

# ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM  
HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,  
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,  
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS V

FASCICULI 1—2



MAGYAR TUDOMÁNYOS AKADÉMIA  
BUDAPEST, 1954

ACTA MATH. HUNG.



# ACTA MATHEMATICA HUNGARICA

A MAGYAR TUDOMÁNYOS AKADEÉMIA III. OSZTÁLYÁNAK  
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY-U. 21

Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok géppel írva, a következő címre küldendők:

*Acta Mathematica, Budapest 62, Postafiók 440.*

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó”-nál (Budapest, V., Alkotmány-utca 21. Bankszámla 04-878-111-48), a külföld számára pedig a „Kultúra” Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin-út 21. Bankszámla 45-790-057-50-032 sz.) vagy külföldi képviselőinél és bizományosainál.

---

„Acta Mathematica” публикует трактаты из области математических наук на русском, французском, английском и немецком языках.

„Acta Mathematica” выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи (в напечатанном на машинке виде) следует направлять по адресу:

*Acta Mathematica* (Венгрия, Будапешт 62, п/я 440).

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена „Acta Mathematica” — 110 форинтов за том. Заказы в стране принимает *Akadémiai Kiadó* (Alkotmány-utca 21. Текущий счет № 04-878-111-48), а для заграницы, предприятие по внешней торговле книг и газет „Kultúra” (Budapest, VI., Sztálin-út 21. Текущий счет № 45-790-057-50-032), или его заграничные представительства и уполномоченные.



## ELEMENTARY PROOFS OF SOME BASIC FACTS CONCERNING ORDER STATISTICS

By

G. HAJÓS (Budapest), member of the Academy, and A. RÉNYI (Budapest),  
corresponding member of the Academy

Let  $\xi_1, \dots, \xi_n$  denote a sample of size  $n$  from a population with the distribution function  $F(x)$ . By other words,  $\xi_1, \dots, \xi_n$  are mutually independent random variables with the common distribution function  $F(x)$ . Let  $\xi_1^*, \dots, \xi_n^*$  be the same set of variables, rearranged in increasing order of magnitude, i. e.

$$\xi_k^* = R_k(\xi_1, \dots, \xi_n),$$

where  $R_k(x_1, \dots, x_n)$  denotes the  $k$ 'th term of the sequence obtained by rearranging the numbers  $x_1, \dots, x_n$  in increasing order of magnitude.

The present paper deals with the *order statistic*  $\xi_k^*$ . Some basic facts will be proved by simple methods. We aim expressively to avoid the calculus and at reduction of any calculation to possibly minimal extent. As consequence, our results may be easily checked by calculation in various different ways, which we are not intended to mention.

Our results seem us mostly to be known, though we did not find some of them explicitly in the literature. We endeavoured to give an elementary and systematic treatment of our subject. Accordingly, our paper may be of methodical interest. As to the literature we refer to the bibliography compiled by S. S. WILKS [3] and by the second named author [4].

1. In order to obtain distribution-free results, i. e. results independent of the distribution function  $F(x)$ , we introduce  $\eta_k = F(\xi_k)$  ( $k=1, \dots, n$ ). If we suppose that  $y=F(x)$  is strictly increasing and continuous, the same holds for the inverse function  $x=F^{-1}(y)$  and we have<sup>1</sup>

$$\mathbf{P}(\eta_k < x) = \mathbf{P}(\xi_k < F^{-1}(x)) = F(F^{-1}(x)) = x \quad (0 \leq x \leq 1),$$

what shows that the variables  $\eta_1, \dots, \eta_n$  are *uniformly distributed* in the interval  $(0, 1)$ . Putting  $\eta_k^* = F(\xi_k^*)$  ( $k=1, \dots, n$ ) we have

$$\eta_k^* = F(\xi_k^*) = F(R_k(\xi_1, \dots, \xi_n)) = R_k(F(\xi_1), \dots, F(\xi_n)) = R_k(\eta_1, \dots, \eta_n).$$

Consequently  $\eta_1^*, \dots, \eta_n^*$  are order statistics of a sample of size  $n$  from a population of uniform distribution in  $(0, 1)$ .

<sup>1</sup>  $\mathbf{P}(A)$  denotes the probability of the event  $A$ .



Accordingly, we confine ourselves in what succeeds to the research of the order statistics  $\eta_{jk}^*$ . Our results may be interpreted as facts concerning the original order statistics  $\xi_k^*$ .

2. The variables  $\eta_{jk}^*$  are not independent, the relation  $\eta_j^* \leq \eta_k^*$  (in case  $j < k$ ) contradicts the independency.

The joint density function of the variables  $\eta_1^*, \dots, \eta_n^*$  is

$$(1) \quad f(x_1, \dots, x_n) = n! \quad (0 \leq x_1 \leq \dots \leq x_n \leq 1).$$

As a matter of fact, if  $E$  denotes any measurable subset of the  $n$ -dimensional simplex defined by the inequalities  $0 \leq x_1 \leq \dots \leq x_n \leq 1$ , we have

$$\mathbf{P}((\eta_1^*, \dots, \eta_n^*) \in E) = \sum \mathbf{P}((\eta_{i_1}, \dots, \eta_{i_n}) \in E),$$

where the summation is extended over all permutations  $i_1, \dots, i_n$  of the indices  $1, \dots, n$  and the density function of  $(\eta_{i_1}, \dots, \eta_{i_n})$  is equal to 1 at any point  $(x_1, \dots, x_n)$  of the cube  $0 \leq x_k \leq 1$  ( $k = 1, \dots, n$ ).

Considering now the case that  $\eta_{jk}^* = c_k, \dots, \eta_n^* = c_n$  ( $2 \leq k \leq n$ ) are fixed, we state that  $\eta_1^*, \dots, \eta_{k-1}^*$  are order statistics of a sample of size  $k-1$  from a population of uniform distribution in the interval  $(0, c_k)$ . In fact, this is true if  $\eta_1^*, \dots, \eta_{k-1}^*$  are furnished by any given  $k-1$  variables out of  $\eta_1, \dots, \eta_n$ , since these  $k-1$  variables are uniformly distributed, even within the cube  $0 \leq x_i \leq c_k$  ( $i = 1, \dots, k-1$ ). Thus, by (1), the joint density function of the variables  $\eta_1^*, \dots, \eta_{k-1}^*$ , under condition  $\eta_{jk}^* = c_k, \dots, \eta_n^* = c_n$ , is

$$(2) \quad f(x_1, \dots, x_{k-1} | c_k, \dots, c_n) = \frac{(k-1)!}{c_k^{k-1}} \quad (0 \leq x_1 \leq \dots \leq x_{k-1} \leq c_k).$$

By the same argument, if  $\eta_1^* = c_1, \dots, \eta_k^* = c_k$  ( $1 \leq k \leq n-1$ ) are fixed,  $\eta_{k+1}^*, \dots, \eta_n^*$  are order statistics of a sample of size  $n-k$  from a population of uniform distribution in the interval  $(c_k, 1)$  and the joint density function of the variables  $\eta_{k+1}^*, \dots, \eta_n^*$ , under condition  $\eta_1^* = c_1, \dots, \eta_k^* = c_k$ , is

$$(3) \quad f(x_{k+1}, \dots, x_n | c_1, \dots, c_k) = \frac{(n-k)!}{(1-c_k)^{n-k}} \quad (c_k \leq x_{k+1} \leq \dots \leq x_n \leq 1).$$

Since (2) and (3) depend only on  $c_k$ , our statements hold also under the only condition  $\eta_{jk}^* = c_k$ , i. e. (2) and (3) give also the values of the functions  $f(x_1, \dots, x_{k+1} | c_k)$  and  $f(x_{k+1}, \dots, x_n | c_k)$ , and the same holds under any restriction on the non-occurring variables. By the same argument, under condition  $\eta_{jk}^* = c_k$ , the sets of variables  $(\eta_1^*, \dots, \eta_{k-1}^*)$  and  $(\eta_{k+1}^*, \dots, \eta_n^*)$  are independent. By other words, order statistics form a *Markov chain*.<sup>2</sup>

3. The joint density function of the variables  $\eta_{i+1}^*, \dots, \eta_k^*$  ( $1 \leq i < k \leq n$ ) is

$$(4) \quad f_{ik}(x_{i+1}, \dots, x_k) = \frac{n!}{i!(n-k)!} x_{i+1}^i (1-x_k)^{n-1} \quad (0 \leq x_{i+1} \leq \dots \leq x_k \leq 1).$$

<sup>2</sup> A. N. KOLMOGOROFF [1] was the first to remark this.



Indeed, the joint density function of the variables  $\eta_1^*, \dots, \eta_i^*, \eta_{k+1}^*, \dots, \eta_n^*$ , under the condition

$$(C_1) \quad \eta_{i+1}^* = x_{i+1}, \dots, \eta_k^* = x_k,$$

is clearly given by

$$f(x_1, \dots, x_i, x_{k+1}, \dots, x_n | x_{i+1}, \dots, x_k) = \frac{f(x_1, \dots, x_n)}{f_{ik}(x_{i+1}, \dots, x_k)}.$$

On the other hand, since the variables  $\eta_1^*, \dots, \eta_i^*$  are by the Markov chain property, under condition (C<sub>1</sub>), independent of the variables  $\eta_{k+1}^*, \dots, \eta_n^*$ , we have

$$\begin{aligned} & f(x_1, \dots, x_i, x_{k+1}, \dots, x_n | x_{i+1}, \dots, x_k) = \\ & = f(x_1, \dots, x_i | x_{i+1}, \dots, x_n) \cdot f(x_{k+1}, \dots, x_n | x_1, \dots, x_k). \end{aligned}$$

Taking into account (1), (2) and (3), comparison of both statements establishes (4).

By (5), taking the special case  $i+1 = k$  the density function of  $\eta_{ik}^*$  is

$$(5) \quad f_k(x_k) = \frac{n!}{(k-1)!(n-k)!} x_k^{k-1} (1-x_k)^{n-k} = \beta_{nk}(x),$$

i. e. order statistics  $\eta_{ik}^*$  have Beta-distribution.

The joint density function of any  $\eta_{k_1}^*, \dots, \eta_{k_r}^*$  ( $1 \leq k_1 < \dots < k_r \leq n$ ) variables is

$$f_{k_1, \dots, k_r}(x_1, \dots, x_r) = C_{k_1, \dots, k_r} x_1^{k_1-1} (x_2-x_1)^{k_2-k_1-1} \dots (x_r-x_{r-1})^{k_r-k_{r-1}-1} (1-x_r)^{n-k_r},$$

where

$$C_{k_1, \dots, k_r} = \frac{n!}{(k_1-1)!(k_2-k_1-1)! \dots (k_r-k_{r-1}-1)!(n-k_r)!}$$

and

$$0 \leq x_1 \leq \dots \leq x_r \leq 1.$$

The proof is given immediately by the above argument, if we consider instead of (C<sub>1</sub>) the condition

$$(C_2) \quad \eta_{k_1}^* = x_1, \dots, \eta_{k_r}^* = x_r,$$

divide (0, 1) by  $x_1, \dots, x_r$  into  $r+1$  subintervals, and take into account that, by the Markov chain property, the sets of variables  $\eta_i^*$  lying in these subintervals are independent under condition (C<sub>2</sub>).

4. We define  $\eta_0^* = 0, \eta_{n+1}^* = 1$ , and introduce the variables

$$\delta_k = \eta_{k-1}^* - \eta_{k+1}^* \quad (k = 1, \dots, n+1).$$

Since  $\delta_1, \dots, \delta_n$  are obtained from  $\eta_1^*, \dots, \eta_n^*$  by a measure-preserving linear transformation, their joint density functions are equal at corresponding places. By corresponding transformation  $y_1 = x_1, y_k = x_k - x_{k-1}$  ( $k = 2, \dots, n$ ) of (1) the joint density function of the random variables  $\delta_1, \dots, \delta_n$  is

$$(6) \quad g(y_1, \dots, y_n) = n! \quad (y_1 \geq 0, \dots, y_n \geq 0; y_1 + \dots + y_n \leq 1).$$



We conclude from (6) that the variables  $\delta_1, \dots, \delta_n$  have the same distribution. As the distribution of  $\eta_k$  is symmetric with respect to the point  $\frac{1}{2}$ , this symmetry holds also for the joint distribution of  $(\eta_1^*, \dots, \eta_n^*)$ . Especially,  $\delta_1 = \eta_1^*$  and  $\delta_{n+1} = 1 - \eta_n^*$  have equal distributions. We draw the conclusion that the variables  $\delta_1, \dots, \delta_{n+1}$  are equally distributed.

Their common density function is that of  $\eta_1^*$ , given by (5),

$$\beta_{n1}(x) = n(1-x)^{n-1} \quad (0 \leq x \leq 1).$$

Their common mean value is, because of  $\delta_1 + \dots + \delta_{n+1} = 1$ ,

$$(7) \quad \mathbf{M}(\delta_k) = \frac{1}{n+1} \quad (k = 1, \dots, n+1).$$

5. The variables  $\delta_1, \dots, \delta_{n+1}$  are not only equally distributed, but are also *equivalent variables*, i. e. their joint distribution remains invariant under any permutation of them. This is established by (6) for permutations of  $\delta_1, \dots, \delta_n$ , by the above mentioned symmetry for the permutation  $(\delta_1, \dots, \delta_{n+1}) \rightarrow (\delta_{n+1}, \dots, \delta_1)$ , and by successive application of these for any permutation of  $\delta_1, \dots, \delta_{n+1}$ .

In consequence, the distribution of the *difference*

$$\eta_{i+k}^* - \eta_i^* = \delta_{i+1} + \dots + \delta_{i+k} \quad (0 \leq i < i+k \leq n+1)$$

depends only on  $k$ , is therefore equal to that of  $\eta_k^*$ , has the density function (5),

and, by (7), the mean value  $\frac{k}{n+1}$ .

Especially, the *range*  $\mathcal{A} = \eta_n^* - \eta_1^*$  has the density function

$$\beta_{n, n-1}(x) = n(n-1)x^{n-2}(1-x) \quad (0 \leq x \leq 1)$$

and the mean value

$$\mathbf{M}(\mathcal{A}) = \frac{n-1}{n+1}.$$

6. As previously stated, under condition  $\eta_k^* = c_k, \dots, \eta_n^* = c_n$  the variable  $\eta_i^*$  ( $1 \leq i < k \leq n$ ) is the  $i$ 'th order statistic of a sample of size  $k-1$  from a population of uniform distribution in the interval  $(0, c_k)$ . Hence, the distribution of the *quotient*  $\frac{\eta_i^*}{\eta_k^*}$  does not depend even on  $c_k$ , remains therefore unaltered if  $\eta_k^*, \dots, \eta_n^*$  are not fixed, and has, by (6), the density function  $\beta_{i+k-1, k}(x)$ . Thus, both differences and quotients of order statistics  $\eta_k^*$  have Beta-distribution.

The variables  $\frac{\eta_k^*}{\eta_{k+1}^*}$  ( $k = 1, \dots, n$ ) are *mutually independent*. Indeed, by above statement, the distribution of  $\frac{\eta_k^*}{\eta_{k+1}^*}$  does not depend on the values of

$\eta_{k+1}^*, \dots, \eta_n^*$ , is therefore independent of  $\frac{\eta_{k+1}^*}{\eta_{k+2}^*}, \dots, \frac{\eta_{n-1}^*}{\eta_n^*}, \frac{\eta_n^*}{\eta_{n+1}^*} = \eta_n^*$ .



Moreover the independent variables

$$\zeta_k^* = \left( \frac{\eta_k^*}{\eta_{k+1}^*} \right)^k \quad (k \leq 1, \dots, n)$$

are uniformly distributed in the interval  $(0, 1)$ .<sup>3</sup> As a matter of fact, the distribution function of  $\eta_n^*$  is

$$\mathbf{P}(\eta_n^* < x) = \prod_{k=1}^n \mathbf{P}(\eta_k < x) = x^n \quad (0 \leq x \leq 1).$$

Correspondingly, since  $\frac{\eta_k^*}{\eta_{k+1}^*}$  is the  $k$ 'th order statistic of a sample of size  $k$  from a population of uniform distribution in  $(0, 1)$ , we have

$$\mathbf{P}\left(\frac{\eta_k^*}{\eta_{k+1}^*} < x\right) = x^k.$$

Whence

$$\mathbf{P}(\zeta_k^* < x) = \mathbf{P}\left(\frac{\eta_k^*}{\eta_{k+1}^*} < \sqrt[k]{x}\right) = x,$$

establishing our statement.

7. We introduce the random variables

$$(8) \quad \mathcal{G}_k = -\ln \zeta_k = -k \ln \frac{\eta_k^*}{\eta_{k+1}^*} \quad (k = 1, \dots, n).$$

These are, according to our preceding result, mutually independent and equally distributed, with the distribution function

$$\mathbf{P}(\mathcal{G}_k < x) = \mathbf{P}(\zeta_k > e^{-x}) = 1 - e^{-x} \quad (x \geq 0),$$

i. e. have exponential distribution with mean value 1.

From equations (8) we get

$$(9) \quad \ln \eta_k^* = -\sum_{j=k}^n \frac{\mathcal{G}_j}{j} \quad (k = 1, \dots, n).$$

Consequently, the *logarithms* of the order statistics  $\eta_k^*$  form not only a Markov chain, but also an *additive chain*, i. e. they are consecutive partial sums of a sequence of mutually independent random variables.

We expressed by (9) the order statistics  $\eta_k^*$  as simple functions of independent and equally distributed random variables. Starting from this fact, *limit theorems* on order statistics may be obtained, by means of the central limit theorem, in a simple and straightforward way. This has been shown by the second named author [4].

(Received 4 May 1954)

<sup>3</sup> This was proved by S. MALMQUIST [2]; his proof is rather complicated.



### Bibliography

- [1] A. N. KOLMOGOROFF, Sulla determinazione empirica di una legge di distribuzione, *Gior. d. Att.*, 4 (1933), pp. 83—91.
- [2] S. MALMQUIST, On a property of order statistics from a rectangular distribution, *Skand. Aktuerietidsskrift*, 33 (1950), pp. 214—222.
- [3] S. S. WILKS, Order statistics, *Bull. Amer. Math. Soc.*, 54 (1948), pp. 6—50.
- [4] A. RÉNYI, On the theory of order statistics, *Acta Math. Acad. Sci. Hung.*, 4 (1953), pp. 191—231.

### ЭЛЕМЕНТАРНОЕ ДОКАЗАТЕЛЬСТВО НЕКОТОРЫХ ОСНОВНЫХ ФАКТОВ ТЕОРИЙ ВАРИАЦИОННЫХ РЯДОВ

Г. ХАЙОШ и А. РЕНЬИ (Будапешт)

(Резюме)

В работе, которая имеет преимущественно методический характер, дается систематическое и элементарное изложение некоторых основных фактов теорий вариационных рядов. Пусть  $\eta_1, \eta_2, \dots, \eta_n$  — независимые случайные величины, которые равномерно распределены в интервале  $(0, 1)$ , пусть  $\eta_1^* \leq \eta_2^* \leq \dots \leq \eta_n^*$  — те же величины, расположенные в возрастающем порядке. Доказывается, между прочим, очень просто, что величины  $\eta_k^*$  ( $k = 1, 2, \dots, n$ ) образуют цепь Маркова (теорема А. Н. Колмогорова, см. [1]), более того, что величины  $\ln \eta_k^*$  ( $k = 1, 2, \dots, n$ ) образуют аддитивную цепь Маркова, так как случайные величины  $\left(\frac{\eta_k^*}{\eta_{k+1}^*}\right)^k$  ( $k = 1, 2, \dots, n; \eta_{n+1}^* \equiv 1$ ) независимы и равномерно распределены в интервале  $(0, 1)$  (теорема С. Малмквиста, см. [2]). С помощью этих фактов возможно доказательство многих предельных теорем теорий вариационных рядов на центральную предельную теорему теории вероятностей (см. [4]).



# BEDINGUNGEN FÜR DIE METRISIERBARKEIT VON AFFINZUSAMMENHÄNGENDEN LINIENELEMENTMANNIGFALTIGKEITEN

Von

O. VARGA (Debrecen), korrespondierendem Mitglied der Akademie

Eine affinzusammenhängende Mannigfaltigkeit von Linienelementen ist bekanntlich durch die Angabe von zwei geometrischen Objekten, nämlich ihren Übertragungsparametern völlig bestimmt. In vorliegender Arbeit wird nun untersucht, unter welchen Bedingungen die Übertragungsparameter diejenigen einer metrischen Linienelementmannigfaltigkeit sind.

Die Frage kommt im wesentlichen auf Angabe der Bedingungen für die Lösbarkeit eines Differentialgleichungssystems hinaus. Diese Bedingungen haben rein algebraischen Charakter und lassen sich in völlig übersichtlicher Weise mit Hilfe der drei Krümmungsaffinoren der Mannigfaltigkeit und der kovarianten Ableitung derselben darstellen.

