo die Schmiegebene Projektionsebene ist (66.).

4°. Geht schliesslich die Schnittebene durch A , aber nicht durch die Tangente a , so ist das Krümmungsverhältnis der Schnittkurven

$$\frac{I_2^{r+1}}{I_1^r}$$

dieselbe Invariante wie für die Projektionen der Raumkurven in Richtung der Tangente (67.).

69. Eine direkte Untersuchung der Relationen zwischen den charakteristischen Zahlen r, s, I_1, I_2 und r', s', I'_1, I'_2 zweier Paare k, k_1 und k', k'_1 von Raumkurven, die einander in einer Polarität entsprechen, zeigt, dass ausser den früher abgeleiteten

$$r' = \frac{r}{s}, \quad s' = \frac{1}{s}$$

die Formeln

$$I'_1 = \left(\frac{I_1}{I_2}\right)^s, \quad I'_2 = \frac{1}{I_2^s}$$

gelten. Hiermit hat man ein Mittel, um von jedem der im vorstehenden gefundenen Resultate zum dualen überzugehen.

70. Die angegebenen Berührungsinvarianten sind mit Hilfe invarianter Methoden, die sofort die betreffenden Relationen in ihrer gegenseitigen Abhängigkeit klar hervortreten lassen, gefunden worden. Sie lassen sich aber natürlich auch mit Hilfe der in 58. besprochenen formalen analytischen Darstellung gewinnen. Um dies zu illustrieren, führen wir die wichtigsten Relationen zwischen den Konstanten der Grenzgleichungen und den betreffenden Invarianten an.

Die beiden Raumkurven seien durch die beiden Paare

$$\left. \begin{aligned} y &= px^{r+1} \\ z &= qx^{r+s+1} \end{aligned} \right\} \text{(I)} \quad \left. \begin{aligned} y &= p_1x^{r+1} \\ z &= q_1x^{r+s+1} \end{aligned} \right\} \text{(II)}$$

von Grenzgleichungen dargestellt. Ihre Projektionen auf die xy -Ebene sind

$$y = px^{r+1}, \quad y = p_1x^{r+1}$$

und die Definition der Invariante I_1 ergibt dann unmittelbar

$$I_1 = \left(\frac{p}{\rho_1} \right)^{\frac{1}{r}} \tag{1}$$

Die Invariante I_2 findet man, indem man mit der Ebene $y = \alpha z$ schneidet. Sind x und x_1 die Abszissen der Schnittpunkte, so ergibt sich

$$I_2 = \frac{x_1}{x} = \left(\frac{q}{q_1} : \frac{p}{p_1} \right)^{\frac{1}{s}}, \tag{2}$$

also wegen (1)

$$\frac{q}{q_1} = I_1^r I_2^s. \tag{3}$$

Mit Hilfe elementarer analytischen Rechnungen und der Formeln (1), (2), (3) ist es nun leicht, alle übrigen Invarianten abzuleiten.

Beispielsweise ist die Invariante der Projektionen

$$z = qx^{r+s+1}, \quad z = q_1 x^{r+s+1}$$

auf die xz -Ebene

$$\left(\frac{q}{q_1} \right)^{\frac{1}{r+s}} = I_1^{r+s} I_2^{\frac{s}{r+s}},$$

womit man die allgemeine Formel aus 66. hat.

Die Projektionen

$$z = \frac{q}{p^{\frac{s}{r+1}+1}} y^{\frac{s}{r+1}+1}, \quad z = \frac{q_1}{p_1^{\frac{s}{r+1}+1}} y^{\frac{s}{r+1}+1}$$

auf die yz -Ebene ergeben die Invariante

$$\left(\frac{q}{q_1} : \left(\frac{p}{p_1} \right)^{\frac{s}{r+1}+1} \right)^{r+1} = I_2^{r+1} : I_1^r,$$

womit das allgemeine Resultat von 67. bestätigt ist.

Indem man die Tangentenflächen analytisch darstellt und ihre Schnittkurven mit den Ebenen $x = 1$, $x = 0$, $y = 0$, $z = 0$ bestimmt, findet man alle Invarianten der ebenen Schnitte der Tangentenflächen, so wie sie in 68. angegeben sind.

71. An Stelle der Grundinvarianten I_1 und I_2 kann man

$$A_1 = I_1^{-r} = \frac{p_1}{p}$$

$$A_2 = I_1^{-r} I_2^{-s} = \frac{q_1}{q}$$

einführen, die sich im Koordinatensystem unmittelbar interpretieren lassen, nämlich A_1 als das Grenzverhältnis der y -Koordinaten von Punkten mit gleicher Abszisse und A_2 als das entsprechende Grenzverhältnis der z -Koordinaten. Da das Koordinatensystem keiner anderen Bedingung unterworfen ist, als dass x in die Halbtangente, y in die Schmiegeebene und z ausserhalb derselben fällt, kann die geometrische Bedeutung von A_1 und A_2 ohne Schwierigkeit unabhängig vom Koordinatensystem formuliert werden.