## § 1. Die Differentialgleichungen, denen der Maßtensor genügt

Eine affinzusammenhängenden Mannigfaltigkeit  $(x^\alpha, v^\alpha)$  von Linienelementen sei durch die Übertragungsparameter  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\cdot\alpha}$  bestimmt. Für diese Objekte bestehen außer dem bekannten Transformationsgesetz<sup>1</sup> noch die Beziehungen

$$(1, 1) \quad C_{\beta\gamma}^{\cdot\alpha} v^\beta = C_{\gamma\beta}^{\cdot\alpha} v^\beta = 0.$$

Die  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\cdot\alpha}$  sind in den  $v^\alpha$  von nullter bzw.  $(-1)$ -ter Dimension homogen. Setzen wir noch

$$(1, 2) \quad G_\beta^\alpha = \Gamma_{\beta\gamma}^{*\alpha} v^\gamma,$$

so ist das einem differenzierbaren Vektorfeld  $\xi^\alpha$  zugeordnete invariante Differential  $\delta\xi^\alpha$  durch

$$(1, 3) \quad \delta\xi^\alpha = d\xi^\alpha + C_{\beta\gamma}^{\cdot\alpha} (dv^\beta + G_\beta^\gamma dx^\gamma) \xi^\gamma + \Gamma_{\beta\gamma}^{*\alpha} dx^\beta \xi^\gamma$$

bestimmt. Durch

$$(1, 4) \quad \delta\xi^\alpha = 0$$

<sup>1</sup> O. VARGA [1], insbes. S. 317.



wird die Parallelübertragung des Vektors  $\xi^x$  von Linienelement  $(x^x, v^x)$  zum Linienelement  $(x^x + dx^x, v^x + dv^x)$  definiert.

Die Linienelementmannigfaltigkeit heißt nach E. CARTAN<sup>2</sup> metrisch, falls es einen symmetrischen Maßtensor  $g_{\alpha\beta}(x, v)$  von der Art gibt, daß erstens

$$(1, 5) \quad \text{Det. } |g_{\alpha\beta}| \neq 0$$

und zweitens die Vektorübertragung metrisch ist, das heißt die Länge eines Vektors bei Parallelübertragung längs einer einparametrischen Schar von Linienelementen ungeändert bleibt. Diese Bedingung ergibt für  $g_{\alpha\beta}$  die Differentialgleichungen

$$(1, 6) \quad \partial_\alpha g_{\beta\gamma} = 2g_{\rho(\beta} \Gamma_{|\alpha|\gamma)}^{\rho\alpha} + 2g_{\rho(\beta} C_{|\alpha|\gamma)}^{\rho\alpha} G_\alpha^x$$

und

$$(1, 7) \quad \partial_{v^\alpha} g_{\beta\gamma} = 2g_{\rho(\beta} C_{|\alpha|\gamma)}^{\rho\alpha}.$$

Falls umgekehrt die Gleichungen (1, 6) und (1, 7) mindestens eine solche Lösung  $g_{\alpha\beta}$  besitzen, für die außerdem (1, 5) gilt, so folgt aus dem Transformationsgesetz der  $\Gamma_{\beta\gamma}^{\alpha\alpha}$  und  $C_{\beta\gamma}^{\alpha\alpha}$ , daß für ein neues Koordinatensystem  $(x', v')$  die Gleichungen (1, 6) und (1, 7) mit Parametern  $\Gamma_{\beta'\gamma'}^{\alpha'\alpha'}$  und  $C_{\beta'\gamma'}^{\alpha'\alpha'}$  in diesem Koordinatensystem,

$$(1, 8) \quad g_{\beta'\gamma'} = g_{\beta\gamma} A_{\beta'}^\beta A_{\gamma'}^\gamma; \quad A_{\beta'}^\beta = \partial_{\beta'} x^\beta(x')$$

als Lösung besitzen Die Lösung  $g_{\alpha\beta}$  ist wegen (1, 6) und (1, 7) in den Zeigern  $\alpha$  und  $\beta$  symmetrisch. Aus (1, 7) und (1, 1) folgt ferner

$$(1, 9) \quad v^\alpha \partial_{v^\alpha} g_{\beta\gamma} = 0.$$

Die Relation (1, 9) besagt, daß  $g_{\alpha\beta}$  in den  $v^x$  von nullter Dimension homogen ist. Aus diesen Feststellungen folgt daher, daß  $g_{\alpha\beta}$  ein kovarianter Tensor zweiter Stufe ist. Hieraus folgt

**SATZ 1.** *Eine notwendige und hinreichende Bedingung dafür, daß der durch  $\Gamma_{\beta\gamma}^{\alpha\alpha}$  und  $C_{\beta\gamma}^{\alpha\alpha}$  bestimmte Affinzusammenhang derjenige einer metrischen Linienelementmannigfaltigkeit ist, besteht darin, daß (1, 6) und (1, 7) ein Lösung  $g_{\alpha\beta}$  besitzt, für die (1, 5) gilt. Durch  $g_{\alpha\beta}$  ist der Maßtensor der Mannigfaltigkeit bestimmt.*

Wir drücken den Tatbestand von Satz 1 dadurch aus, daß wir die affinzusammenhängende Linienelementmannigfaltigkeit als metrisierbar bezeichnen.

Es sei  $L(x, v)$  eine in den  $v^\alpha$  positiv homogene Funktion von erster Dimension, die mindestens dreimal stetig hinsichtlich sämtlicher Argumente differenzierbar ist. Außerdem gebe diese Funktion zu einem regulären Variationsproblem Anlaß, d. h. die quadratische Form

$$(1, 10) \quad \frac{1}{2} \partial_{v^\alpha} \partial_{v^\beta} L^2(x, v) u^\alpha u^\beta$$

<sup>2</sup> E. CARTAN [1], insbes. S. 4—5.



sei in den Hilfsvariablen  $u^\alpha$  positiv definit. Betrachtet man  $L(x, v)$  als die Grundfunktion eines Finslerschen Raumes, für die also das Bogenelement durch  
(1, 11) 
$$ds = L(x, dx)$$

bestimmt ist, dann läßt sich diesem Finslerschen Raum nach E. CARTAN<sup>3</sup> eine metrische Linienelementmannigfaltigkeit eindeutig auf Grund der folgenden vier Bedingungen zuordnen:

C<sub>1</sub>. Jeder Vektor  $\xi^\alpha$ , der die Richtung seines Linienelementes hat, besitzt die Länge  $L(x, \xi)$ .

C<sub>2</sub>. Sind  $\xi^\alpha$  und  $\eta^\alpha$  zwei Vektoren mit festen Komponenten,  $\delta\xi^\alpha$  und  $\delta\eta^\alpha$ , ihre invarianten Differentiale, gebildet für den Fall, daß das gemeinsame Linienelement  $(x^\alpha, v^\alpha)$  sich bloß um sein Zentrum  $x^\alpha$  nach  $(x^\alpha, v^\alpha + dv^\alpha)$  dreht, dann gelte

$$(1, 12) \quad g_{\alpha\beta} \xi^\alpha \delta\eta^\beta = g_{\alpha\beta} \eta^\alpha \delta\xi^\beta.$$

C<sub>3</sub>. Das invariante Differential eines Vektors mit festen Komponenten, der die Richtung seines Linienelementes besitzt, ist Null, wenn das Linienelement sich infinitesimal um sein Zentrum dreht.

C<sub>4</sub>. Wird das Linienelement  $(x^\alpha, v^\alpha)$  parallel nach  $(x^\alpha + dx^\alpha)$  übertragen, dann sind die zu dieser Übertragung gehörige Parameter  $\Gamma_{\beta\gamma}^{*\alpha}$  in den Zeigern  $\alpha$  und  $\beta$  symmetrisch.

Setzen wir zur Abkürzung

$$(1, 13) \quad L^2(x, v) = 2F(x, v),$$

dann bedeuten diese Forderungen analytisch

$$(1, 14) \quad g_{\alpha\beta} = \partial_{v^\alpha} \partial_{v^\beta} F,$$

$$(1, 15) \quad C_{\alpha\beta\gamma} = g_{\gamma\varrho} C_{\dot{\alpha}\dot{\beta}}^{\varrho} = \frac{1}{2} \partial_{v^\alpha} \partial_{v^\beta} \partial_{v^\gamma} F,$$

$$(1, 16) \quad \Gamma_{\alpha\beta\gamma}^{*\alpha} = g_{\varrho\beta} \Gamma_{\alpha\gamma}^{*\varrho} = \frac{1}{2} (\partial_\alpha g_{\beta\gamma} + \partial_\gamma g_{\alpha\beta} - \partial_\beta g_{\alpha\gamma}) + \\ + C_{\alpha\gamma\alpha} G_\beta^\alpha - C_{\alpha\gamma\beta} G_\alpha^\alpha - C_{\alpha\beta\alpha} G_\gamma^\alpha.$$

Hieraus ergibt sich, daß die Übertragungsparameter  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\dot{\beta}\dot{\gamma}}^{\dot{\alpha}}$  außer (1, 1) noch den Bedingungen

$$(1, 17) \quad \Gamma_{\beta\gamma}^{*\alpha} = \Gamma_{\gamma\beta}^{*\alpha},$$

$$(1, 18) \quad C_{\dot{\beta}\dot{\gamma}}^{\dot{\alpha}} = C_{\dot{\gamma}\dot{\beta}}^{\dot{\alpha}}$$

und

$$(1, 19) \quad g_{\varrho\gamma} C_{\dot{\alpha}\dot{\beta}}^{\varrho} = g_{\varrho\beta} C_{\dot{\alpha}\dot{\gamma}}^{\varrho}$$

genügen müssen. Wegen (1, 19) können die Differentialgleichungen (1, 7) jetzt

<sup>3</sup> CARTAN [1], S. 10. Dort sind allerdings fünf Forderungen angegeben, von denen jedoch die mit bezeichnete, eine Folge der übrigen ist.



auf die Gestalt

$$(1, 7') \quad \partial_{v^\alpha} g_{\beta\gamma} = 2g_{\alpha\beta} C_{\alpha\gamma}^{\cdot\cdot\cdot\alpha}$$

gebracht werden.

Wir beweisen nun folgenden

SATZ 2. Besitzen die Differentialgleichungen (1, 6) und (1, 7), in den die  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\cdot\cdot\cdot\alpha}$  nicht nur die Übertragungsparameter eines Affinzusammenhanges sind, sondern noch den Relationen (1, 17) und (1, 18) genügen, eine Lösung  $g_{\alpha\beta}(x, v)$ , für die (1, 19) gilt und die quadratische Form

$$(1, 20) \quad g_{\alpha\beta}(x, v) u^\alpha u^\beta$$

der Hilfsvariablen  $u^\alpha$  positiv definit ist, dann stellen die  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\cdot\cdot\cdot\alpha}$  den Affinzusammenhang des zur Grundfunktion

$$(1, 21) \quad L(x, v) = \sqrt{g_{\alpha\beta}(x, v) v^\alpha v^\beta}$$

gehörigen Finslerschen Raumes dar.

BEWEIS. Aus den vor Satz 1 angestellten Betrachtungen folgt zunächst, daß  $g_{\alpha\beta}$  ein kovarianter Tensor zweiter Stufe ist. Wir müssen nun zeigen, daß die durch (1, 21) festgelegte Grundfunktion  $L(x, v)$  mit  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\cdot\cdot\cdot\alpha}$ , entsprechend den vier Cartanschen Forderungen  $C_1 - C_4$  zusammenhängt. Aus den Gleichungen (1, 7) und (1, 19) ergibt sich bei Beachtung von (1, 1)

$$(1, 22) \quad \partial_{v^\alpha} g_{\beta\gamma} \cdot v^\gamma = \partial_{v^\alpha} g_{\gamma\beta} \cdot v^\gamma = 0.$$

Quadrieren wir die Gleichung (1, 21) und differenzieren sie hinsichtlich  $v^\alpha$ , so erhält man wegen (1, 22) bei Beachtung von (1, 13)

$$(1, 23) \quad \partial_{v^\alpha} F = g_{\alpha\beta} v^\beta$$

und eine nochmalige Differentiation gibt

$$(1, 24) \quad \partial_{v^\alpha} \partial_{v^\beta} = g_{\alpha\beta}.$$

Wegen der Homogenität nullter Dimension der  $g_{\alpha\beta}$  ist  $F$  von zweiter Dimension homogen, so daß die Bedingung  $C_1$  unmittelbar aus (1, 24) folgt. Die Bedingung  $C_2$  ist zufolge der Definition (1, 3) des invarianten Differentials mit (1, 19) identisch. Die Bedingung  $C_3$  ist wegen (1, 1) erfüllt. Zuzufolge der Definition der Parallelübertragung und (1, 1) besteht die Bedingung  $C_4$  wegen (1, 17). Damit ist der Beweis von Satz 2 beendigt.

Aus den Forderungen  $C_1 - C_4$  bestimmt E. CARTAN die  $G_\beta^\alpha$  aus  $L$ . Es ergibt sich dabei,<sup>4</sup> falls

$$(1, 25) \quad G^\alpha = \frac{1}{2} G_\beta^\alpha v^\beta$$

gesetzt wird, die wichtige Beziehung

$$(1, 26) \quad \partial_{v^\beta} G^\alpha = G_\beta^\alpha.$$

<sup>4</sup> CARTAN [1] S. 16, Formel X.



Die  $G^\alpha$  sind schließlich durch die Funktion  $L$  auf folgende Weise festgelegt. Man bestimmt die Eulerschen Gleichungen des zu  $L$  gehörigen Variationsproblems und löst die so entstehenden Gleichungen nach den zweiten Ableitungen  $\frac{d^2 x'}{ds^2}$  auf. Es kommt dann

$$(1, 27) \quad \frac{d^2 x^\alpha}{ds^2} + 2 G^\alpha \left( x, \frac{dx}{ds} \right) = 0.$$

BEMERKUNG. Daß unter den Bedingungen von Satz 2, die  $\Gamma_{\beta\gamma}^{*\alpha}$  und  $C_{\beta\gamma}^{\alpha}$  tatsächlich die zur Grundfunktion  $\sqrt{g_{\alpha\beta} v^\alpha v^\beta}$  gehörigen Übertragungsparameter eines Finslerschen Raumes im Sinne der Cartanschen Theorie sind, d. h. daß sich diese Größen gemäß (1, 15), (1, 16) und (1, 26) aus der Grundfunktion bestimmen lassen, kann man auch durch eine unmittelbare Rechnung feststellen. Wir führen dieselbe deswegen durch, weil es sich dabei im wesentlichen nur um eine Umformung der Gleichungen (1, 6) und (1, 7) handelt, bei der noch (1, 1), (1, 17), (1, 18) und (1, 19) gilt. Es wird dabei die Existenz einer Lösung  $g_{\alpha\beta}$  gar nicht benötigt.

Vertauschen wir die Indizes  $\alpha, \beta, \gamma$  in (1, 6) zyklisch und subtrahieren wir die letzte Gleichung von der Summe der beiden ersten Gleichungen, so ergibt sich wegen (1, 17), (1, 18), (1, 19) und (1, 7)

$$(1, 28) \quad g_{\alpha\gamma} \Gamma_{\beta\alpha}^{*\alpha} = \frac{1}{2} (\partial_\alpha g_{\beta\gamma} + \partial_\beta g_{\gamma\alpha} - \partial_\gamma g_{\alpha\beta}) - \frac{1}{2} \partial_{v^\alpha} g_{\beta\gamma} \cdot G_\alpha^\alpha - \\ - \frac{1}{2} \partial_{v^\alpha} g_{\gamma\alpha} \cdot G_\alpha^\alpha + \frac{1}{2} \partial_{v^\alpha} g_{\alpha\beta} \cdot G_\alpha^\alpha.$$

Differenzieren wir nun

$$(1, 29) \quad \gamma_{\alpha\gamma\beta} = \frac{1}{2} (\partial_\alpha g_{\gamma\beta} + \partial_\beta g_{\alpha\gamma} - \partial_\gamma g_{\alpha\beta})$$

nach  $v^\sigma$  und überschieben mit  $v^\alpha$  und  $v^\beta$ , so folgt wegen (1, 22)

$$(1, 30) \quad \partial_{v^\sigma} \gamma_{\alpha\gamma\beta} v^\alpha v^\beta = 0.$$

Aus (1, 28) erhalten wir nach Überschiebung mit  $v^\alpha$  und  $v^\beta$  bei Benützung der Bezeichnungen (1, 2) und (1, 25)

$$(1, 31) \quad g_{\alpha\gamma} G_\beta^\alpha = \gamma_{\alpha\gamma\beta} v^\alpha - \partial_{v^\alpha} g_{\beta\gamma} \cdot G^\alpha$$

und

$$(1, 32) \quad 2g_{\alpha\gamma} G^\alpha = \gamma_{\alpha\gamma\beta} v^\alpha v^\beta.$$

Partielle Ableitung von (1, 32) nach  $v^\beta$  gibt wegen (1, 30) und (1, 31) die Gleichung

$$(1, 33) \quad g_{\alpha\gamma} (\partial_{v^\beta} G^\alpha - G_\beta^\alpha) = 0.$$

Haben nun die Gleichungen (1, 6) und (1, 7), in denen jetzt die  $\Gamma_{\beta\gamma}^{*\alpha}$  der Symmetriebedingung (1, 17) genüge leisten, eine Lösung  $g_{\alpha\beta}$ , die außer



(1, 19) noch die in Vergleich zu (1, 20) schwächere Bedingung (1, 5) erfüllen, dann sind die  $I_{\beta\gamma}^{*\alpha}$  wegen (1, 33) in der Tat die Cartanschen Übertragungsparameter der zu (1, 29) gehörigen Grundfunktion, da sie dann die Relationen (1, 28) befriedigen. Für die  $C_{\alpha\beta}^{\gamma}$  folgt dies schon aus den Gleichungen (1, 7') selbst.

## § 2. Algebraische Bedingungen für die Existenz von Lösungen $g_{\alpha\beta}$ .

Wir wollen nun feststellen, unter welchen Bedingungen die Differentialgleichungen (1, 6) und (1, 7) erstens unter den Nebenbedingungen (1, 1) und zweitens unter derselben Nebenbedingung und den weiteren Nebenbedingungen (1, 17), (1, 18) und (1, 19) Lösungen besitzen.

Es handelt sich in beiden Fällen um ein System von partiellen Differentialgleichungen, auf die der Satz von O. VEBLEN und J. M. THOMAS angewandt werden kann.<sup>5</sup> Entsprechend dem Verfahren dieses Satzes müssen die Integrabilitätsbedingungen der Gleichungen (1, 6) und (1, 7) in Bezug auf sämtliche Argumente, d. h. sowohl bezüglich  $x^*$  als  $v^*$  unter Beachtung dieser Gleichungen selbst gebildet werden. Gelangt man so zu Relationen, die noch bei Beachtung der Nebenbedingungen und deren Ableitungen Identitäten sind, dann ist die Existenz von Lösungen  $g_{\alpha\beta}$  gewährleistet. Anderenfalls müssen die erhaltenen Relationen und die Nebenbedingungen weiter differenziert werden, und dabei sind sämtliche schon erhaltenen Relationen zu berücksichtigen. Die so entstehende Gleichungskette ist nun entweder identisch erfüllt, oder wird entsprechend weiter behandelt. Es handelt sich dann schließlich um die Feststellung der Verträglichkeit der so entstehenden Gleichungskette.

Bei der Ausführung der entsprechenden Rechnungen beachten wir, daß wegen der Definition<sup>6</sup> der kovarianten Ableitung und des Umstandes, daß  $g_{\alpha\beta}$  ein Tensor ist, die Relationen (1, 6) auf die einfache Gestalt

$$(2, 1) \quad \nabla_{\times} g_{\alpha\beta} = 0$$

gebracht werden können.

Um gleich koordinateninvariant arbeiten zu können, benützen wir an Stelle der partiellen Ableitung hinsichtlich  $x^*$  die kovariante Ableitung dieser Variablen. Die Ableitung nach  $v^*$  muß nicht durch einen anderen Differentiationsprozeß ersetzt werden, da sie den tensoriellen Charakter von Größen nicht zerstört. Dem entsprechend können die Integrabilitätsbedingungen von (1, 6) und (1, 7) auf die Gestalt

$$(2, 2) \quad \nabla_{[\sigma} \nabla_{\kappa]} g_{\alpha\beta} = 0,$$

$$(2, 3) \quad \partial_{\nu} [\sigma \partial_{\nu^*}] g_{\alpha\beta} = 0$$

<sup>5</sup> J. M. THOMAS—O. VEBLEN [1], S. 288.

<sup>6</sup> O. VARGA [2], S. 10—11.



und

$$(2, 4) \quad \nabla_{\kappa} \partial_{\nu} \sigma g_{\alpha\beta} - \partial_{\nu} \kappa \nabla_{\sigma} g_{\alpha\beta} = 0$$

gebracht werden.

Wir wollen nun die Integrabilitätsbedingungen (2, 2), (2, 3) und (2, 4) auf eine Form bringen, in der die drei durch

$$(2, 5) \quad P_{\sigma\kappa\alpha}^{\varrho} = 2\partial_{[\sigma} \Gamma_{\kappa]\alpha}^{\varrho} - 2G_{[\sigma}^{\mu} \partial_{\nu|\mu|} \Gamma_{\kappa]\alpha}^{\varrho} + 2\Gamma_{[\kappa|\alpha]}^{*\mu} \Gamma_{\sigma] \mu}^{\varrho} + \\ + 2C_{\mu\alpha}^{\cdot\varrho} (\partial_{\sigma} G_{\kappa}^{\mu} + G_{\sigma\tau}^{\mu} G_{\varrho}^{\tau}); \quad G_{\varrho\tau}^{\mu} = \partial_{\nu\sigma} G_{\tau}^{\mu},$$

$$(2, 6) \quad \Sigma_{\sigma\kappa\alpha}^{\varrho} = 2\partial_{\nu[\sigma} C_{\kappa]\alpha}^{\cdot\varrho} + 2C_{[\sigma|\mu]}^{\cdot\varrho} C_{\kappa]\alpha}^{\cdot\mu},$$

$$(2, 7) \quad \Pi_{\sigma\kappa\alpha}^{\varrho} = \partial_{\nu\sigma} \Gamma_{\kappa\alpha}^{*\varrho} + \partial_{\nu\sigma} \Gamma_{\kappa\lambda}^{*\tau} \cdot v^{\lambda} C_{\tau\alpha}^{\cdot\varrho} - \nabla_{\kappa} C_{\sigma\alpha}^{\cdot\varrho}$$

bestimmten Krümmungsaffinoren<sup>7</sup> der affinzusammenhängenden Linieelementmannigfaltigkeit eingehen. Zunächst betrachten wir den ersten Fall, indem für die Differentialgleichungen (1, 6) und (1, 7) die Nebenbedingungen (1, 17) und (1, 19) nicht vorausgesetzt werden.

Für einen beliebigen Affinor zweiter Ordnung  $t_{\alpha\beta}$  gilt die Vertauschungsformel

$$(2, 8) \quad \nabla_{[\sigma} \Delta_{\kappa]} = -2t_{\varrho(\beta} P_{|\sigma\kappa|\alpha)}^{\varrho} + 2t_{\varrho(\beta} C_{|\mu|\alpha)}^{\cdot\varrho} P_{\sigma\kappa\tau}^{\mu} v^{\tau} - \partial_{\nu\mu} t_{\alpha\beta} \cdot P_{\sigma\kappa\tau}^{\mu} v^{\tau}.$$

Für

$$t_{\alpha\beta} = g_{\alpha\beta}$$

nimmt die Integrabilitätsbedingung (2, 2) wegen (2, 1) und (2, 8) die Form

$$(2, 9) \quad g_{(\alpha|\varrho} P_{\sigma\kappa|\beta)}^{\varrho} = 0$$

an.

Für die Integrabilitätsbedingung (2, 3) gibt eine einfache Rechnung wegen (2, 6) die Relation

$$(2, 10) \quad g_{(\alpha|\varrho} \Sigma_{\sigma\kappa|\beta)}^{\varrho} = 0.$$

Bei der Umformung der Integrabilitätsbedingungen (2, 4) beachten wir daß für einen beliebigen Affinor  $t_{\alpha\beta}$  die Vertauschungsformel

$$(2, 11) \quad \nabla_{\kappa} \partial_{\nu\sigma} t_{\alpha\beta} - \partial_{\nu\sigma} \nabla_{\kappa} t_{\alpha\beta} = \partial_{\nu\sigma} \Gamma_{\kappa\lambda}^{*\mu} \cdot v^{\lambda} \partial_{\mu} t_{\alpha\beta} + \partial_{\nu\sigma} \Gamma_{\kappa\alpha}^{*\mu} \cdot t_{\mu\beta} + \partial_{\nu\sigma} \Gamma_{\kappa\beta}^{*\mu} \cdot t_{\alpha\mu}$$

gilt. Für

$$t_{\alpha\beta} = g_{\alpha\beta}$$

folgt aus (2, 11) wegen (2, 1) und (2, 7) für die Integrabilitätsbedingungen (2, 4) der Ausdruck

$$(2, 12) \quad g_{\mu(\alpha} \Pi_{|\sigma\kappa|\beta)}^{\mu} = 0.$$

Im zweiten Falle müssen die  $g_{\alpha\beta}$  außer den Differentialgleichungen (1, 6) und (1, 7) noch die Skalarenrelationen (1, 19) erfüllen.

<sup>7</sup> O. VARGA [2], S. 12. An Stelle der im Text verwendeten Bezeichnungen steht dort  $P_{\alpha}^{\varrho}{}_{\kappa\sigma}$ ,  $\Sigma_{\alpha\kappa\sigma}^{\varrho}$ ,  $\Pi_{\alpha\kappa\sigma}^{\varrho}$ .



Entsprechend den Bemerkungen, die wir im Anschluß an den Satz von J. M. THOMAS und O. VEBLEN machten, müssen die kovarianten bzw. gewöhnlichen Ableitungen dieser skalaren Relationen gebildet werden. Kombiniert man die sich so ergebenden Relationen mit (2, 10) bzw. (2, 12), so kommen zu den Integrabilitätsbedingungen (2, 9), (2, 10) und (2, 12) die weiteren Relationen

$$(2, 13) \quad g_{\alpha\varrho}(\Sigma_{\sigma\kappa\beta}^{\varrho} - \bar{S}_{\sigma\kappa\beta}^{\varrho}) = 0$$

und

$$(2, 14) \quad (2H_{\sigma\kappa\alpha}^{\varrho} - \partial_{v^{\sigma}}\Gamma_{\kappa\lambda}^{*\varrho})g_{\varrho\beta} + \partial_{v^{\sigma}}\Gamma_{\kappa\beta}^{*\varrho}\cdot g_{\varrho\alpha} = 0$$

hinzu. In (2, 13) ist  $\bar{S}_{\sigma\kappa\beta}^{\varrho}$  bis auf den Faktor  $1/L^2$  gerade derjenige Affinor, auf den sich  $\Sigma_{\sigma\kappa\beta}^{\varrho}$  im Falle des Finslerschen Raumes reduziert. Derselbe ist bekanntlich<sup>8</sup> durch

$$(2, 15) \quad \bar{S}_{\sigma\kappa\beta}^{\varrho} = 2C_{[\kappa|\mu]|\beta}^{\varrho}C_{\sigma] \mu}^{\mu}$$

definiert.

Im ersten Falle bestehen zufolge unserer Ergebnisse die Integrabilitätsbedingungen dann identisch, falls die drei Krümmungsaffinoren (2, 5), (2, 6) und (2, 7) der Linielementmannigfaltigkeit identisch verschwinden. Da die Lösung  $g_{\alpha\beta}$  dann in einem beliebigen Linielement  $(x_0^{\alpha}, v_0^{\alpha})$  willkürlich vorgeschrieben werden kann, so wird man  $g_{\alpha\beta}(x, v)$  so wählen können, daß für dieses Linielement die Bedingung (1, 5) erfüllt ist. Daß hingegen

$$\text{Det } |g_{\alpha\beta}(x_0, v)| \neq 0$$

(zitiert als Bedingung A) für sämtliche Linielemente durch  $x_0^{\alpha}$  gilt, folgt natürlich nicht.

Im zweiten Falle besitzen die Differentialgleichungen (2, 6) und (1, 7) stets dann Lösungen  $g_{\alpha\beta}$ , für die noch die Werte in einem Linielement beliebig vorgeschrieben werden können, falls außer den drei Krümmungsaffinoren — wegen (2, 13) und (2, 14) — noch der Affinor (2, 15) und der Affinor  $\partial_{v^{\sigma}}\Gamma_{\kappa\beta}^{*\varrho}$  identisch verschwinden.

Erfüllt eine Lösung noch die Bedingung (1, 20) für sämtliche Linielemente durch einen Punkt (zitiert als Bedingung B), dann ist der Finslersche Raum ein Minkowskischer<sup>9</sup> mit konstanter Winkelmetrik.<sup>10</sup> Diese Ergebnisse ergeben zusammengefaßt:

**SATZ 3.** Falls die Krümmungsaffinoren einer affinzusammenhängenden Linielementmannigfaltigkeit identisch verschwinden, so läßt sich diese Mannigfaltigkeit dann metrisieren, falls es unter den (sicher vorhandenen) Lösungen  $g_{\alpha\beta}$  von (1, 6) und (1, 7) eine solche gibt, für die die Bedingung A gilt.

<sup>8</sup> E. CARTAN [1], S. 34.

<sup>9</sup> E. CARTAN [1], S. 35, Formel XVII und S. 36, sowie 39.

<sup>10</sup> E. CARTAN [1], S. 34—35.



Erfüllen die Übertragungsparameter noch die Symmetriebedingungen (1, 17) und (1, 18), und verschwinden außer den drei Krümmungsaffinoren noch die Affinoren  $\bar{S}_{\sigma \times \beta}^{\circ}$  und  $\partial_{v^{\sigma}} \Gamma_{\times \beta}^{\circ}$ , so läßt sich eine Metrisierung so durchführen, daß der Raum ein Minkowskischer wird, falls die Lösungen  $g_{\alpha \beta}$  von (1, 6) und (1, 7) noch die Bedingung B erfüllen.

Bestehen die Integrabilitätsbedingungen (2, 9), (2, 10) und (2, 12) nicht identisch, so müssen diese Relationen hinsichtlich  $v^{\times}$  gewöhnlich und hinsichtlich  $x^{\times}$  kovariant abgeleitet werden. Beachten wir die Form (2, 1) der Differentialgleichungen (1, 6) und weiter, daß die Ableitungen der  $g_{\alpha \beta}$  nach  $v^{\times}$  wegen (1, 7) linear in diesen Größen sind, so folgt, daß man durch diesen Prozess wieder lineare Relationen in den  $g_{\alpha \beta}$  erhält. Auf Grund der Herleitung der Relationen sind dieselben, wieder koordinateninvariante Affinorrelationen.

Aus (2, 9) erhält man so die Beziehungen

$$(2, 16) \quad g_{(\alpha|\rho} \Delta_{\lambda} P_{\sigma \times|\beta)}^{\circ} = 0$$

und

$$(2, 17) \quad g_{\mu(\alpha} C_{|\lambda|\rho)}^{\mu} P_{\sigma \times \beta}^{\circ} + g_{\mu(\beta} C_{|\lambda|\rho)}^{\mu} P_{\sigma \times \alpha}^{\circ} + g_{(\alpha|\rho} \partial_{v^{\lambda}} P_{\sigma \times|\beta)}^{\circ} = 0.$$

Aus (2, 10) und (2, 11) folgen je zwei Relationen, die genau (2, 16) und (2, 17) entsprechen mit der Änderung, daß jetzt an Stelle des Affinors  $P_{\sigma \times \beta}^{\circ}$  die Affinoren  $\Sigma_{\sigma \times \beta}^{\circ}$  bzw.  $\Pi_{\sigma \times \alpha}^{\circ}$  treten.

Da die drei Affinoren homogene Funktionen in den  $v^{\times}$  sind, folgt aus (2, 17) und den, entsprechend mit den  $\Sigma_{\sigma \times \beta}^{\circ}$  und  $\Pi_{\sigma \times \alpha}^{\circ}$  gebildeten Relationen bei Beachtung von (1, 1) nach Überschiebung mit  $v^{\times}$  wieder (2, 9), bzw. (2, 10) und (2, 12).

Dieses Differentiationsverfahren kann beliebig fortgesetzt werden. Dabei ist zu beachten, daß die durch Differentiation nach  $v^{\times}$  erhaltenen, neuen Relationen stets die vorangehenden Relationen als Folge enthalten, so daß diese dann überflüssig werden. Aus dem zitierten Satz von J. M. THOMAS und O. VELEN folgt nun

**SATZ 4.** Die affinzusammenhängende Linielementmannigfaltigkeit läßt sich metrisieren, falls es erstens ein Zahl  $N$  von der Art gibt, daß die ersten  $N$  Ketten von Relationen (2, 9), (2, 10), (2, 12), (2, 16), (2, 17) usw. ein verträgliches Gleichungssystem für die  $g_{\alpha \beta}$  geben und daß eine Lösung auch die  $(N+1)$ -te Kette von Relationen erfüllt, und zweitens die  $g_{\alpha \beta}$  die Bedingung A erfüllen.

Wir bemerken, daß für den Fall symmetrischer Übertragungsparameter ein zu Satz 4 entsprechender Satz gilt. In diesem Falle wird die Mannigfaltigkeit ein Finslerscher Raum. Jetzt müssen noch die zusätzlichen Gleichungen (2, 13) und (2, 14) ebenso differenziert werden, wie oben (2, 9), (2, 10) und (2, 12). Weiter muß statt Bedingung A jetzt die Bedingung B gelten.

(Eingegangen am 4. Januar 1954.)



### Literaturverzeichnis

- E. CARTAN [1], *Les espaces de Finsler*, Actualités sci. et industr. **79**, (Paris, 1934).  
J. M. THOMAS—O. VELEN [1], Projective invariants of affine geometry of paths, *Annals of Math.*, **27** (1926), S. 249—296.  
O. VARGA [1], Affinzusammenhängende Mannigfaltigkeiten von Linienelementen, die ein Inhaltsmaß besitzen, *Kon. Akad. v. Wetensch. Amsterdam. Proc.*, **52** (1949), S. 316—322  
O. VARGA [2], Über affinzusammenhängende Mannigfaltigkeiten von Linienelementen, insbesondere deren Äquivalenz, *Publ. Math., Debrecen*, **1** (1949), S. 7—17.

### ОБ УСЛОВИЯХ МЕТРИЗУЕМОСТИ АФФИННО-СВЯЗНЫХ МНОГО- ОБРАЗИЙ ЛИНЕЙНЫХ ЭЛЕМЕНТОВ

О. ВАРГА (Дебрецен)

(Резюме)

Известно, что аффинно-связное многообразие линейных элементов полностью определено заданием двух геометрических объектов, его двух параметров переноса. В настоящей работе исследуются условия, при которых параметры переноса соответствуют метрическому многообразию линейных элементов. Полученные условия носят чисто алгебраический характер, и они выражаются с помощью трёх аффиноров кручения многообразия и его ковариантной производной во весьма прозрачном виде.



# ÜBER DIE KANTENBASEN FÜR ENDLICHE VOLLSTÄNDIGE GERICHTETE GRAPHEN

Von

LADISLAUS RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

Wenn nichts anderes gesagt wird, so halten wir uns bezüglich der graphentheoretischen Grundbegriffe, Fachausdrücke und Bezeichnungen an die Monographie von KÖNIG [2].<sup>1</sup>

Unter der *Ordnung* eines endlichen Graphen verstehen wir die Anzahl seiner Knotenpunkte, die wir im folgenden kurz die Punkte des Graphen nennen. Mit  $\mathfrak{N}$  bezeichnen wir das *Netz*  $n$ -ter Ordnung, d. h. den *vollständigen gerichteten Graphen*  $n$ -ter Ordnung. (Das bedeutet, daß in  $\mathfrak{N}$  jede zwei Punkte mit genau zwei entgegengesetzt gerichteten Kanten verbunden sind, weshalb  $\mathfrak{N}$  genau  $n(n-1)$  Kanten hat.) Unter einer *Kantenbasis* für (oder von)  $\mathfrak{N}$  verstehen wir nach KÖNIG [2] (S. 92) einen Teilgraphen  $\mathfrak{B}$  von  $\mathfrak{N}$  mit den folgenden zwei Eigenschaften:

**a)** In  $\mathfrak{B}$  führt von jedem Punkt von  $\mathfrak{N}$  nach jedem anderen Punkt von  $\mathfrak{N}$  mindestens eine *Bahn*.

**b)** Nach Streichen einer beliebigen Kante von  $\mathfrak{B}$  geht die Eigenschaft **a)** verloren.

Wenn wir kurz über eine Kantenbasis  $\mathfrak{B}$  sprechen (ohne die Angabe des Netzes, wofür  $\mathfrak{B}$  eine Kantenbasis ist), so meinen wir damit, daß  $\mathfrak{B}$  ein (gerichteter) Graph ist, der eine Kantenbasis für irgendein Netz  $\mathfrak{N}$  ist; da aber dann  $\mathfrak{B}$  und  $\mathfrak{N}$  dieselben Punkte enthalten, so ist dieses  $\mathfrak{N}$  durch  $\mathfrak{B}$  eindeutig bestimmt.

Die Bedeutung der Kantenbasen liegt vor allem in ihren Anwendungen in der Logik.<sup>2</sup> Handelt es sich nämlich um  $n$  äquivalente Aussagen

$$(1) \quad A_1, \dots, A_n$$

d. h. gelten für diese Aussagen alle (logischen) Implikationen

$$(2) \quad \overrightarrow{A_i A_k} \quad (i \neq k; k = 1, \dots, n)$$

<sup>1</sup> Mit [ ] verweisen wir an das Literaturverzeichnis am Schluß unserer Arbeit.

<sup>2</sup> Wir bemerken, daß sich die Kantenbasen auch für beliebige gerichtete Graphen definieren lassen (s. KÖNIG [2], S. 92) und daß sie auch im allgemeinen Fall die entsprechende Bedeutung in der Logik haben. S. hierüber die wichtigen Arbeiten von HERTZ [1], von denen ein Teil auch in KÖNIG [2] (S. 92—109) enthalten ist. Den Begriff der Kantenbasis werden wir hier weiter auch nur im oben definierten beschränkten Sinn gebrauchen.



(man lese: aus  $A_i$  folgt  $A_k$ ), deutet man dann (1), (2) als die Punkte und Kanten eines Netzes  $\mathfrak{N}$ , so entsprechen den Kantenbasen  $\mathfrak{B}$  von  $\mathfrak{N}$  (wegen der Transitivität der Implikationen (2)) eben die sämtlichen vollständigen unabhängigen Axiomensysteme für das Satzsystem (2). Das bedeutet mit anderen Worten, daß jede Kantenbasis  $\mathfrak{B}$  für  $\mathfrak{N}$  sozusagen ein „Beweisschema“ zum Beweis aller Implikationen (2), d. h. der Äquivalenz der Aussagen (1) liefert, wobei kein „Beweisschritt“ überflüssig ist; unter den Beweisschritten verstehen wir die in  $\mathfrak{B}$  enthaltenen Implikationen.<sup>3</sup>

Die allereinfachsten Kantenbasen für  $\mathfrak{N}$  sind von der folgenden Form (veranschaulicht wird den Fall  $n=6$ ):



Figur 1.



Figur 2.



Figur 3.

Solche Kantenbasen nennen wir (im Fall eines beliebigen  $n$ ) der Reihe nach *zyklisch*, *linear*, bzw. *sternförmig*. Für diese ist die Anzahl der Kanten gleich  $n$ ,  $2(n-1)$ , bzw.  $2(n-1)$ . Eine leichte Folgerung aus Satz 2 (s. unten) ist, daß die Anzahl der Kanten einer Kantenbasis stets zwischen diesen Zahlen  $n$ ,  $2(n-1)$  liegt und nur für die zyklischen Kantenbasen gleich  $n$  ist. (Das letztere sieht man auch unmittelbar leicht ein.)

In den Fällen  $n=2, 3$  gibt es im wesentlichen offenbar keine weiteren Kantenbasen als die obigen drei. (Von ihnen fallen natürlich im Fall  $n=2$  alle drei, im Fall  $n=3$  die letzten zwei zusammen.)

Im Fall  $n=4$  sind die sämtlichen Kantenbasen offenbar von der folgenden Form:



Figur 4.



Figur 5.



Figur 6.



Figur 7.



Figur 8.

<sup>3</sup> Andere Deutungen der Kantenbasen liegen auf der Hand, wie z. B. die folgenden zwei, auf die mich die Herren L. KALMÁR bzw. A. RÉNYI freundlichst aufmerksam gemacht haben. *Erstens* sollen  $A_1, \dots, A_n$  reelle Zahlen bedeuten, ferner bedeute (2) dasselbe wie  $A_i \geq A_k$ , und es handele sich um den Beweis von  $A_1 = \dots = A_n$ . *Zweitens* sollen  $A_1, \dots, A_n$  verschiedene Punkte bezeichnen, wischen denen man ein Verkehrssystem mit einseitigem Verkehr anlegen will. (Mit dem letzteren meinen wir, daß auf jeder Verkehrsstrecke eine Verkehrsrichtung vorgeschrieben ist.) In beiden Fällen liefern die Kantenbasen wieder die „besten“ Lösungen der Aufgabe.