A_1 und A_2 sind die Segreschen Invarianten der Raumkurven¹.

Als Normalformen der Grenzgleichungen der Kurven kann man hiernach für k :

$$y = x^{r+1}, \quad z = x^{r+s+1}$$

und für k_1 :

$$x = A_1 x^{r+1}, \quad z = A_2 x^{r+s+1}$$

einführen.

72. Als Anwendungsbeispiel wollen wir folgende Aufgabe lösen: Auf dem Kurvenzweige k sind drei gegen A strebende Punkte gegeben, deren Abszissen durch die Zahlen x_1, x_2, x_3 repräsentiert seien. Die Ebene durch die drei Punkte konvergiert gegen die Schmiegeebene in A und schneidet den anderen Kurvenzweig k_1 in einem Punkt, der mit den anderen gegen A konvergiert. Gefragt wird nach der die Abszisse dieses Punktes repräsentierenden Zahl x . Die Aufstellung der Bedingung dafür, dass die vier Punkte in derselben Ebene liegen, ergibt unmittelbar die Gleichung

$$\begin{vmatrix} x & A_1 x^{r+1} & A_2 x^{r+s+1} & 1 \\ x_1 & x_1^{r+1} & x_1^{r+s+1} & 1 \\ x_2 & x_2^{r+1} & x_2^{r+s+1} & 1 \\ x_3 & x_3^{r+1} & x_3^{r+s+1} & 1 \end{vmatrix} = 0$$

¹ C. SEGRE, Sugli elementi curvilinei, che hanno comuni la tangente e il piano osculatore, Rendiconti Lincei (5) 23, 1924;

B. SEGRE, Sugli elementi curvilinei che hanno comuni le origini ed i relativi spazi osculatori, Rendiconti Lincei (6) 22, 1935.

74. Als Beispiel wollen wir die in 72. behandelte Aufgabe für zwei Kurvenzweige lösen, die in A einen Scheitelpunkt bilden (d. h. die Halbtangenten sind entgegengesetzt, die Schmiegehalbebenen identisch und die Schmiegeebene ist Stützebene der ganzen Kurve k, k_1).

In Bezug auf das mit k verknüpfte Koordinatensystem hat die Ebene durch die drei Punkte auf k mit den Abszissen x_1, x_2, x_3 die Gleichung

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & x_1^{r+1} & x_1^{r+s+1} & 1 \\ x_2 & x_2^{r+1} & x_2^{r+s+1} & 1 \\ x_3 & x_3^{r+1} & x_3^{r+s+1} & 1 \end{vmatrix} = 0.$$

Wir suchen nun einen Schnittpunkt dieser Ebene mit k_1 . Es seien $x, A_1x^{r+1}, A_2x^{r+s+1}$ die absolut genommenen Koordinaten dieses Punktes. Durch eine einzige Spiegelung, nämlich an der yz -Ebene, wird k_1 in eine mit k gleich orientierte Kurve übergeführt. Daher muss x die Gleichung

$$\begin{vmatrix} -x & A_1x^{r+1} & A_2x^{r+s+1} & 1 \\ x_1 & x_1^{r+1} & x_1^{r+s+1} & 1 \\ x_2 & x_2^{r+1} & x_2^{r+s+1} & 1 \\ x_3 & x_3^{r+1} & x_3^{r+s+1} & 1 \end{vmatrix} = 0$$

befriedigen, die genau eine reelle Wurzel hat.

Für die anderen Bogentypen löst man die analoge Aufgabe, indem man diejenigen Vorzeichenänderungen der Koordinaten $x, A_1x^{r+1}, A_2x^{r+s+1}$ vornimmt, die den Spiegelungen an Koordinatenebenen entsprechen, durch welche k_1 in eine mit k gleich orientierte Kurve übergeführt wird. Die acht Vorzeichenkombinationen entsprechen gerade den acht Bogentypen.

zur Bestimmung von x . Sie kann bis zu drei Wurzeln haben im Einklang damit, dass die Ebene bis zu drei Schnittpunkten mit k_1 haben kann.

Der Grenzfall, wo die drei Zahlen x_1, x_2, x_3 zusammenfallen und die betreffende Ebene in die Schmiegeebene im Punkt x_1 übergeht, erledigt sich ebenso leicht, und man erhält zur Bestimmung des Schnittpunktes x die Gleichung

$$A_2 \xi^{r+s+1} - A_1 \frac{(r+s)(r+s+1)}{r(r+1)} \xi^{r+1} + \frac{s(r+s+1)}{r} \xi - \frac{s(r+s)}{r+1} = 0$$

in $\xi = \frac{x}{x_1}$.