Im Fall  $n = 5$  findet man mit wenig Mühe, daß dann die Kantenbasen von der folgenden Form sind:



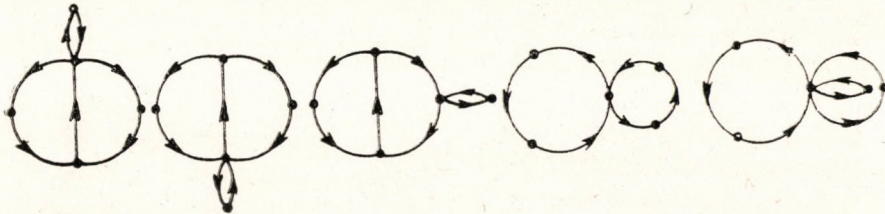
Figur 9.

Figur 10.

Figur 11.

Figur 12.

Figur 13.



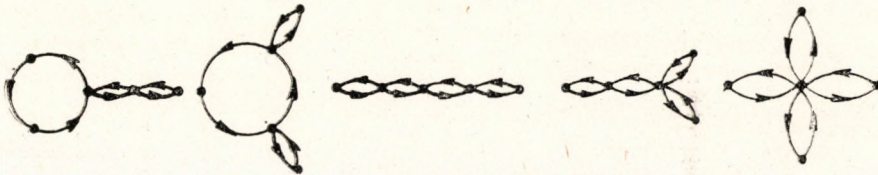
Figur 14.

Figur 15.

Figur 16.

Figur 17.

Figur 18.



Figur 19.

Figur 20.

Figur 21.

Figur 22.

Figur 23.

Für den allgemeinen Fall machen wir die folgende nützliche Bemerkung. Ersetzt man eine Kante  $\overrightarrow{XY}$  von einem Zyklus durch  $\overrightarrow{YX}$ , so nennen wir den entstandenen Graphen einen *Fastzyklus*. Dann ist klar, daß eine *Kantenbasis keinen Fastzyklus enthält*.

Unser Hauptresultat wird der folgende:<sup>4</sup>

**SATZ 1.** *Jede Kantenbasis ist ein ebener Graph.* (Unter einem ebenen Graphen verstehen wir einen solchen, der sich ohne Durchkreuzung in einer Ebene zeichnen läßt.)

Einen guten Überblick über alle Kantenbasen konnten wir uns nicht verschaffen, aber im unten folgenden Satz 2 geben wir ein rekursives Verfahren an, mit dem man alle Kantenbasen für ein gegebenes Netz bestimmen kann. Um diesen Satz formulieren zu können, schicken wir folgendes voran.

Ist  $\mathfrak{B}$  eine Kantenbasis und  $\overrightarrow{XY}$  eine Kante von ihr, so bezeichnen wir mit  $\mathfrak{B}_{XY}$  denjenigen Graphen, der aus  $\mathfrak{B}$  nach der Entfernung der Kante  $\overrightarrow{XY}$  übrigbleibt. Man nehme den Punkt  $X$ , ferner die sämtlichen Bahnen in  $\mathfrak{B}_{XY}$  mit dem Anfangspunkt  $X$ , und addiere sie; ähnlich addiere man den Punkt

<sup>4</sup> Für die in <sup>2</sup> erwähnte allgemeinere Fassung der Kantenbasis ist Satz 1 offenbar falsch.



$Y$  und die sämtlichen Bahnen in  $\mathfrak{B}_{XY}$  mit dem Schlußpunkt  $Y$ .<sup>5</sup> Wir bezeichnen die so entstandenen zwei Teilgraphen von  $\mathfrak{B}_{XY}$  mit  $\mathfrak{B}'_{XY}$  bzw.  $\mathfrak{B}''_{XY}$ . Diese zwei Graphen sind stets fremd, denn sonst würde  $\mathfrak{B}$  einen Fastzyklus enthalten, was nicht möglich ist.

**SATZ 2.** *Es sei gegeben ein Netz  $\mathfrak{N}$  von der Ordnung  $n (\geq 3)$ . Um die nichtzyklischen Kantenbasen für  $\mathfrak{N}$  zu gewinnen, verfähre man so. Man nehme eine in  $\mathfrak{N}$  enthaltene Kantenbasis  $\mathfrak{B}$ , die nicht alle Punkte von  $\mathfrak{N}$  enthält. Man nehme dann zwei Punkte  $P, Q$  von  $\mathfrak{B}$ , die im Fall  $P \neq Q$  der Bedingung unterworfen sind, daß für keine Kante  $\overrightarrow{XY}$  von  $\mathfrak{B}$  gleichzeitig  $P$  in  $\mathfrak{B}'_{XY}$  und  $Q$  in  $\mathfrak{B}''_{XY}$  gehört,<sup>6</sup> nehme ferner eine (in  $\mathfrak{N}$  enthaltene) Bahn  $b$ , die von  $P$  nach  $Q$  führt und (außer  $P, Q$ ) nur noch die sämtlichen, außerhalb  $\mathfrak{B}$  liegenden Punkte von  $\mathfrak{N}$  enthält. Dann ist  $\mathfrak{B} + b$  eine Kantenbasis für  $\mathfrak{N}$ . Alle nichtzyklischen Kantenbasen für  $\mathfrak{N}$  entstehen auf diesem Wege.*

**ERGÄNZUNG.** *Es genügt im Satz 2 diejenigen Kanten  $\overrightarrow{XY}$  zu berücksichtigen, für die  $X$  gleich  $P$  oder ein Verzweigungspunkt<sup>7</sup> von  $\mathfrak{B}$  und zugleich der Anfangspunkt einer von  $\overrightarrow{XY}$  verschiedenen Kante in  $\mathfrak{B}$  ist, ferner  $Y$  gleich  $Q$  oder ein Verzweigungspunkt von  $\mathfrak{B}$  und zugleich der Schlußpunkt einer von  $\overrightarrow{XY}$  verschiedenen Kante in  $\mathfrak{B}$  ist.*

Der Satz 2 scheint nicht geeignet zu sein, um mit seiner Hilfe die Anzahl aller Kantenbasen für ein Netz  $n$ -ter Ordnung zu bestimmen. Unmittelbar auf Grund der Definition läßt sich leicht für diese Anzahl die Formel

$$N_n = \sum_{\substack{c_{uv}=0,1 \\ u,v=1,\dots,n \\ u \neq v}} \prod_{\substack{r,s=1,\dots,n \\ r \neq s}} \left( \prod_{\substack{i,k=1,\dots,n \\ i \neq k}} (1 - \prod_{C_{ik}} (1 - C_{ik})) - c_{rs} \prod_{\substack{i,k=1,\dots,n \\ i \neq k}} (1 - \prod_{C_{ik}^{rs}} (1 - C_{ik}^{rs})) \right)$$

aufstellen. Hier bedeutet  $C_{ik}$  alle Produkte von der Form

$$C_{ik} = c_{ix_1} c_{x_1 x_2} \cdots c_{x_{t-1} x_t} c_{x_t k} \quad (0 \leq t \leq n-2; x_1, \dots, x_t = 1, \dots, n),$$

wobei die  $x_1, \dots, x_t$  voneinander und von  $i, k$  verschieden sein sollen; die  $C_{ik}^{rs}$  bezeichnen diejenigen  $C_{ik}$ , in denen der Faktor  $c_{rs}$  nicht auftritt. Da uns aber nicht gelang, diese Formel auf eine einfachere Form zu bringen, so halten wir es nicht der Mühe wert, sie zu beweisen. Als Beispiele erwähnen wir  $N_2 = 1, N_3 = 5, N_4 = 58, N_5 = 1069$ .

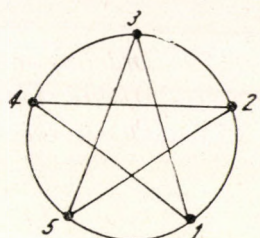
<sup>5</sup> Ist  $b$  eine Bahn, die vom Punkt  $P$  nach dem Punkt  $Q$  führt, so nennen wir  $P$  den Anfangspunkt und  $Q$  den Schlußpunkt von  $b$ . (Insbesondere wenn  $b$  ein Zyklus ist, so kann  $P = Q$  ein beliebiger Punkt von  $b$  sein.)

<sup>6</sup> Folglich kommen nur solche  $P, Q$  in Frage, für die  $\mathfrak{B}$  keine Kante  $\overrightarrow{PQ}$  enthält.

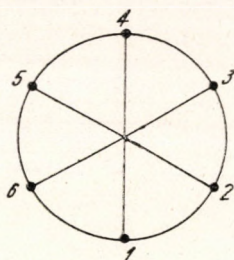
<sup>7</sup> Unter einem Verzweigungspunkt eines Graphen verstehen wir einen solchen Punkt dieses Graphen, der zu mindestens drei Kanten gehört.



BEWEIS VOM SATZ 1. Wir betrachten die folgenden zwei Graphen 5-ter bzw. 6-ter Ordnung mit 10 bzw. 9 Kanten (Figuren 24, 25):



Figur 24.



Figur 25.

Wir bezeichnen sie mit  $\mathbb{G}_5$  bzw.  $\mathbb{G}_6$ .<sup>8</sup> Nach KURATOWSKI [3] lassen sich die endlichen ebenen Graphen dadurch charakterisieren, daß sie keine  $\mathbb{G}_5, \mathbb{G}_6$  enthalten.

Werden alle Kanten eines Graphen  $\mathbb{G}$  irgendwie gerichtet, so bezeichnen wir den entstandenen gerichteten Graphen mit  $\vec{\mathbb{G}}$ . Wir haben dann zu beweisen, daß eine Kantenbasis  $\mathfrak{B}$  weder ein  $\vec{\mathbb{G}}_5$  noch ein  $\vec{\mathbb{G}}_6$  enthalten kann.

Mit der ersten dieser Behauptungen werden wir leicht fertig sein. Wir nehmen an, daß ein  $\vec{\mathbb{G}}_5$  in einer Kantenbasis  $\mathfrak{B}$  enthalten ist, um hieraus einen Widerspruch abzuleiten. Wie man auch die drei Seiten eines Dreiecks mit Richtungen versieht, entsteht immer ein Zyklus oder ein Fastzyklus. Da eine Kantenbasis keine Fastzyklen enthält, so müssen die Kanten von  $\vec{\mathbb{G}}_5$  so gerichtet sein, daß für alle zehn Punkttripel

$$i, j, k \quad (1 \leq i < j < k \leq 5)$$

entweder  $\vec{ij}, \vec{jk}, \vec{ki}$  oder  $\vec{ji}, \vec{kj}, \vec{ik}$  einen Zyklus in  $\vec{\mathbb{G}}_5$  bilden. Man darf nun annehmen, daß  $\vec{\mathbb{G}}_5$  die gerichtete Kante  $\vec{12}$  enthält. Dann muß  $\vec{\mathbb{G}}_5$  auch  $\vec{23}, \vec{25}$  enthalten. Das ergibt für das Punkttripel 2, 3, 5 einen Widerspruch mit dem vorher Gesagten.

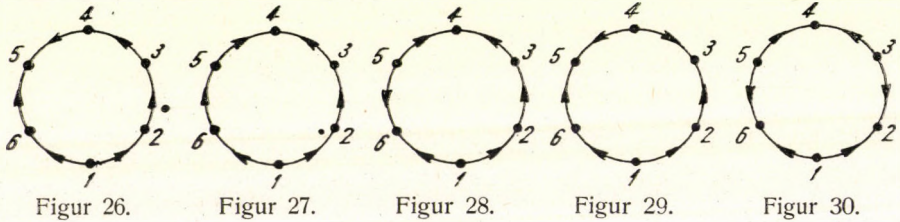
Um die zweite Behauptung zu beweisen, nehmen wir jetzt an, daß ein  $\vec{\mathbb{G}}_6$  in einer Kantenbasis  $\mathfrak{B}$  enthalten ist. Wir haben auch hieraus einen Widerspruch abzuleiten. Da jede Kante von  $\mathbb{G}_6$  auf zwei Arten gerichtet werden kann, so gibt es für  $\vec{\mathbb{G}}_6$  insgesamt  $2^9$  Fälle, aber wir werden es in der Wirklichkeit nur mit ganz wenigen Fällen zu tun haben. Vor allem bemerken wir, daß zwei entgegengesetzte Graphen<sup>9</sup> stets nur gleichzeitig Kantenbasen sein

<sup>8</sup> Hiervon ist  $\mathbb{G}_5$  das „vollständige Fünfeck“ und  $\mathbb{G}_6$  der zum sogenannten Dreibrunnenproblem gehörige Graph.

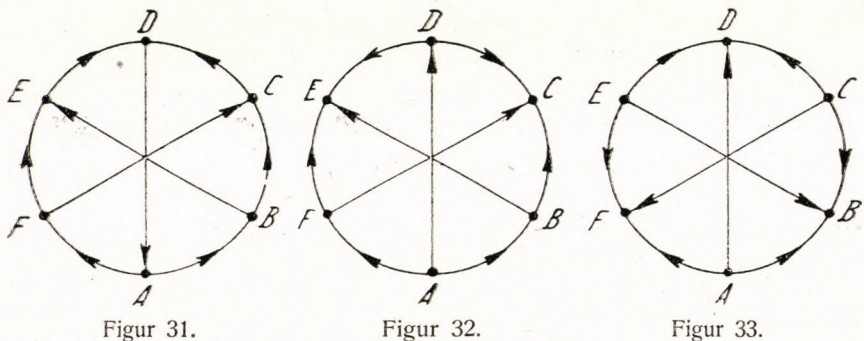
<sup>9</sup> So nennen wir zwei gerichtete Graphen, wenn der eine aus dem anderen durch Umkehrung aller Kantenrichtungen entsteht.



können. Folglich brauchen wir von zwei entgegengesetzten  $\vec{\mathcal{G}}_6$  immer nur das eine zu berücksichtigen. Ferner nehmen wir in Acht, daß jeder Zyklus eine Kantenbasis ist. Aus diesem Grunde können die Punkte von einem in  $\mathfrak{B}$  enthaltenen Zyklus mit keinen Kanten verbunden sein, außer den Kanten, die diesen Zyklus bilden. Da aber  $\vec{\mathcal{G}}_6$  mehr Kanten als Punkte enthält, so folgt hieraus, daß es in  $\vec{\mathcal{G}}_6$  keinen Zyklus mit sechs Punkten geben kann. Auch enthält  $\vec{\mathcal{G}}_6$  keinen Fastzyklus, denn ein solcher müßte auch in  $\mathfrak{B}$  enthalten sein, was nach obigem unmöglich ist. Nunmehr wollen wir vorläufig nur beachten, wie in  $\vec{\mathcal{G}}_6$  die Kanten des Sechsecks 12...6 gerichtet sind. Nach obigem brauchen wir nur die folgenden Fälle zu berücksichtigen:



(Diese fünf Graphen und ihre entgegengesetzten erschöpfen nämlich bis auf Lage alle überhaupt möglichen Fälle.) Diese fünf Graphen müssen noch mit je drei gerichteten Kanten ergänzt werden, die die gegenüberliegenden Punkte miteinander verbinden. Da aber  $\vec{\mathcal{G}}_6$  keinen Fastzyklus bestehend aus vier Kanten enthalten kann, so erweisen sich der erste und dritte dieser fünf Fälle als unmöglich, ferner muß  $\vec{\mathcal{G}}_6$  aus demselben Grund in den übriggebliebenen drei Fällen so aussehen:



(Dabei haben wir die Punkte des Graphen mit  $A, \dots, F$  umbezeichnet.) Der erste dieser drei Fälle ist unmöglich, da dann  $\mathcal{G}_6$  den Fastzyklus  $BCDAFE$  enthält. Für den mittleren Fall bemerken wir, daß es dann in  $\mathfrak{B}$



offenbar eine Bahn  $b$  geben muß, die von  $C$  nach einem von  $C$  verschiedenen Punkt  $H$  von  $\mathfrak{G}_6$  führt und sonst keinen gemeinsamen Punkt mit  $\mathfrak{G}_6$  hat. Es ist  $H=A$  unmöglich, da dann  $\mathfrak{B}$  den Fastzyklus enthielte, der aus der Bahn  $b$  und den Kanten  $\overrightarrow{AF}, \overrightarrow{FE}, \overrightarrow{BE}, \overrightarrow{BC}$  besteht. Ebenso leicht ergibt sich, daß  $H$  aus ähnlichem Grunde keiner der übrigen Punkte  $B, D, E, F$  sein kann. Auf den ähnlichen Widerspruch führt auch der letzte der obigen drei Fälle, wobei nämlich jetzt z. B. der Punkt  $B$  die vorige Rolle von  $C$  übernehmen kann. Damit haben wir Satz 1 bewiesen.

BEWEIS VOM SATZ 2. Wir zeigen zuerst, daß auf dem im Satz 2 beschriebenen Wege stets eine Kantenbasis für  $\mathfrak{N}$  entsteht. Wir haben zu zeigen, daß  $\mathfrak{B}+b$  eine Kantenbasis für  $\mathfrak{N}$  ist.

Da  $\mathfrak{B}$  eine Kantenbasis und  $b$  eine Bahn ist, die von  $P$  ( $\in \mathfrak{B}$ ) nach  $Q$  ( $\in \mathfrak{B}$ ) führt, ferner  $\mathfrak{B}+b$  alle Punkte von  $\mathfrak{N}$  enthält, so ist klar, daß für  $\mathfrak{B}+b$  die Eigenschaft **a)** gilt.

Um noch die Eigenschaft **b)** auszuweisen, haben wir zu zeigen, daß die Eigenschaft **a)** von  $\mathfrak{B}+b$  verlorengeht, wenn aus ihm eine Kante entfernt wird. Handelt es sich um eine Kante von  $b$ , so ist die Behauptung klar. Es ist nur noch der Fall übrig, daß es sich um eine Kante von  $\mathfrak{B}$  handelt. Diese bezeichnen wir (wie im Satz) mit  $\overrightarrow{XY}$ . Wie auch früher bezeichnen wir mit  $\mathfrak{B}_{XY}$  denjenigen Graphen, der aus  $\mathfrak{B}$  nach Entfernung der Kante  $\overrightarrow{XY}$  übrigbleibt. Dann ist  $\mathfrak{B}_{XY}+b$  eben derjenige Graph, der aus  $\mathfrak{B}+b$  nach Entfernung von  $\overrightarrow{XY}$  übrigbleibt. Es genügt zu zeigen, daß in  $\mathfrak{B}_{XY}+b$  keine Bahn von  $X$  nach  $Y$  führt.

Nehmen wir das Gegenteil an, daß nämlich in  $\mathfrak{B}_{XY}+b$  eine Bahn  $b_1$  von  $X$  nach  $Y$  führt. Andererseits führt in  $\mathfrak{B}_{XY}$  keine Bahn von  $X$  nach  $Y$ , denn sonst würde diese Bahn mit  $\overrightarrow{XY}$  zusammen einen in  $\mathfrak{B}$  enthaltenen Fastzyklus bilden, das aber unmöglich ist, da  $\mathfrak{B}$  eine Kantenbasis ist. Aus beiden folgt offenbar, daß  $b$  in  $b_1$  enthalten ist. Folglich zerlegt sich  $b_1$  in drei (anschließende) Bahnen, von denen die erste von  $X$  nach  $P$  führt, die zweite gleich  $b$  ist, die dritte von  $Q$  nach  $Y$  führt. (Die erste und dritte kann sich im Punkt  $P$  bzw.  $Q$  ausarten.) Dann gehört  $P$  in  $\mathfrak{B}'_{XY}$  und  $Q$  in  $\mathfrak{B}''_{XY}$ . Das verstößt in die Definition von  $P, Q$ , womit wir bewiesen haben, daß  $\mathfrak{B}+b$  auch die Eigenschaft **b)** hat, also eine Kantenbasis für  $\mathfrak{N}$  ist.

Zum vollständigen Beweis des Satzes 2 fehlt nur noch der Nachweis, daß umgekehrt jede nichtzyklische Kantenbasis  $\mathfrak{B}_1$  von  $\mathfrak{N}$  auf dem im Satz beschriebenen Wege entsteht. Da für irgendzwei Punkte  $A, B$  von  $\mathfrak{B}_1$  zwei Bahnen in diesem vorhanden sind, die von  $A$  nach  $B$  bzw. von  $B$  nach  $A$  führen, so ist klar, daß  $\mathfrak{B}_1$  einen Zyklus enthält. Dieser ist wegen der Annahme ein echter Teilgraph von  $\mathfrak{B}_1$ . Wir haben bekommen, daß es unter den echten Teilgraphen von  $\mathfrak{B}_1$  jedenfalls auch Kantenbasen vorkommen.



Wir betrachten also alle echten Teilnetze  $\mathcal{N}'$  von  $\mathcal{N}$  von der Eigenschaft, daß  $\mathfrak{B}_1$  eine Kantenbasis für  $\mathcal{N}'$  enthält. Unter diesen  $\mathcal{N}'$  wählen wir ein maximales heraus und bezeichnen mit  $\mathfrak{B}$  eine Kantenbasis für dieses  $\mathcal{N}'$ . Dann gibt es einen Punkt  $A$  von  $\mathcal{N}$  außerhalb von  $\mathfrak{B}$ . Da  $\mathfrak{B}_1$  eine Kantenbasis für  $\mathcal{N}$  ist, so gibt es zwei Bahnen  $b_1, b_2$  in  $\mathfrak{B}_1$ , so daß  $b_1$  von einem Punkt  $P (\in \mathfrak{B})$  nach  $A$  und  $b_2$  von  $A$  nach einem Punkt  $Q (\in \mathfrak{B})$  führt. Offenbar darf angenommen werden, daß  $b_1, b_2$  außer  $P$  bzw.  $Q$  kein Element mit  $\mathfrak{B}$  gemeinsam haben. Die Vereinigung von  $b_1, b_2$  enthält also eine Bahn  $b$ , die von  $P$  nach  $Q$  führt und sonst kein Element von  $\mathfrak{B}$  enthält, ferner durch mindestens einen außerhalb von  $\mathfrak{B}$  liegenden Punkt von  $\mathcal{N}$  geht. Dabei ist mit  $b_1, b_2$  zusammen auch  $b$  in  $\mathfrak{B}_1$  enthalten. Wir zeigen, daß  $\mathfrak{B}_1 = \mathfrak{B} + b$  ist.

Da nämlich  $\mathfrak{B}$  eine Kantenbasis ist und die Bahn  $b$  zwei Punkte von  $\mathfrak{B}$  miteinander verbindet, so enthält  $\mathfrak{B} + b$  offenbar eine Kantenbasis für ein Netz  $\mathcal{N}_0 (\subseteq \mathcal{N})$ , das dieselben Punkte hat wie  $\mathfrak{B} + b$ . Andererseits enthält  $\mathcal{N}'$  dieselben Punkte wie  $\mathfrak{B}$ . Da ferner  $b$  mindestens einen Punkt außerhalb  $\mathfrak{B}$  enthält, so folgt, daß  $\mathcal{N}'$  in  $\mathcal{N}_0$  echt enthalten ist. Da  $\mathfrak{B}_1$  nach vorigem eine Kantenbasis für  $\mathcal{N}_0$  enthält, so muß wegen der Maximaleigenschaft von  $\mathcal{N}'$  die Gleichung  $\mathcal{N}_0 = \mathcal{N}'$  gelten. Somit enthält  $\mathfrak{B} + b (\subseteq \mathfrak{B}_1)$  eine Kantenbasis für  $\mathcal{N}$ . Da aber auch  $\mathfrak{B}_1$  eine Kantenbasis für  $\mathcal{N}$  ist, so folgt die Richtigkeit der Behauptung  $\mathfrak{B}_1 = \mathfrak{B} + b$ .

Dies hat auch zur Folge, daß  $b$  die sämtlichen, außerhalb  $\mathfrak{B}$  liegenden Punkte von  $\mathcal{N}$  enthält. Würde nun für eine Kante  $\overrightarrow{XY} (\in \mathfrak{B})$  sowohl  $P \in \mathfrak{B}'_{XY}$  als auch  $Q \in \mathfrak{B}''_{XY}$  gelten, so würde hieraus offenbar folgen, daß durch die Kante  $\overrightarrow{XY}$  ein in  $\mathfrak{B}$  enthaltener Fastzyklus geht, das aber unmöglich ist, da  $\mathfrak{B}$  eine Kantenbasis ist. Wir haben bekommen, daß die Kantenbasis  $\mathfrak{B}_1$  auf dem im Satz 2 beschriebenen Wege entsteht. Das beendet den Beweis vom Satz 2.

Auch die Ergänzung vom Satz 2 ist richtig, denn sind für eine Kante  $\overrightarrow{XY} (\in \mathfrak{B})$  die in dieser Ergänzung gesagten Bedingungen nicht erfüllt, so können  $P \in \mathfrak{B}'_{XY}, Q \in \mathfrak{B}''_{XY}$  offenbar nicht beide erfüllt sein.

(Eingegangen am 3. Februar 1954.)

### Literaturverzeichnis

- [1] P. HERTZ, Über Axiomensysteme für beliebige Satzsysteme, Teile I, II, *Math. Annalen*, **87** (1922), S. 246—269 bzw. **89** (1923), S. 76—102.
- [2] D. KÖNIG, *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936), S. 1—258.
- [3] C. KURATOWSKI, Sur le problème des courbes gauche en Topologie, *Fundamenta Math.*, **15** (1930), S. 271—283.



## О ФАЗИСАХ РЕБЕР КОНЕЧНЫХ ОРИЕНТИРОВАННЫХ ПОЛНЫХ ГРАФОВ

Л. РЭДЭИ (Сегед)

(Резюме)

Обозначим через  $\mathfrak{N}$  ориентированный полный граф, состоящий из  $n$  точек ( $n = 2, 3, \dots$ ) (т. е. каждая пара точек из  $\mathfrak{N}$  соединена двумя противоположно направленными ребрами). Как обычно, мы называем базисом ребер  $\mathfrak{B}$  графа  $\mathfrak{N}$  его подграф, обладающий следующими свойствами:

- а) из любой точки в любую другую его точку ведет путь, содержащийся в  $\mathfrak{B}$ ;
- б) при отбрасывании любого ребра из  $\mathfrak{B}$  теряется свойство а).

(Знание базисов ребер равносильно знанию всех независимых систем аксиом системы теорем

$$A_1 \rightarrow A_2 \quad (i, k = 1, \dots, n; ik)$$

и таким образом оно дает способы доказательства эквивалентности конечного числа утверждений  $A_1, \dots, A_n$ . См. Герц [1].)

**Теорема.** Любой базис ребер  $\mathfrak{B}$  графа  $\mathfrak{N}$  является полным графом.

Доказательство основано на известной теореме Куратовского [3].

Трудно дать полный обзор всех базисов ребер, однако можно указать простой способ определения всех базисов ребер с помощью рекуррентного процесса.







## ÜBER DAS KREISTEILUNGSPOLYNOM

Von

LADISLAUS RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

Bezeichne  $I$  den Ring der ganzen rationalen Zahlen,  $I[x]$  den Ring der Polynome  $f(x)$  mit Koeffizienten in  $I$ , ferner  $n (> 1)$  eine natürliche Zahl und  $F_n(x) (\in I[x])$ , das  $n$ -te Kreisteilungspolynom. In einer früheren Arbeit<sup>1</sup> habe ich den folgenden Satz ausgesprochen: *Das Hauptideal  $(F_n(x))$  wird durch die  $F_p\left(x^{\frac{n}{p}}\right) = (x^n - 1)\left(x^{\frac{n}{p}} - 1\right)^{-1}$  ( $p$  Primzahl,  $p|n$ ) erzeugt:*

$$(1) \quad (F_n(x)) = \left(F_p\left(x^{\frac{n}{p}}\right), \dots\right).$$

Mein Beweis war aber fehlerhaft, wie das BRUIJN<sup>2</sup> bemerkt hat. Er teilte für den Satz gleichzeitig auch einen Beweis mit. Der hier folgende Beweis wird anders und leichter als der Bruijnsche.

Ist  $n$  die Potenz einer Primzahl  $p$ , so ist  $F_n(x) = F_n\left(x^{\frac{n}{p}}\right)$ , weshalb jetzt der Satz trivial ist. Wir nehmen (1) für ein  $n$  an. Es genügt zu zeigen, daß dann (1) für  $nQ$  (statt  $n$ ) richtig ist, wobei  $Q$  die Potenz einer Primzahl  $q$  ist, für die  $q \nmid n$  gilt.

Aus (1) folgt offenbar

$$(2) \quad (F_n(x^Q)) = \left(F_p\left(x^{\frac{nQ}{p}}\right), \dots\right).$$

Wir setzen

$$(3) \quad f(x) = F_n(x^Q)F_{nQ}^{-1}(x), \quad g(x) = F_q\left(x^{\frac{nQ}{q}}\right)F_{nQ}^{-1}(x).$$

Hieraus folgt

$$(4) \quad f(x) = \prod_{\varrho} (x - \varrho), \quad g(x) = \prod_{\sigma} (x - \sigma),$$

wobei  $\varrho, \sigma$  die komplexen Einheitswurzeln mit

$$(5) \quad O(\varrho) = nQ, \quad O(\sigma) = n'Q \quad (n'|n, n' < n, Q'|Q, Q' < Q)$$

<sup>1</sup> L. RÉDEI, Ein Beitrag zum Problem der Faktorisierung von endlichen Abelschen Gruppen, *Acta Math. Acad. Sci. Hung.*, 1 (1950), S. 197–207, Hilfssatz 4.

<sup>2</sup> N. G. DE BRUIJN, On the factorization of cyclic groups, *Proc. Kon. Ned. Akad. v. Wetensch.*, Ser. A, 56 (1953), S. 370–377, § 2. — Sowohl ich in der Arbeit<sup>1</sup> als auch BRUIJN in seiner Arbeit<sup>2</sup> haben wir beide den obigen Satz (1) zu Untersuchungen eines Faktorisierungsproblems von HAJÓS verwendet.



durchlaufen und  $O$  die Ordnung der Elemente in der Gruppe aller komplexen Einheitswurzeln bezeichnet. Wegen (3), (4) gilt auch  $f(x), g(x) \in I[x]$ . Die Resultante der Polynome (4) ist:

$$R = \prod_{\varrho, \sigma} (\varrho - \sigma).$$

Da nach (5)  $O(\varrho\sigma^{-1})$  durch mindestens zwei Primzahlen teilbar ist, so ist die ganze algebraische Zahl  $\varrho - \sigma (= \sigma(\varrho\sigma^{-1} - 1))$  eine Einheit, weshalb  $R = \pm 1$  sein muß. Also gilt (1)  $= (f(x), g(x))$ . Hierfür läßt sich nach (3)

$$(F_{n\varrho}(x)) = \left( F_n(x^\varrho), F_q \left( x^{\frac{n\varrho}{q}} \right) \right)$$

schreiben. Dies besagt wegen (2) die Richtigkeit von (1) für  $nQ$  (statt  $n$ ), womit der Satz bewiesen ist.

(Eingegangen am 3. Februar 1954.)

## О ЦИКЛОТОМИЧЕСКОМ ПОЛИНОМЕ

Л. РЭДЭИ (Сегед)

(Резюме)

В настоящей работе сообщается краткое доказательство теоремы, соответственно которой имеет место равенство между идеалами

$$(F_n(x)) = \left( F_p \left( x^{\frac{n}{p}} \right), \dots \right),$$

где  $F_n(x)$  означает  $n$ -ый циклотомический полином, а  $p$  пробегает все различные простые делители  $n$ . Эта теорема раньше была использована автором при исследовании теоретико-групповой факторизационной проблемы, поставленной Hajós. Она нашла аналогичное применение и недавно, у N. G. de Bruijn. Первое доказательство этой теоремы, данное автором, было неполно, как на это указал de Bruijn, дающий полное доказательство. Настоящее доказательство является более простым, чем доказательство de Bruijn.



# ÜBER DIE TRIGONOMETRIE DES POINCARÉSCHEN KREIS- MODELLS DER HYPERBOLISCHEN EBENEN GEOMETRIE

Von

PAUL SZÁSZ (Budapest)

(Vorgelegt von G. HAJÓS)

Eine bekannte Verwirklichung der hyperbolischen ebenen Geometrie im Rahmen der euklidischen Geometrie ist die folgende *Pseudogeometrie* (oder *Bildgeometrie*), die sich durch die Arbeiten von H. POINCARÉ<sup>1</sup> verbreitete und *Poincarésches Kreismodell* dieser Geometrie genannt wird.

Es sei in der euklidischen Ebene ein Kreis  $K$  als *Fundamentalkreis* vorgelegt. Die inneren Punkte von  $K$  sollen *Pseudopunkte*, die im Innern von  $K$  liegenden Teile der  $K$  rechtwinklig schneidenden Kreise bzw. Geraden (die wir zusammen *Orthogonalkreise* nennen wollen) *Pseudogeraden* heißen. Den *Pseudoabstand* zweier Pseudopunkte  $P_1, P_2$  heie die Groe

$$(1) \quad \overline{P_1 P_2} = \log (UVP_2 P_1),$$

wobei  $U, V$  die Schnittpunkte von  $K$  und dem durch  $P_1$  und  $P_2$  gelegten Orthogonalkreis sind, so bezeichnet, da  $P_2$  auf dem Bogen  $\widehat{UV}$  zwischen

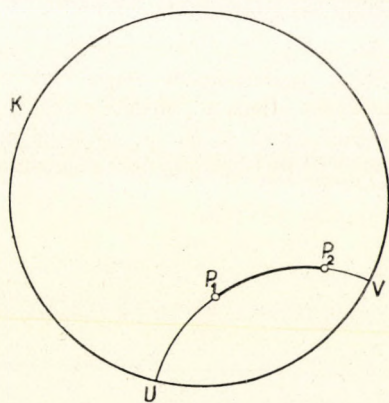


Fig. 1.

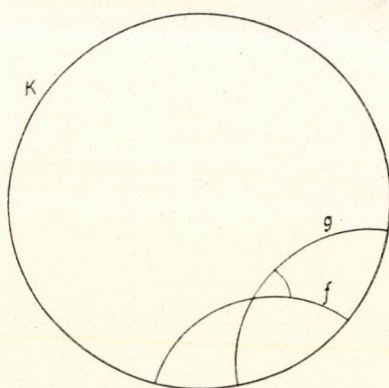


Fig. 2.

<sup>1</sup> H. POINCARÉ, Théorie des groupes fuchsien, *Acta Mathematica* (Stockholm), 1 (1882), S. 1—62, besonders § 2, S. 6—8, und § 12, S. 58—61; Mémoire sur les fonctions fuchsiennes, *ebenda* S. 193—294, besonders S. 201—202; Mémoire sur les groupes kleinéens, *ebenda*, 3 (1883), S. 49—92, besonders S. 55—56.



$P_1$  und  $V$  liegt (Fig. 1), und  $(UVP_2P_1)$  das Doppelverhältnis

$$(UVP_2P_1) = \frac{UP_2}{P_2V} : \frac{UP_1}{P_1V}$$

bedeutet. Dem gegenüber soll der *Pseudowinkel* zweier Pseudogeraden  $f, g$  einfach durch den Winkel der sie enthaltenden Orthogonalkreise gemessen werden (Fig. 2). Man überzeugt sich leicht, daß in der durch diese Festsetzungen erklärten Pseudogeometrie alle Postulate der hyperbolischen ebenen Geometrie erfüllt sind.

HOWARD EVES und V. E. HOGGATT<sup>2</sup> haben die Trigonometrie dieses POINCARÉschen Kreismodells von neuem hergeleitet. Im folgenden wird für diese Herleitung wieder eine andere Methode angegeben, die uns viel einfacher zu sein scheint. Man darf aber nicht vergessen, daß eine derartige Herleitung wie die von HOWARD EVES und V. E. HOGGATT oder von uns, erst dann für einen Beweis der Formeln der hyperbolischen Trigonometrie gelten kann, wenn man vorher die Äquivalenz der hyperbolischen ebenen Geometrie mit dieser Poincaréschen Bildgeometrie schon erwiesen hat. Das kann in verschiedener Weise geschehen.<sup>3</sup>

### § 1. Vorbereitung

Der Fundamentalkreis  $K$  habe den Radius  $r$  und bezeichne  $O$  seinen Mittelpunkt. Für einen Pseudopunkt  $P \neq O$  kann der Abstand  $OP$  durch den Pseudoabstand  $\overline{OP} = t$  leicht ausgedrückt werden. Schneidet nämlich die Gerade  $OP$  den Fundamentalkreis in den Punkten  $U_1, V_1$  wobei  $P$  zwischen  $O$  und  $V_1$  liegt (Fig. 3), so ist im Sinne der Definition unter (1)

$$t = \log(U_1V_1PO) = \log \frac{U_1P}{PV_1} \frac{OV_1}{U_1O}$$

oder wegen  $U_1O = OV_1$

$$t = \log \frac{U_1P}{PV_1} = \log \frac{1 + \frac{OP}{r}}{1 - \frac{OP}{r}}$$

woher sich

$$\frac{OP}{r} = \frac{e^t - 1}{e^t + 1}$$

<sup>2</sup> HOWARD EVES und V. E. HOGGATT, Hyperbolic trigonometry derived from the Poincaré model, *The American Mathematical Monthly*, **58** (1951), S. 469—474.

<sup>3</sup> Siehe z. B. PAUL SZÁSZ, Über die Hilbertsche Begründung der hyperbolischen Geometrie, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 243—250.



d. h.

$$(2) \quad OP = r \operatorname{th} \frac{t}{2}$$

ergibt.

Um weiter gehen zu können, schicken wir einen Hilfssatz voran.

HILFSSATZ. Wird durch  $P$  ein Kreis gelegt, der  $K$  in den Punkten  $U, V$  rechtwinklig schneidet, so besteht, falls die Halbgerade  $OX$  den  $\sphericalangle UOV = 2\beta$  halbiert, für  $\sphericalangle XOP = \lambda$  die Formel

$$(3) \quad \frac{\cos \beta}{\cos \lambda} = \frac{2rOP}{r^2 + OP^2}.$$

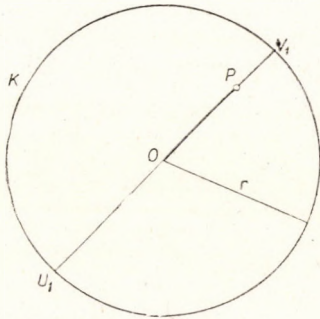


Fig. 3.

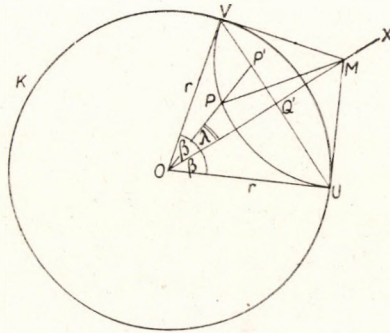


Fig. 4.

Bezeichnet nämlich  $M$  den Mittelpunkt des durch  $P$  gelegten Kreises (Fig. 4), so ist offenbar

$$PM = VM = r \operatorname{tg} \beta, \quad OM = \frac{r}{\cos \beta},$$

also ergibt der Cosinussatz für das Dreieck  $OMP$

$$r^2 \operatorname{tg}^2 \beta = OP^2 + \frac{r^2}{\cos^2 \beta} - \frac{2rOP}{\cos \beta} \cos \lambda,$$

woraus (3) folgt.

Wird die Gerade  $UV$  von  $OP$  in  $P'$ , von  $OX$  in  $Q'$  geschnitten, so ist

$$OP' = \frac{OQ'}{\cos \lambda} = \frac{r \cos \beta}{\cos \lambda},$$

also zieht (3) die Formel

$$(4) \quad OP' = \frac{2r^2 OP}{r^2 + OP^2}$$

nach sich. Setzt man hier  $OP$  aus (2) ein, so erhält man

$$(4^*) \quad OP' = r \operatorname{th} t.$$



Damit ist auch  $OP'$  durch den Pseudoabstand  $\overline{OP} = t$  ausgedrückt. Auf Grund dieser Formel (4) bzw. (4\*) ist der Schnittpunkt  $P'$  von der Wahl des durch  $P$  gelegten Orthogonalkreises unabhängig.<sup>4</sup>

Die Formeln (3) und (4\*) sind für das Folgende von grundlegender Bedeutung.

## § 2. Die Grundformeln der Trigonometrie des Poincaréschen Kreismodells

Ist  $P_1O_1Q_1$  ein bei  $Q_1$  rechtwinkliges Pseudodreieck und  $O_1 \neq O$ , so gibt es (wie man leicht einsieht) ein rechtwinkliges Pseudodreieck  $POQ$  ( $\sphericalangle Q = 90^\circ$ ), das mit  $P_1O_1Q_1$  pseudokongruent ist, d. h.

$$\overline{PO} = \overline{P_1O_1}, \quad \overline{OQ} = \overline{O_1Q_1}, \quad \overline{QP} = \overline{Q_1P_1}, \quad \sphericalangle P = \sphericalangle P_1, \quad \sphericalangle O = \sphericalangle O_1$$

ausfällt, wobei  $\overline{PO}$  etc. die unter (1) erklärten Pseudoabstände bedeuten.<sup>5</sup> Demnach können wir uns bei der Herleitung der trigonometrischen Grundformeln des rechtwinkligen Pseudodreiecks auf ein rechtwinkliges Pseudodreieck  $POQ$  ( $\sphericalangle Q = 90^\circ$ ) beschränken (Fig. 5).