73. Bisher haben wir ausschliesslich zwei Kurvenzweige k und k_1 betrachtet, die mit derselben räumlichen Orientierung von A ausgehen. Wir wollen nun etwas weiter gehen. Wir halten daran fest, dass die Kurvenzweige vom selben Punkt A ausgehen und dort gemeinsame Tangente und gemeinsame Schmiegeebene haben, verlangen aber nicht mehr, dass sie dieselbe Halbtangente und dieselbe Schmiegehalbebene haben und auf derselben Seite der letzteren liegen. Wir lassen mit anderen Worten zu, dass die Bögen k und k_1 in A einen beliebigen der acht v. Staudtschen Bogentypen bilden.

Bei allen Typen kann man von den Invarianten I_1 und I_2 , und damit auch von A_1 und A_2 sprechen, was man ganz wie für ebene Kurven einsieht. Das Koordinatensystem x, y, z sei mit Bezug auf k in der bisherigen Weise gewählt. Hat nun k_1 nicht dieselbe Halbtangente wie k , so spiegele man k_1 an der yz -Ebene; hat die dabei entstehende Kurve nicht dieselbe Schmiegehalbebene wie k , so spiegele man sie an der xz -Ebene; liegt die neue Kurve nicht auf derselben Seite der Schmiegeebene wie k , so spiegele man sie schliesslich an der xy -Ebene. Auf diese Weise sieht man, dass höchstens drei Spiegelungen an Koordinatenebenen hinreichend sind, um k_1 in eine Kurve k'_1 mit derselben Orientierung wie k überzuführen. Die Invarianten I_1, I_2 , und damit auch A_1, A_2 , von k und k_1 werden nun als die entsprechenden Invarianten von k und k'_1 definiert.

Wir rechnen die Invarianten nicht mit Vorzeichen, sie sind also stets grösser oder gleich 0. Dasselbe soll auch für die Koordinaten der Punkte auf den beiden Kurvenzweigen gelten.

VIII. Kurven in höheren Räumen.

75. Wie Kurven in höheren Räumen zu behandeln sind, liegt nach dem Vorstehenden auf der Hand: Eine monotone Kurve k in einem n -dimensionalen Raum, die von einem Punkt A ausgeht, dort die Halbtangente α^1 , die Schmieghalbebene α^2 , die Schmieghalbräume $\alpha^3, \alpha^4, \dots, \alpha^{n-1}$ hat, ist durch die positiven Krümmungsexponenten r_0, r_1, \dots, r_{n-2} zu charakterisieren, die zu den hyperprojektiven Beziehungen zwischen der Punktschar (P) auf der Kurve und den sie von $A, \alpha^1, \alpha^2, \dots$ bzw. α^{n-2} aus projizierenden Büscheln gehören. Die Kurve kann dann in einem ihr angepassten Koordinatensystem x_0, x_1, \dots, x_{n-1} durch Grenzgleichungen der Form

$$x_1 = p_1 x_0^{r_0+1}, \quad x_2 = p_2 x_0^{r_0+r_1+1}, \quad \dots, \quad x_{n-1} = p_{n-1} x_0^{r_0+r_1+\dots+r_{n-2}+1}$$

dargestellt werden, und dann können elementare analytische Methoden herangezogen werden.

Für zwei vom Punkt A ausgehende Kurven mit derselben Halbtangente, denselben Schmieghalbräumen und denselben Exponenten r_0, r_1, \dots, r_{n-2} in A kann man dann die Berührungsinvarianten I_1, I_2, \dots, I_{n-1} als Grenzverhältnisse der bei den Zentralprojektionen von $A, \alpha^1, \alpha^2, \dots$ bzw. α^{n-2} aus geschnittenen Vektoren einführen.

Auf diese Weise gelangt man unmittelbar zu den von B. SEGRE (für speziellere Kurventypen) aufgestellten Invarianten¹.

76. Auf Einzelheiten, die schwachen Figuren in höheren Räumen betreffend, wollen wir hier nicht eingehen, sondern nur zum Schluss eine andere Verallgemeinerung erwähnen, die daran anknüpft und durch geläufige Abbildungen in die obige eingeordnet werden kann, nämlich die schwachen Figuren höherer Ordnung. Was wir darunter verstehen, wollen wir uns begnügen, an Hand von zwei der einfachsten Beispiele anzudeuten:

1°. Unter einer schwachen Figur zweiter Ordnung in der projektiven Ebene wollen wir eine variable Figur verstehen, deren Punkte sämtlich gegen einen bestimmten Punkt und deren (durch je fünf Punkte der Figur bestimmte) Kegelschnitte sämt-

¹ L. c. Fussnote 1, S. 60.

lich gegen einen bestimmten eigentlichen Kegelschnitt konvergieren.

2°. Im »zyklischen Raume« wollen wir unter einer schwachen zyklischen Figur eine variable Figur verstehen, deren Punkte sämtlich gegen einen bestimmten Punkt und deren (durch je drei Punkte der Figur bestimmte) Kreise sämtlich gegen einen bestimmten Kreis konvergieren.

Diese und analoge schwache Figuren bieten ein interessantes Arbeitsfeld dar, das es sich gewiss lohnen würde zu erforschen.