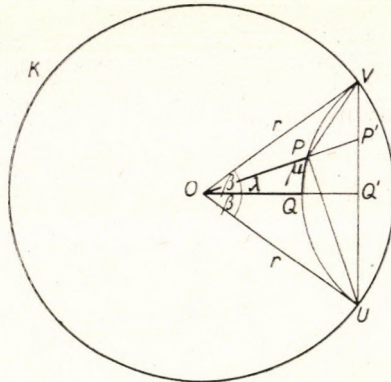


Fig. 5.

Es seien die Bestimmungsstücke dieses Pseudodreiecks

$$\overline{QP} = a, \quad \overline{OQ} = b, \quad \overline{OP} = c, \quad \sphericalangle QOP = \lambda, \quad \sphericalangle OPQ = \mu.$$

Die Pseudogerade  $PQ$  enthaltender Orthogonalkreis soll den Fundamentalkreis  $K$  in den Punkten  $U, V$  schneiden, wobei  $P$  auf dem Bogen  $\widehat{UV}$  zwischen  $Q$  und  $V$  liegt, und man setze wieder  $\sphericalangle UOV = 2\beta$ . Bezeichnet  $P'$

<sup>4</sup> Diese Tatsache ist übrigens eine Folge des bekannten Satzes, laut welchem die Potenzlinien dreier Kreise sich in einem Punkte schneiden.

<sup>5</sup> Vgl. HOWARD EVES and V. E. HOGGATT, a. a. O. <sup>2</sup>, S. 470—471. Dieser Satz wird schon bei dem Beweis des Hilbertschen Kongruenzaxioms III<sub>5</sub> (betreffs der Pseudokongruenz) verwendet.



bzw.  $Q'$  den Schnittpunkt der Geraden  $OP$  resp.  $OQ$  mit  $UV$ , so ist

$$\cos \lambda = \frac{OQ'}{OP'}$$

also auf Grund von (4\*) durch die Pseudoabstände  $\overline{OQ} = b$ ,  $\overline{OP} = c$  ausgedrückt

$$(1) \quad \cos \lambda = \frac{\operatorname{th} b}{\operatorname{th} c}.$$

Für den Pseudoabstand  $\overline{QP} = a$  besteht auf Grund der Definition unter (1)

$$e^{2a} = (UV/PQ)^2$$

d. h. wegen  $UQ = QV$

$$(5) \quad e^{2a} = \left( \frac{UP}{PV} \right)^2.$$

Wird aber der Cosinussatz auf die Dreiecke  $POU$  und  $POV$  angewandt, so erhält man mit Rücksicht auf (3)

$$\left( \frac{UP}{PV} \right)^2 = \frac{r^2 + OP^2 - 2rOP \cos(\beta + \lambda)}{r^2 + OP^2 - 2rOP \cos(\beta - \lambda)} = \frac{1 - \frac{\cos \beta}{\cos \lambda} \cos(\beta + \lambda)}{1 - \frac{\cos \beta}{\cos \lambda} \cos(\beta - \lambda)},$$

oder nach trivialer Umformung

$$\left( \frac{UP}{PV} \right)^2 = \frac{\operatorname{tg} \beta + \operatorname{tg} \lambda}{\operatorname{tg} \beta - \operatorname{tg} \lambda}.$$

(5) nimmt daher die Gestalt

$$e^{2a} = \frac{\operatorname{tg} \beta + \operatorname{tg} \lambda}{\operatorname{tg} \beta - \operatorname{tg} \lambda}$$

an. Aus dieser Gleichung ergibt sich

$$(6) \quad \operatorname{tg} \lambda = \frac{e^{2a} - 1}{e^{2a} + 1} \operatorname{tg} \beta = \operatorname{th} a \operatorname{tg} \beta.$$

Im Sinne von (4\*) ist aber

$$\cos \beta = \frac{OQ'}{r} = \operatorname{th} b,$$

woraus folgt

$$\operatorname{tg} \beta = \frac{1}{\operatorname{sh} b}.$$

<sup>6</sup> Diese Formel hat offenbar (Fig. 5) die Bedeutung

$$\left( \frac{UP}{PV} \right)^2 = \frac{UP'}{P'V}.$$



(6) geht somit in die Formel

$$(II) \quad \operatorname{tg} \lambda = \frac{\operatorname{th} a}{\operatorname{sh} b}$$

über.

Die Formeln (I) und (II) haben schon die ganze Trigonometrie des Poincaréschen Kreismodells der hyperbolischen ebenen Geometrie zur Folge.

(Eingegangen am 20. April 1953.)

ZUSATZ (während der Korrektur am 22. Mai 1954.) Früher ist erschienen unsere nachfolgende Note, Herleitung der hyperbolischen Trigonometrie in der Poincaréschen Halbebene, *Acta Scientiarum Mathematicarum* (Szeged), 15(1954), S. 126—129. Die Mitteilung von H. MESCHKOWSKI, Die Ableitung der trigonometrischen Formeln im Poincaréschen Modell der hyperbolischen Geometrie, *Elemente der Mathematik*, 7(1952), S. 130—132, zum Teil ähnlichen Inhalts, ist uns erst nachher bekannt geworden.

## О ТРИГОНОМЕТРИИ КРУГОВОЙ МОДЕЛИ ПУАНКАРЕ ДЛЯ ГИПЕРБОЛИЧЕСКОЙ ПЛОСКОЙ ГЕОМЕТРИИ

П. САС (Будапешт)

(Резюме)

Автор выводит тригонометрию известной круговой модели Пуанкаре для гиперболической плоской геометрии из следующей леммы.

Лемма. Если через точку  $P$  лежащую внутри круга  $K$  радиуса  $r$  провести окружность пересекающую  $K$  в точках  $U$  и  $V$  под прямыми углами, то, разделяя угол  $UOV \sphericalangle = 2\beta$  пополам полупрямой  $OX$ , получаем для угла  $XOP \sphericalangle = \lambda$  формулу

$$\frac{\cos \beta}{\cos \lambda} = \frac{2r OP}{r^2 + OP^2}.$$

Подобные рассуждения были проведены раньше, но менее просто, в работе Howard Eves and V. E. Hoggatt, Hyperbolic Trigonometry Derived from the Poincaré Model, *Amer. Math. Monthly*, 58 (1951), стр. 469—474.



# ÜBER DIE VERTEILUNG DER VIELFACHEN EINER KOMPLEXEN ZAHL NACH DEM MODUL DES EINHEITSQUADRATS

Von

P. SZÜSZ (Budapest)  
(Vorgelegt von A. RÉNYI)

In dieser Arbeit bezeichnet  $(v)$  stets den Bruchteil von  $v$ , d. h. :  
 $(v) = v - [v]$ .

Es seien  $\alpha$  und  $\beta$  zwei zwischen Null und Eins gelegene Zahlen, ferner

$$(1) \quad z_\nu = (\nu\alpha) + i(\nu\beta) \quad (\nu = 1, 2, \dots).$$

Es bezeichne  $J$  ein im Einheitsquadrat der  $z$ -Ebene (d. h.  $z = x + iy$ ,  $0 \leq x < 1$ ,  $0 \leq y < 1$ ) gelegenes, nicht notwendigerweise achsenparalleles Parallelogramm,  $M(J)$  dessen Flächeninhalt  $N_{z_1}(n, J)$  die Anzahl derjenigen,  $\nu \leq n$ , für die die Relation

$$z_\nu \in J$$

stattfindet.

Sind  $\alpha, \beta, 1$  linear unabhängig, so gilt bekanntlich

$$(2) \quad \lim_{n \rightarrow \infty} \frac{N_{z_1}(n, J)}{n} = M(J).$$

Wir betrachten nun die Differenz

$$(3) \quad |N_{z_1}(n, J) - nM(J)|.$$

Sind  $\alpha, \beta, 1$  linear unabhängig, so gilt wegen (2)

$$|N_{z_1}(n, J) - nM(J)| = o(n),$$

mehr läßt sich aber ohne weitere Voraussetzungen über  $z_1$  oder  $J$  nicht sagen.

In der vorliegenden Note wird die Differenz (3) untersucht im Falle, wenn für  $J$  spezielle Intervalle  $J_{z_\nu}$  gesetzt werden.

Es bezeichne  $J_z$  das Parallelogramm, definiert durch die Vektoren

$$\frac{\min(x, y)}{\max(x, y)}, \quad \min(x, y) + i \max(x, y) \quad (z = x + iy).$$

Die spezielle Lage von  $J_z$  ist dabei gleichgültig.  $J_z$  kann über das Einheitsquadrat der  $z$ -Ebene hinausragen, muß aber dann nach dem Einheitsquadrat reduziert werden. Dann gilt folgender

SATZ. Für jedes natürliche  $q$

$$(4) \quad |N_{z_1}(n, J_{z_q}) - n \min((q\alpha), (q\beta))| = O(1),$$



wobei  $z_q$  die Zahlen aus (1) bedeuten; die Schranke, die der absolute Betrag der Differenz  $N_{z_1}(n, J_{z_q}) - n \min((q\alpha), (q\beta))$  nicht übersteigt, hängt lediglich von  $\alpha, \beta$  und  $q$  ab, also weder von  $n$  noch von der speziellen Lage von  $J_{z_q}$ .

Es sei darauf hingewiesen, daß hier nicht einmal die lineare Unabhängigkeit von  $\alpha, \beta, 1$  gefordert wird.

Das hier behandelte Problem hängt mit einer von S. HARTMAN aufgeworfenen Frage zusammen. HARTMAN hat gefragt, ob die Relation (4) gilt, wenn  $J_{z_q}$  durch ein achsenparalleles Intervall ersetzt wird, dessen Seiten  $(q\alpha)$ , bzw.  $(q\beta)$  sind; (in (4) muß natürlich in diesem Falle  $\min((q\alpha), (q\beta))$  durch  $(q\alpha)(q\beta)$  ersetzt werden). Neulich ist mir gelungen, die Hartmansche Frage im verneinenden Sinne zu beantworten.<sup>1</sup> Der Beweis wird über die Kettenbruchlehre geführt und ist nicht so einfach, wie der folgende Beweis von (4).

Bevor ich zum Beweis von (4) übergehe, möchte ich eine Anwendung geben.

Für linear unabhängige  $\alpha, \beta, 1$  leite ich aus dem bekannten Kronecker'schen Satze und aus (4) einen sehr einfachen Beweis für die bekannte Limesrelation (2) her.

Es sei  $J$  ein beliebiges Teilintervall von  $0 \leq x < 1$ ,  $0 \leq y < 1$  und nehmen wir an, (4) sei bereits bewiesen. Wegen der linearen Unabhängigkeit von  $\alpha, \beta, 1$  ist es möglich, ein  $q$  mit

$$(5) \quad \begin{cases} (q\alpha) < \varepsilon^2 \\ \varepsilon - \varepsilon^2 < (q\beta) < \varepsilon \end{cases} \quad (\varepsilon > 0, \text{ beliebig klein})$$

zu finden. Nimmt man  $J_{z_q}$  mit diesem  $q$  ( $J_{z_q}$  hat dieselbe Bedeutung, wie bei (4)), so unterscheidet sich der minimale Gesamtflächeninhalt derjenigen  $J_{z_q}$ , die zu einer vollständigen Überdeckung von  $J$  nötig sind, von dem maximalen Gesamtflächeninhalt derjenigen  $J_{z_q}$ , die von  $J$  überdeckt werden, um weniger, als  $\varepsilon'$  ( $\varepsilon' > 0$ , beliebig klein) falls nur  $\varepsilon$  genügend klein ist. Es bezeichne  $\bar{J}$  die Gesamtheit der überdeckenden Intervalle,  $\underline{J}$  die Gesamtheit derjenigen  $J_{z_q}$ , die von  $J$  überdeckt werden. Nach dem Gesagten ist also

$$(6) \quad M(\bar{J}) - M(\underline{J}) < \varepsilon',$$

falls nur  $\varepsilon$  (aus (5)) genügend klein gewählt ist.

Die Intervallfunktion  $N(n, J)$  ist offenbar additiv, es gilt also

$$(7) \quad N_{z_1}(n, \underline{J}) < N_{z_1}(n, J) < N_{z_1}(n, \bar{J}),$$

Es gilt aber wegen (4)

$$N_{z_1}(n, \underline{J}) = M(\underline{J})n + O(1), \quad N_{z_1}(n, \bar{J}) = M(\bar{J})n + O(1),$$

also wegen (7) und der Additivität von  $N(n, J)$

$$M(\underline{J})n + O(1) < N_{z_1}(n, J) < M(\bar{J})n + O(1).$$

<sup>1</sup> Erscheint in *Studia Mathematica*.



Wegen (6) folgt nun

$$N_{z_1}(n, J) = M(J)n + n\mathcal{G}\varepsilon' + O(1)$$

mit irgendeinem  $\mathcal{G}$  ( $|\mathcal{G}| \leq 1$ ). Da  $\varepsilon'$  beliebig klein gewählt werden darf, ist (2) bewiesen.

Nun wende ich mich dem Beweis von (4) zu.

1) Es sei zuerst  $q = 1$ ; ohne Beschränkung der Allgemeinheit sei  $\alpha \leq \beta$  angenommen.  $N_{z_1}(n, J_{z_1})$  kann folgendermaßen charakterisiert werden: man stelle jede Zahl  $z (= x + iy)$  in der Form  $u + vz_1 = u + v(\alpha + i\beta)$  mit reellem  $u$  und  $v$  dar. Eine solche Darstellung ist eindeutig möglich; durch eine einfache Rechnung bestätigt man

$$(8) \quad u = x + y \frac{\alpha}{\beta}, \quad v = \frac{y}{\beta}.$$

Man setze  $x = r\alpha$ ,  $y = (r\beta)$  ( $r = 1, 2, \dots$ ), d. h. wir führen die Reduktion noch nicht nach dem Einheitsquadrat durch, sondern reduzieren nur die Komponente  $y$  modulo 1; die Punkte  $r\alpha + i(r\beta)$  ( $r = 1, 2, \dots$ ) liegen dann im Streifen  $0 \leq x < \infty$ ,  $0 \leq y < 1$  verteilt. Aus (8) folgt

$$(8') \quad u_r = r\alpha - (r\beta) \frac{\alpha}{\beta} = r\alpha - \frac{\alpha}{\beta} (r\beta - [r\beta]) = \frac{\alpha}{\beta} [r\beta].$$

$N_{z_1}(n, J_{z_1})$  ist offenbar gleich der Anzahl derjenigen  $r \leq n$ , für die Relation

$$(9) \quad (u_r) \in J_1, \quad (r\beta) \in J_2$$

gleichzeitig stattfinden, wobei  $J_1$  ein eindimensionales Intervall von der Länge  $\frac{\alpha}{\beta}$ ,  $J_2$  ein ebenfalls eindimensionales Intervall von der Länge  $\beta$  bedeutet. Aus (8') und aus der Voraussetzung  $0 < \beta < 1$  folgt, daß für  $u_r$  nur die Werte  $k \frac{\alpha}{\beta}$  ( $k = 1, 2, \dots$ ) in Frage kommen und jedes  $k$  mindestens einmal angenommen wird. Es sei  $k$  gegeben. Ich betrachte diejenigen  $r$ , für die

$$[r\beta] = k$$

gilt, also die Zahlen  $\left[ \frac{k}{\beta} + 1 \right], \left[ \frac{k}{\beta} + 2 \right], \dots, \left[ \frac{k+1}{\beta} \right]$ . Für genau eine dieser Zahlen, die für einen Augenblick mit  $r'$  bezeichnet werden soll, gilt

$$(r'\beta) \in J_2,$$

denn je zwei konsekutive der Zahlen  $(r\beta)$  ( $r = \left[ \frac{k}{\beta} + 1 \right], \left[ \frac{k}{\beta} + 2 \right], \dots, \left[ \frac{k+1}{\beta} \right]$ ) haben voneinander den gleichen Abstand  $\beta$ ; da

$$\beta \left( \left[ \frac{k+1}{\beta} \right] + 1 - \left[ \frac{k}{\beta} \right] \right) > \beta \left( \frac{k+1}{\beta} - \left[ \frac{k}{\beta} \right] \right) = 1 + \beta \left( \frac{k}{\beta} \right) > 1,$$

liegt in jedem Intervall der Länge  $\beta$  genau eine der Zahlen  $\beta \left[ \frac{k}{\beta} + 1 \right]$ ,



$\beta \left[ \frac{k}{\beta} + 2 \right], \dots, \beta \left[ \frac{k+1}{\beta} \right]$ .  $N_{z_1}(n, J_{z_1})$  ist daher nicht größer, als die Anzahl der  $k \leq n\beta$  mit  $\left( k \frac{\alpha}{\beta} \right) \in J_1$  und nicht kleiner als die Anzahl der  $k \leq n\beta - 1$  mit

$$\left( k \frac{\alpha}{\beta} \right) \in J_1.$$

Da  $J_1$  die Länge  $\frac{\alpha}{\beta}$  hat, liegt zwischen  $\nu$  und  $\nu+1$  genau ein  $k'$ , für welches

$$\left( k' \frac{\alpha}{\beta} \right) \in J_1$$

gilt, es gilt also

$$N_{z_1}(n, J_{z_1}) = [n\beta] \frac{\alpha}{\beta} + O(1) = n\alpha + O(1),$$

womit unser Satz für  $q=1$  bewiesen ist.

2) Es sei nun  $q \geq 2$ . Dieser Fall läßt sich einfach auf den unter 1) behandelten Fall zurückführen. Es ist

$$(10) \quad N_{z_1}(n, J_{z_q}) = \sum_{\substack{r \equiv n \\ \nu z_1 \in z_1}} 1 = \sum_{s=0}^{q-1} \sum_{\substack{r \equiv 0 \\ \nu z_1 \in J_{z_q} \\ r \equiv s \pmod{q}}} 1,$$

wobei  $\nu z_1 \in J_{z_q}$  so zu verstehen ist, daß  $\nu z_1$  nach dem Einheitsquadrat reduziert zu  $J_{z_q}$  gehört. Aus (10) folgt

$$(11) \quad \begin{aligned} N_{z_1}(n, J_{z_q}) &= \sum_{s=0}^{q-1} \sum_{\substack{r \equiv 1 \\ \nu z_1 \in J_{z_q} \\ r \equiv s \pmod{q}}} 1 = \sum_{s=0}^{q-1} \sum_{\substack{r \equiv \frac{n-s}{q} \\ (rq+s)z_1 \in J_{z_q}}} 1 = \sum_{s=0}^{q-1} N_{z_q} \left( \frac{n-s}{q}, J_{z_q}^{(s)} \right) = \\ &= \sum_{s=1}^{q-1} N_{z_q} \left( \frac{n}{q}, J_{z_q}^{(s)} \right) + O(q), \end{aligned}$$

wobei  $J_{z_q}^{(s)}$  das um  $sz_1$  „verschobene“ und nachher nach dem Einheitsquadrat reduzierte Intervall  $J_{z_q}$  bedeutet, also die Punktmenge  $z + sz_q$  mit  $z \in J_{z_q}$ , die noch nach dem Einheitsquadrat reduziert werden soll. Wendet man auf jedes Glied der letzten summe in (11) das Ergebnis von 1) mit  $z_q$  statt des dortigen  $z_1$ ,  $\frac{n}{q}$  statt des dortigen  $n$  und  $J_{z_q}^{(s)}$  statt des dortigen  $J_{z_1}$  an, so kommt

$$N_{z_1}(n, J_{z_q}) = n \min((q\alpha), (q\beta)) + O(q) = n \min(q\alpha, q\beta) + O(1),$$

(da  $q$  fest gegeben), womit wir unseren Beweis zu Ende geführt haben.

(Eingegangen am 15. Juli 1953.)



О РАСПРЕДЕЛЕНИИ КРАТНЫХ КОМПЛЕКСНОГО ЧИСЛА ПО  
МОДУЛЮ КВАДРАТА ЕДИНИЦЫ

П. СЮС (Будапешт)

(Резюме)

В следующем ( $u$ ) означает дробную часть  $u$ . Пусть  $z_1 = \alpha + i\beta$ , где  $\alpha$  и  $\beta$  числа, лежащие между 0 и 1.  $J_q$  есть (произвольно расположенный) параллелограмм, стороны которого характеризуются векторами

$$\frac{\min [(q\alpha), (q\beta)]}{\max [(q\alpha), (q\beta)]} \text{ и } \min [(q\alpha), (q\beta)] + i \max [(q\alpha), (q\beta)].$$

$N_{z_1}(n, J_q)$  означает число значений  $v \leq n$ , для которых

$$(v\alpha) + i(v\beta) = z_v \in J_q.$$

При таких предположениях имеет место следующая теорема:

$$|N_{z_1}(n, J_q) - n \min [(q\alpha), (q\beta)]| = O(1),$$

если  $n \rightarrow \infty$ . Граница, не превосхожденная этой разницей, зависит только от  $\alpha, \beta$  и  $q$ , т. е. не зависит и от положения  $J_q$ .

В качестве применения отсюда получается очень простое доказательство известного факта, что точки  $z_v$  ( $v = 1, 2, \dots$ ) равномерно распределены в единичном квадрате плоскости  $z$ .







# ÜBER DIE DICHTESTE HOROZYKLENLAGERUNG

Von

L. FEJES TÓTH (Budapest)

(Vorgelegt von G. Hajós)

Wir betrachten in einer hyperbolischen Ebene drei einander gegenseitig berührende Kreise vom Radius  $r$ . Der Inhalt des durch die Kreismittelpunkte bestimmten Dreiecks sei  $A$ , die Inhaltssumme der im Dreieck liegenden Kreis-sektoren  $T$ . Wir betrachten die Dichte der Kreise im Dreieck, d. h. den Quotienten  $D = T/A$ . Es läßt sich leicht zeigen, daß  $D = D(r)$  für  $r \rightarrow \infty$  ständig zunehmend dem Grenzwert  $\lim_{r \rightarrow \infty} D(r) = 3/\pi$  zustrebt, so daß  $D < 3/\pi$  ausfällt.

Es sei nun in der hyperbolischen Ebene ein beliebiges System von nicht übereinandergreifenden Kreisen vom Radius  $r$  vorgegeben. In einer vorigen Arbeit<sup>1</sup> habe ich gezeigt, daß sich dann die Ebene so in Dreiecke zerlegen läßt, daß die Dichte der Kreise in jedem Dreieck  $\leq D(r)$  ausfällt. Wird daher die Lagerungsdichte  $d$  in der ganzen Ebene als irgendein Mittelwert der Dichten bezüglich dieser Dreiecke definiert, so ist  $d < 3/\pi$ . In der genannten Arbeit habe ich ohne Beweis darauf hingewiesen, daß diese Ungleichung ihre Gültigkeit mit Zulassung des Gleichheitszeichens auch im Falle von Grenzkreisen behält. Jedoch bedarf diese Behauptung eines Beweises, den wir hier durch eine direkte Behandlung einer Horozyklenpackung gewähren wollen. Unser Resultat lautet folgenderweise:

*Zu einer beliebig vorgegebenen Lagerung von nicht übereinandergreifenden Horozyklen läßt sich stets eine Zerlegung der hyperbolischen Ebene in asymptotische Dreiecke so angeben, daß die Dichte der Horozyklen in jedem Dreieck  $\leq 3/\pi$  ausfällt.*

Wir können voraussetzen, daß unser Horozyklensystem  $\{H\}$  gesättigt ist, d. h. daß jeder Horozyklus der Ebene gemeinsame innere Punkte mit einem Horozyklus von  $\{H\}$  aufweist. Wir betrachten einen Kreis  $K$ , der wenigstens drei Horozyklen von  $\{H\}$  berührt, ohne gemeinsame innere Punkte mit irgendeinem Horozyklus des Systems zu besitzen. Die „Mittelpunkte“ der

<sup>1</sup> Kreisausfüllungen der hyperbolischen Ebene, *Acta Math. Acad. Sci. Hungaricae*, 4 (1953), S. 103—110.



durch  $K$  berührten Horozyklen sind Ecken eines asymptotischen Polygons  $P$ . Wir zeigen, daß die Gesamtheit  $\{P\}$  der so konstruierten Polygone die Ebene schlicht und lückenlos bedeckt.

Wir zeigen zunächst, daß zwei Polygone  $P$  und  $P'$  von  $\{P\}$  nicht übereinandergreifen. Aus der Konstruktion der Polygone  $\{P\}$  folgt nämlich, daß zwei Polygone von  $\{P\}$  nicht mehr als zwei gemeinsame „Eckpunkte“ besitzen können. Haben daher  $P$  und  $P'$  gemeinsame innere Punkte, so gibt es je eine Seite  $AB$  und  $A'B'$  von  $P$  bzw.  $P'$ , die einander schneiden. Bezeichnen wir die Horozyklen von  $\{H\}$  von den Mittelpunkten  $A, B, A', B'$  mit  $H_A, H_B, H_{A'}, H_{B'}$ , so gibt es nach Definition der Polygone  $\{P\}$  einen Kreis  $K$ , der  $H_A$  und  $H_B$  berührt, ohne in  $H_{A'}$  oder  $H_{B'}$  hineinzugreifen. Ebenso gibt es einen von  $K$  verschiedenen Kreis  $K'$ , der  $H_{A'}$  und  $H_{B'}$  berührt, aber weder in  $H_A$  noch in  $H_B$  hineindringt. Ein Blick auf das Poincarésche Kreismodell (Abb. 1) überzeugt uns davon, daß diese Annahme sich als unmöglich erweist, indem  $K$  und  $K'$  mehr als zwei gemeinsame Punkte aufwiesen.

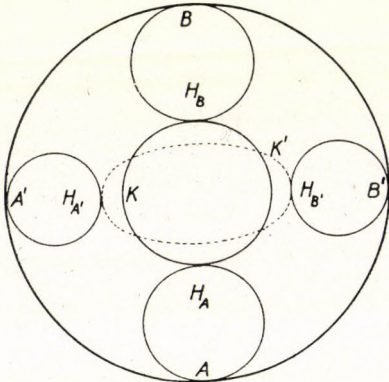


Abb. 1.

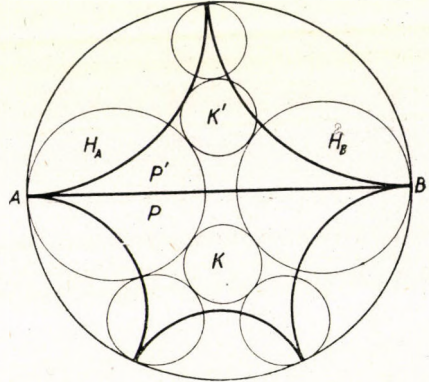


Abb. 2.

Wir haben noch zu zeigen, daß die Polygone  $\{P\}$  die Ebene völlig bedecken. Wegen der Gesättigtheit von  $\{H\}$  gibt es unter den  $H_A$  und  $H_B$  berührenden Kreisen, außer dem  $P$  definierenden Kreis  $K$ , einen zweiten Kreis  $K'$ , der außer  $H_A$  und  $H_B$  noch einen Horozyklus von  $\{H\}$  berührt, ohne in irgendeinen Horozyklus von  $\{H\}$  hineinzugreifen (Abb. 2). Das bedeutet, daß sich zu  $AB$ , also zu jeder Seite eines Polygons von  $\{P\}$ , ein anderes Polygon  $P'$  von  $\{P\}$  anschließt. Wir betrachten ein Polygon  $P$ , sowie die zu den Seiten von  $P$  anschließenden Polygone. Dann fügen wir zu den freien Seiten wieder die anschließenden Polygone hinzu und setzen dieses Verfahren unbegrenzt fort. Wir setzen voraus, daß ein Punkt  $Q$  der Ebene nach beliebig vielen Schritten dieses Prozesses unbedeckt bleibt. Dann muß diejenige, im  $n$ -ten Schritt auftretende Polygonseite  $A_n B_n$ , die  $Q$  von den schon hingefügten Polygonen trennt, mit  $n \rightarrow \infty$  gegen eine Grenzlage  $AB$  heranrücken. Es sei  $K_n$  ein Kreis, der die Horozyklen  $H_{A_n}$  und  $H_{B_n}$  von  $\{H\}$



von den Mittelpunkten  $A_n$  und  $B_n$  berührt, aber die übrigen Horozyklen nicht erreicht. Die Existenz eines solchen Kreises folgt aus der Definition der Polygone  $\{P\}$ . Sind die Punkte  $A_n$  und  $B_n$  von keinem Index an mit  $A$  bzw.  $B$  identisch, so läßt es sich leicht zeigen, daß die Horozyklen  $H_{A_n}$  und  $H_{B_n}$  im Kreismodell gegen die Punkte  $A$  bzw.  $B$  konvergieren. Dann strebt aber  $K_n$  gegen den Begrenzungskreis des Modells, woraus folgen würde, daß im Modell überhaupt kein Horozyklus vorhanden ist. Wir müssen also annehmen, daß etwa die Horozyklen  $H_{B_n}$  von einem Index  $N$  an identisch sind:  $H_{B_n} \equiv H_B$  ( $n > N$ ). Dann strebt  $K_n$  gegen den  $H_B$  berührenden Horozyklus  $H_A$  vom Mittelpunkt  $A$  (Abb. 3). Da aber  $H_A$  mit keinem Horozyklus von  $\{H\}$  gemeinsame innere Punkte aufweist, wäre  $\{H\}$  nicht gesättigt.

Zerlegen wir die mehr als dreieitigen Polygone von  $\{P\}$  irgendwie durch Diagonale in Dreiecke, so entsteht ein Netz asymptotischer Dreiecke, von dem wir noch zu zeigen haben, daß es der Forderung unseres Satzes Genüge leistet.

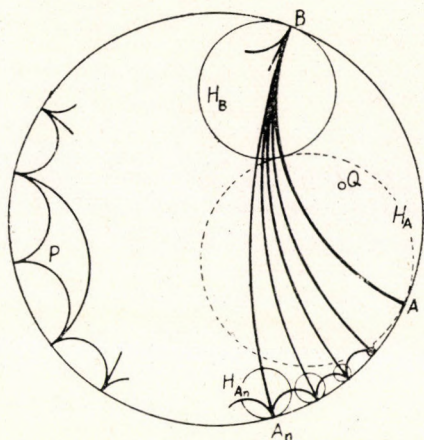


Abb. 3.

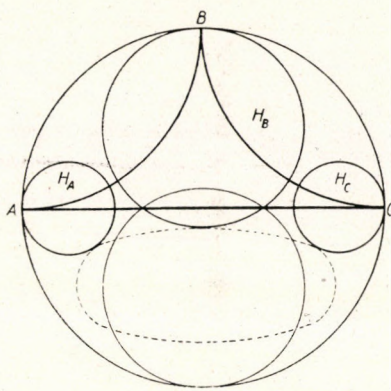


Abb. 4.

Es sei  $ABC$  ein Dreieck unseres Netzes und  $H_A, H_B, H_C$  die entsprechenden Horozyklen. Wir behaupten, daß kein Horozyklus die gegenüberliegende Dreiecksseite treffen kann. Erreicht nämlich etwa  $H_B$  die Seite  $AC$ , so gibt es keinen Kreis, der  $H_A, H_B$  und  $H_C$  berührt, weil dieser mit dem Spiegelbild von  $H_B$  an  $AC$  mehr als zwei Punkte gemein hätte (Abb. 4). Ragt aber von einem beliebigen Dreieck unseres Netzes kein Horozyklenabschnitt heraus, so kann umgekehrt in  $ABC$  kein fremder (d. h. von  $H_A, H_B, H_C$  verschiedener) Horozyklus von  $\{H\}$  hineingreifen. Folglich kommen bei der Abschätzung der Dichte  $d$  in  $ABC$  nur die Horozyklen  $H_A, H_B$  und  $H_C$  in Betracht.

Offenbar erreicht  $d$  ihr Maximum, wenn ein Horozyklus, sagen wir  $H_B$ , von den beiden anderen berührt wird. Wir zerlegen  $ABC$  in zwei inhalts-



gleiche Dreiecke  $AMB$  und  $BMC$  und bezeichnen die in  $AMB$  liegenden Sektoren von  $H_A$  und  $H_B$  samt ihren Inhalten mit  $t_A$  bzw.  $t_B$  (Abb. 5). Es handelt sich um das Maximum von  $t_A + t_B$  unter der Voraussetzung eines festen rechtwinkligen Dreiecks  $AMB$  und einander berührender Horozyklen  $H_A$  und  $H_B$ , von denen keiner  $M$  als inneren Punkt enthält. Dabei kann  $M$  am Rand von  $H_A$ , aber nicht am Rand von  $H_B$  liegen. Liegt  $M$  außerhalb  $H_A$ ,

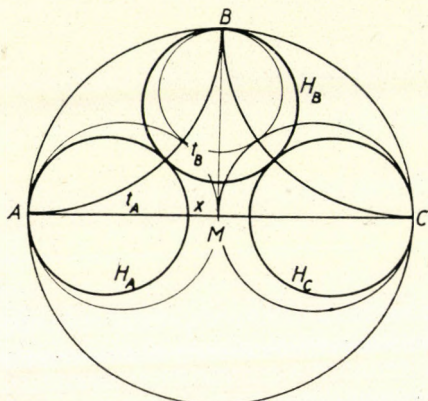


Abb. 5.

so ersetzen wir  $H_A$  durch den durch  $M$  hindurchgehenden Horozyklus und  $H_B$  durch den diesen berührenden Horozyklus. Dann kommt zu  $t_A$  ein Horozyklenringsektor  $h_A$  hinzu, während von  $t_B$  ein Horozyklenringsektor  $h_B$  derselben „Dicke“ wegfällt. Da aber der äußere Horozyklenbogen bei  $h_A$  größer ist als bei  $h_B$ , enthält  $h_A$  einen zu  $h_B$  kongruenten Sektor. Folglich nimmt  $t_A + t_B$  zu. Damit ist gezeigt, daß die maximale Dichte  $\bar{d}$  im Falle von drei einander gegenseitig berührenden Horozyklen  $H_A, H_B, H_C$  erreicht wird.

Zur Berechnung von  $\bar{d}$  fassen wir eine Ebene der Krümmung  $-1$  ins Auge. Dann ist der Inhalt von  $ABC$  gleich  $\pi$  und die Bogenlänge des in  $ABC$  liegenden Teilbogens von  $H_A$  gleich  $e^{-x}$ , wo  $x$  den Abstand des Punktes  $M$  von  $H_A$  bedeutet. Folglich haben wir, unserer Behauptung entsprechend,

$$\bar{d} = \frac{3}{\pi} \int_0^{\infty} e^{-x} dx = \frac{3}{\pi}.$$

(Eingegangen am 12. Oktober 1953.)

## О САМОМ ПЛОТНОМ РАСПОЛОЖЕНИИ ГОРОЦИКЛОВ

Л. ФЕЕШ ТОТ (Веспрем)

(Резюме)

Автором доказывается следующий результат: К любой системе непересекающихся гороциклов можно задать разложение гиперболической плоскости на асимптотические треугольники, такое, что плотность гороциклов в каждом из треугольников  $\leq \frac{3}{\pi}$ .



# THE THEORY OF GROUPS WITH FINITE CLASSES OF CONJUGATE ELEMENTS

By

J. ERDŐS (Debrecen)

(Presented by L. RÉDEI)

## § 1. Introduction

In recent investigations concerning the structure of infinite groups, several important results were obtained for groups which are in one sense or another close to abelian groups resp. to finite groups. Such a category of groups is, for instance, the groups in which every element has a finite number of conjugates. In what follows such groups will be called briefly FC-groups. Trivial examples of FC-groups are the abelian groups, the finite groups, and any direct product of an arbitrary set of abelian or finite groups. But we shall see in § 8 below that there exist FC-groups, even ones with two generators, which cannot be obtained by such a trivial construction.

The theory of FC-groups is due to R. BAER and B. H. NEUMANN and is developed in [2] and [6].<sup>1</sup> The present paper has arisen from the observation that NEUMANN'S very clear and excellent treatment in [6] admits some simplifications. In particular, in the present exposition of NEUMANN'S theory I have succeeded in avoiding the use of a deep result of SCHREIER in [7] and the extension of the group under consideration by constructing a direct product with amalgamation (this plays an important role in NEUMANN'S treatment). Thus the development of the theory of FC-groups becomes thoroughly elementary, so that the reading of this paper supposes merely the knowledge of the simplest basic concepts of group theory. In addition to NEUMANN'S theory, we give here also a new application and we make some remarks on the converses of the results.

The main results of the theory of FC-groups are the following. In an FC-group the elements of finite order form a (characteristic) subgroup whose factor group in the whole group is an abelian torsion-free group (§ 4). A group is always an FC-group if its center has a finite index, or if its derived group is finite (§ 2, § 6). However, we shall see in § 8 that none of

<sup>1</sup> The numbers in brackets refer to the Bibliography at the end of this paper.



these two sufficient conditions is necessary, unless the group under consideration is finitely generated. On the other hand, if a group is an FC-group then its derived group is a torsion group (§ 3); NEUMANN's counterexample — the group of finite permutations of an infinite set — shows that the converse of this statement is not true in general. As an important application we get, in the same way as NEUMANN, the theorem of BAER and NEUMANN [1], [3], [6] stating that the finiteness of the index of the center of a group  $G$  implies the finiteness of the derived group  $G'$  (§ 6). Another interesting application, also due to NEUMANN, is the following theorem of DICMAN and NEUMANN [4], [6]: an element of an arbitrary group  $G$  belongs to a finite normal subgroup of  $G$  if and only if it has finite order and a finite number of conjugates (§ 5). Recently FJODOROV has succeeded in proving the following very nice theorem: if any subgroup with more than one element of an infinite group  $G$  has finite index, then  $G$  is cyclic. Now we show in § 7 that the theory of FC-groups leads to a quite simple and short proof of this theorem.

Finally, we outline briefly the way, in which the theory of the FC-groups will be developed in the present paper (§§ 2—4). In our treatment a central place is taken by the theorem, according to which the derived group of an FC-group is a torsion group (§ 3). We obtain this main theorem in the following way: first we show that it is sufficient to prove the theorem for finitely generated groups. Then we show that the center of a finitely generated FC-group  $G$  always contains a torsion-free subgroup  $A$  which has a finite index  $m$  in  $G$ . Thus the  $m$ th power of any element of  $G$  is contained in  $A$  (§ 2). In § 3 (cf. Lemma 3.3) we prove that the  $m$ th power of a product of elements of  $G$  can be obtained by raising to the  $m$ th power each of the factors. This is the essentially new feature in our treatment of NEUMANN's theory, and this implies that the  $m$ th power of any element of the derived group of  $G$  is equal to the unit element, any such  $m$ th power being contained in the center of  $G$ .

## § 2. Preliminaries

In the present section we briefly summarize the fundamental facts on which the theory of FC-groups will be based in §§ 3—4. Well-known simple lemmas will be stated without proof.

The notations employed are as follows: groups are denoted by Latin capitals, elements of groups by the small letters  $a, b, \dots, g$ . The other small Latin letters will denote rational integers. A group is called *finitely generated*, if there exists a finite system of group elements which are not all contained in a proper subgroup of the group. A group is called *torsion-free*, if it contains no element of finite order other than the unity. On the other hand, groups every element of which is of finite order are called *torsion groups*.



The groups which are neither torsion groups nor torsion-free groups are said to be *mixed groups*. The elements of a group which commute with a given group element, form a subgroup called the *normalizer* of the given element. The *center* of a group consists of the elements which commute with any group element. If  $a$  and  $b$  are arbitrary elements of the group  $G$ , then the element

$$[a, b] = a^{-1}b^{-1}ab$$

is called the *commutator* of the elements  $a$  and  $b$ . The *derived group* of  $G$ , denoted by  $G'$ , is the subgroup of  $G$  generated by all the commutators of  $G$ . The center and the derived group of a group are always normal subgroups. The unit element of any group occurring will be denoted by  $1$ , since this gives rise to no confusion.

The following lemma is a statement well-known from the elements.

LEMMA 2.1. *The cardinal number of the set of conjugates of a group element is equal to the index of its normalizer.*

Owing to this, the FC-groups can be characterized by the property that the normalizer of any of their elements has a finite index in the group. As the center of a group is the intersection of the normalizers of all group elements, we have obviously the following theorem, giving a sufficient condition for a group to be an FC-group:

LEMMA 2.2. *If the center of a group has a finite index in the group, then the group is an FC-group.*

This sufficient condition is, however, not necessary in general, as we shall see in § 8. The converse of Lemma 2.2 is valid only for finitely generated groups:

LEMMA 2.3. *The center of a finitely generated FC-group has a finite index.*

PROOF. Let  $G$  be a finitely generated FC-group, and let us consider a finite system of generators of  $G$ . It is clear that the center of  $G$  is the intersection of the normalizers of the generator elements. Since  $G$  is an FC-group, these normalizers are subgroups of finite index in  $G$ . On the other hand, the intersection of a finite number of subgroups with finite index is also a subgroup of finite index, as the index of the intersection cannot exceed the product of the indices of the subgroups. Thus Lemma 2.3 is proved.

LEMMA 2.4. *Any subgroup of finite index of a finitely generated FC-group is itself finitely generated.*

PROOF.<sup>2</sup> Let  $G$  be a finitely generated FC-group, and let  $H$  be a subgroup of finite index  $r$  in  $G$ . From a fixed finite system of generators of  $G$ , by the

<sup>2</sup> I am indebted to T. SZELE for this simple proof.



FC-property, we may obtain a finite system of generators,

$$(1) \quad g_1, g_2, \dots, g_n,$$

containing, together with any element, its inverse and all of its conjugates too. Let us now form all the products

$$(2) \quad a_1 a_2 \dots a_k \quad (a_i = g_j; k \leq r)$$

with at most  $r$  factors  $a_i$  belonging to the system (1). (The same element may occur among the factors several times, but the order of the factors is essential.) We denote the set of all these products (2) by  $K$ .  $K$  is evidently a finite set. Now we prove Lemma 2.4 by showing that the elements of  $K$  belonging to  $H$  form a system of generators of  $H$ .

For this purpose let us consider an arbitrary element  $b$  of the subgroup  $H$ . Since  $b \in G$ ,  $b$  can be written in the form

$$(3) \quad b = b_1 b_2 \dots b_s \quad (b \in H)$$

where each factor  $b_i$  belongs to the system (1). The existence of a decomposition of  $b$  into a product of elements of the set  $K$  can be proved by induction on  $s$ . Evidently, we can restrict ourselves to the case  $s > r$ , since for  $s \leq r$  the element  $b$  of  $H$  in (3) is itself an element of  $K$ . Now, by  $s > r$ , the system of the elements

$$b_1, b_1 b_2, \dots, b_1 b_2 \dots b_s$$

contains two elements, say,  $b_1 b_2 \dots b_u$  and  $b_1 b_2 \dots b_{u+v}$ , belonging to the same right coset modulo  $H$ . Thus

$$(4) \quad c = b_{u+1} \dots b_{u+v} \in H.$$

Now  $b$  in (3) can be written also in the form

$$(5) \quad b = b_1 \dots b_u \cdot c \cdot b_{u+r+1} \dots b_s = c[(c^{-1} b_1 c) \dots (c^{-1} b_u c) \cdot b_{u+r+1} \dots b_s].$$

Since  $b, c \in H$ , the product in brackets also belongs to  $H$ . On the other hand, this product contains less than  $s$  factors from the system (1), since not only the  $b_i$ 's, but also the elements  $c^{-1} b_i c$  belong to the system (1). Then, by our induction hypothesis, the product in brackets can be obtained as a product of some elements of  $K$ . The same is true for the element  $c$  in (4), having fewer factors than  $s$ , and therefore by virtue of (5), also for the element  $b$  under consideration. This completes the proof of Lemma 2.4.

Applying Lemma 2.4 to the center of a finitely generated FC-group, we obtain the following

LEMMA 2.5. *An arbitrary finitely generated FC-group  $G$  has a torsion-free subgroup  $A$  of finite index, contained in the center of  $G$ .*

PROOF. Let  $G$  be a finitely generated FC-group. By Lemma 2.3, the center  $Z$  of  $G$  has a finite index  $r$  in  $G$ . By virtue of Lemma 2.4,  $Z$  is then a finitely generated abelian group admitting, by the fundamental theorem of finitely generated abelian groups, a decomposition into the direct product of



a finite group and a torsion-free group  $A$ . If this finite group has order  $s$ , then the torsion-free group  $A$  has the finite index  $rs$  in  $G$ . Thus Lemma 2.5 is proved.

LEMMA 2.6. *The derived group of an arbitrary group  $G$  is a torsion group if and only if the derived group of any finitely generated subgroup of  $G$  has this property.*

PROOF. It is evidently sufficient to prove that if the derived group of any finitely generated subgroup of  $G$  is a torsion group, then  $G'$  is a torsion group too. This is clear, considering that an arbitrary element of  $G'$  is the product of a finite number of commutators, thus being contained in the derived group of a finitely generated subgroup of  $G$ .

Owing to the lemma just proved, we shall have to prove the main theorem in § 3 instead for arbitrary FC-groups only for finitely generated FC-groups.

LEMMA 2.7. *If the derived group of a group  $G$  is a torsion group, then the elements of finite order of  $G$  form a group.*

PROOF. Let the derived group  $G'$  of  $G$  be a torsion group, and let us consider two arbitrary elements of finite order, say  $a$  and  $b$ , of  $G$ . We have only to show that the product  $ab$  is of finite order too, since the inverse of an element of finite order is evidently an element of finite order. We denote by  $m$  the product of orders of the elements  $a$  and  $b$ . Thus

$$(ab)^m = a^m b^m c = c,$$

where  $c$  is an element of the derived group  $G'$ , for each interchange of two contiguous elements of the product  $(ab)^m$  induces a commutator-element, which may be written after the elements interchanged. On the other hand,  $G'$  is by hypothesis a torsion group, i. e. there exists a natural number  $n$  such that  $c^n = 1$ . Therefore

$$(ab)^{mn} = c^n = 1,$$

completing the proof of Lemma 2.7.

The validity of the next lemma is obvious:

LEMMA 2.8. *If in a group  $G$  the elements of finite order form a subgroup  $H$ , then  $H$  is characteristic in  $G$ , and the factor group  $G/H$  is torsion-free.*

It is also easy to prove the following lemma.

LEMMA 2.9. *A group is an FC-group if and only if it has a system of generators each element of which has a finite number of conjugates in the group.*

PROOF. Since the set of all elements of an FC-group is a system of generators having the property formulated in the lemma, it suffices to show that the existence of a system of generators with this property implies that



the group is an FC-group. Clearly, this follows from the fact that an arbitrary group element can be written as a product of a finite number of generator elements and of their inverses. Considering that the conjugate of a product is equal to the product of the corresponding conjugates of the factors, and by our hypothesis the factors have only a finite number of conjugates, it is therefore evident that each element has only a finite number of conjugates in the group, i. e. the group is an FC-group.

### § 3. The main theorem

Now we are going to prove that the derived group of an FC-group is always a torsion group. The proof is based on the lemmas of the preceding section.

First we consider the case in which the group is finitely generated. Once proved this special case, the theorem can easily be established for arbitrary FC-groups. Let us denote by  $G$  an arbitrary finitely generated FC-group throughout this section. By Lemma 2.5 we know that  $G$  has a torsion free subgroup  $A$  of finite index  $m$ , contained in the center of  $G$ . Our immediate aim is to show that the  $m$ th power of any product of elements of  $G$  can be obtained by raising to the  $m$ th power each of the factors separately. This we establish by the first three lemmas below.

Since  $A$  is contained in the center of  $G$ , it is a normal subgroup of  $G$ . We denote by small Greek letters the elements of the factor group  $G/A$  (the cosets of  $G$  modulo  $A$ ) and by  $g_\varrho$  an element of the coset  $\varrho$ . Let  $g_\varrho, g_\sigma \in G$ . As the  $m$ th power of an arbitrary element of  $G$  belongs to  $A$ , the element  $a_{\varrho, \sigma}$  defined by the equation

$$(6) \quad (g_\varrho g_\sigma)^m = a_{\varrho, \sigma} g_\varrho^m g_\sigma^m$$

also belongs to  $A$ . We show that this element  $a_{\varrho, \sigma}$  depends only on the cosets  $\varrho$  and  $\sigma$ , but not on the representatives  $g_\varrho$  and  $g_\sigma$  of these cosets. For if, instead of  $g_\varrho$  and  $g_\sigma$  we choose  $g'_\varrho = a_1 g_\varrho$  and  $g'_\sigma = a_2 g_\sigma$  as representatives of these cosets ( $a_1$  and  $a_2$  denoting suitable elements of  $A$ , i. e. elements of the center of  $G$ ), we find that for the element  $a'_{\varrho, \sigma}$ , defined by the equation

$$(g'_\varrho g'_\sigma)^m = a'_{\varrho, \sigma} g'^m_\varrho g'^m_\sigma,$$

we have

$$\begin{aligned} a'_{\varrho, \sigma} &= (g'_\varrho g'_\sigma)^m g'^{-m}_\sigma g'^{-m}_\varrho = (a_1 g_\varrho \cdot a_2 g_\sigma)^m (a_2 g_\sigma)^{-m} (a_1 g_\varrho)^{-m} = \\ &= a_1^m a_2^m (g_\varrho g_\sigma)^m a_2^{-m} g_\sigma^{-m} a_1^{-m} g_\varrho^{-m} = \\ &= a_1^m a_2^m a_2^{-m} a_1^{-m} (g_\varrho g_\sigma)^m g_\sigma^{-m} g_\varrho^{-m} = a_{\varrho, \sigma}, \end{aligned}$$

i. e.  $a_{\varrho, \sigma}$  depends only on  $\varrho$  and  $\sigma$ , in fact.

The following three lemmas refer to the (evidently finite) "factor system" consisting of the elements  $a_{\varrho, \sigma}$ . Our aim is to prove that all elements  $a_{\varrho, \sigma}$  co-



incide with the unit element of  $G$ . In the proof of the following lemmas an important role is played by the fact that *all elements  $a_{\rho, \sigma}$  of the "factor system" just introduced are lying in the subgroup  $A$ , and are thus elements of the center of  $G$ , and if any of them is different from the unit element, it has infinite order.*

LEMMA 3.1. *For arbitrary elements  $\rho, \sigma, \tau$  of the factor group  $G/A$  the relation*

$$(7) \quad a_{\rho, \sigma} \cdot a_{\rho\sigma, \tau} = a_{\rho, \sigma\tau} \cdot a_{\sigma, \tau}$$

holds.

Using our previous remark, this follows immediately by comparing the equations

$$\begin{aligned} (g_\rho g_\sigma g_\tau)^m &= ((g_\rho g_\sigma) g_\tau)^m = a_{\rho\sigma, \tau} \cdot (g_\rho g_\sigma)^m \cdot g_\tau^m = \\ &= a_{\rho\sigma, \tau} \cdot a_{\rho, \sigma} \cdot g_\rho^m \cdot g_\sigma^m \cdot g_\tau^m \end{aligned}$$

and

$$\begin{aligned} (g_\rho g_\sigma g_\tau)^m &= (g_\rho (g_\sigma g_\tau))^m = a_{\rho, \sigma\tau} \cdot g_\rho^m (g_\sigma g_\tau)^m = \\ &= a_{\rho, \sigma\tau} \cdot g_\rho^m \cdot a_{\sigma, \tau} \cdot g_\sigma^m \cdot g_\tau^m = a_{\rho, \sigma\tau} \cdot a_{\sigma, \tau} \cdot g_\rho^m \cdot g_\sigma^m \cdot g_\tau^m. \end{aligned}$$

LEMMA 3.2. *If the elements  $\rho$  and  $\sigma$  of the factor group  $G/A$  commute, then  $a_{\rho, \sigma} = 1$ .*

PROOF. By (6) it is enough to show that for  $\rho\sigma = \sigma\rho$

$$(8) \quad g_\rho g_\sigma = g_\sigma g_\rho$$

holds. Hypothesis implies that there is an element  $a$  of  $A$  with

$$(9) \quad g_\rho g_\sigma = a g_\sigma g_\rho.$$

Therefore

$$g_\rho g_\sigma g_\rho^{-1} = a g_\sigma,$$

i. e.

$$\begin{aligned} (g_\rho g_\sigma g_\rho^{-1})^m &= (a g_\sigma)^m, \quad g_\rho g_\sigma^m g_\rho^{-1} = a^m g_\sigma^m, \\ g_\sigma^m &= a^m g_\sigma^m, \quad a^m = 1, \end{aligned}$$

$a$  and  $g_\sigma^m$  being elements of the subgroup  $A$ , and therefore of the center of  $G$ . The equation  $a^m = 1$  implies  $a = 1$ , for  $A$  is torsion free, and thus Lemma 3.2 is proved.

LEMMA 3.3. *All elements  $a_{\rho, \sigma}$  coincide with the unit element of the group.*

PROOF. Let  $\rho$  be an arbitrary fixed element of the factor group  $G/A$ , and let  $\tau$  run over the set of elements of this factor group. We define

$$(10) \quad a(\rho) = \prod_{\tau} a_{\rho, \tau}$$

where the order of the  $m$  factors is irrelevant for they lie in the center of  $G$ . Taking fixed values for  $\rho$  and for  $\sigma$ , let us write down the relation (7) for all elements  $\tau$  of the factor group  $G/A$ . Multiplying corresponding members



of these equations we find

$$(11) \quad a_{\varrho, \sigma}^m \cdot a(\varrho\sigma) = a(\varrho) \cdot a(\sigma),$$

for together with  $\tau$ , also  $\sigma\tau$  runs over the set of elements of the factor group  $G/A$ . If  $\varrho$  and  $\sigma$  commute, then by Lemma 3.2 we have from (11)

$$a(\varrho\sigma) = a(\varrho) \cdot a(\sigma).$$

Putting here  $\sigma = \varrho, \sigma = \varrho^2, \dots$  successively, we obtain

$$(12) \quad \begin{aligned} a(\varrho^2) &= a(\varrho)^2; \\ a(\varrho^3) &= a(\varrho \cdot \varrho^2) = a(\varrho) \cdot a(\varrho^2) = a(\varrho)^3; \\ &\vdots \\ a(\varrho^m) &= a(\varrho)^m. \end{aligned}$$

Now  $\varrho^m = 1$ , and by Lemma 3.2 and by (10) we see that  $a(1)$  is identical with the unit element of the group  $G$ . Therefore it follows from (12) that  $a(\varrho)^m = 1$  for each  $\varrho \in G/A$ . Then  $a(\varrho) = 1$ , for  $a(\varrho) \in A$  and  $A$  is torsion-free. Applying this result to (11), we obtain

$$a_{\varrho, \sigma}^m = 1$$

and this implies  $a_{\varrho, \sigma} = 1$ . This completes the proof of Lemma 3.3.

From our previous statements we obtain easily

LEMMA 3.4. *The derived group of a finitely generated FC-group  $G$  is always a torsion group.*

PROOF. Using the previous notations, from Lemma 3.3 we infer by (6) that for arbitrary elements  $g_\varrho, g_\sigma$  of the group  $G$  we have

$$(g_\varrho g_\sigma)^m = g_\varrho^m g_\sigma^m.$$

A simple induction on the number of factors leads us to the following statement: *a product of arbitrary elements of the group  $G$  can be raised to the  $m$ th power by raising to the  $m$ th power each factor separately.* Now we make use of this rule in the case of an arbitrary element

$$c = [a_1, b_1] \dots = a_1^{-1} b_1^{-1} a_1 b_1 \dots$$

of the derived group  $G'$ . We see that the  $m$ th power of this element, i. e.

$$c^m = a_1^{-m} b_1^{-m} a_1^m b_1^m \dots$$

is equal to the unit element of  $G$ , for the  $m$ th powers  $a_1^{-m}, b_1^{-m}, \dots$  lie in the center of  $G$ . Thus the  $m$ th power of any element of  $G'$  is equal to the unit element, which proves Lemma 3.4.

THEOREM 3.5. *The derived group of an FC-group is always a torsion group.*

PROOF. Let us consider an arbitrary FC-group. Any finitely generated subgroup of such a group is itself an FC-group, and so the validity of the theorem is an immediate consequence of Lemmas 3.4 and 2.6.



### § 4. Properties of FC-groups

Our aim in the present section is to derive from the main theorem, by making use of the lemmas previously stated, the most important properties of FC-groups.

**THEOREM 4.1.** *A torsion free FC-group is always abelian.*

The **PROOF** is easily obtained from the main Theorem 3.5 implying that the derived group of a torsion-free FC-group consists of one element.

**THEOREM 4.2.** *The elements of finite order of an FC-group form a characteristic subgroup; the corresponding factor group is a torsion-free abelian group.*

**PROOF.** Let  $G$  be an arbitrary FC-group. By Theorem 3.5 and by Lemmas 2.7 and 2.8 the elements of finite order of  $G$  evidently form a characteristic subgroup  $H$  of  $G$ , and the factor group  $G/H$  is torsion-free. Theorem 3.5 moreover shows that  $G' \subseteq H$ , and thus  $G/H$  is clearly abelian.

**THEOREM 4.3.** *An FC-group  $G$  generated by elements of finite order is always a torsion group.*

**PROOF.** By Theorem 4.2 the elements of finite order of  $G$  form a subgroup which cannot be a proper one, as the whole group is generated by the set of elements of finite order.

**THEOREM 4.4.** *The elements of finite order of a finitely generated FC-group form a finite subgroup.*

**PROOF.** Let  $G$  be a finitely generated FC-group. It is sufficient to show that  $G$  has only a finite number of elements of finite order, for the fact that these elements form a subgroup, is a consequence of Theorem 4.2. By Lemma 2.5  $G$  has a torsion-free subgroup  $A$  of finite index, contained in the center of  $G$ . We are going to show that any coset of  $G \bmod A$  contains at most one element of finite order. Let us assume that  $c$  and  $d$  are elements of finite order of  $G$ , belonging to the same coset of  $G \bmod A$ . Then  $c = ad$  for some  $a \in A$ . If  $n$  is a common multiple of the orders of  $c$  and  $d$ , then  $c^n = d^n = 1$ , and by  $c = ad$  we thus obtain  $c^n = a^n d^n$ , i. e.

$$(13) \quad a^n = 1,$$

$a$  being an element of the center of  $G$ . On the other hand, we know that  $A$  is torsion-free, so that (13) implies  $A = 1$ , i. e.  $c = d$ . Now as  $A$  has finite index in  $G$ , we see that  $G$  can have only a finite number of elements of finite order. This completes the proof of Theorem 4.4.

**THEOREM 4.5.** *An FC-group generated by a finite number of elements of finite order is always finite.*

The **PROOF** runs along the same lines as that of Theorem 4.3.



**THEOREM 4.6.** *The elements of finite order of an FC-group form a locally finite subgroup.*

**PROOF.** We have to show that a finite number of elements of finite order of an FC-group is always contained in a finite subgroup. This fact is an immediate consequence of Theorem 4.5, any subgroup of an FC-group being also an FC-group.

### § 5. Finite normal subgroups

In this section we are going to prove the theorem of DICMAN and NEUMANN mentioned above, together with some applications. The theorem of DICMAN and NEUMANN reads as follows [4]:

**THEOREM 5.1.** *A finite set of elements of an arbitrary group is contained in a finite normal subgroup of the group if and only if all its elements are of finite order, and have only a finite number of conjugates in the group.*

**PROOF.** The necessity of the conditions is obvious, so we can restrict ourselves to the proof of the sufficiency. Let us consider a finite set  $V$  of elements of a group  $G$  such that the elements of  $V$  are of finite order and have a finite number of conjugates in  $G$ . Now if we adjoin to the set  $V$  all the conjugates of its elements, we obtain a finite set  $V'$  of group elements of finite order. Thus, by Theorem 4.5, the elements of the set  $V'$  generate a finite subgroup  $H$  in  $G$ , for, by Lemma 2.9,  $H$  is obviously an FC-group. On the other hand,  $H$  is clearly a normal subgroup of  $G$ . This proves Theorem 5.1.

**THEOREM 5.2.** *Each element of a group  $G$  belongs to a finite normal subgroup of the group if and only if  $G$  is a torsion FC-group.*

**PROOF.** An application of the preceding theorem to a set of one element show that each element of the group  $G$  is contained in a finite normal subgroup of  $G$  if and only if all elements of  $G$  are of finite order and have only a finite number of conjugates in  $G$ , i. e. if  $G$  is a torsion FC-group.

**THEOREM 5.3.** *Let  $n$  be a fixed natural number. If a group  $G$  contains but a finite number of elements of order  $n$ , then these elements generate a finite characteristic subgroup of  $G$ .*

The **PROOF** follows immediately from Theorem 5.1, for the set of all elements of order  $n$  of a group evidently contains together with any element all its conjugates too. The finite normal subgroup of  $G$  generated by all elements of order  $n$  is characteristic in  $G$ : this follows from the fact that any element of order  $n$  is carried by any automorphism of the group into an element of the same order.



**THEOREM 5.4.** *If the group  $G$  contains but a finite number of elements of finite order, then these elements form a characteristic subgroup  $H$  of  $G$ , and the factor group  $G/H$  is torsion-free.*

**PROOF.** The elements of finite order of  $G$  form a set which contains, together with any of its elements, all of its conjugates. Now if this set is finite, then by Theorem 5.1 the subgroup  $H$  generated by it is also finite, and so this subgroup  $H$  necessarily coincides with the set of all elements of finite order of  $G$ .  $H$  is evidently a characteristic subgroup of  $G$ , and  $G/H$  is torsion-free.

## § 6. Connection between the center and the derived group

Our main aim in this section is to prove a theorem of BAER and NEUMANN. Before doing this we establish the following theorem which gives a sufficient condition for a group to be an FC-group.

**THEOREM 6.1.** *If the derived group of a group  $G$  is finite, then  $G$  is an FC-group.*

**PROOF.** Let us assume that the derived group of the group  $G$  is finite, and let  $a$  be an arbitrary fixed element of  $G$ . We show that  $a$  has only a finite number of conjugates in  $G$ . A conjugate of  $a$  can be written in the form

$$g^{-1}ag = a(a^{-1}g^{-1}ag) = a[a, g].$$

By our hypothesis there is but a finite number of the commutators  $[a, g]$  whence  $a$  has only a finite number of conjugates.

For finitely generated groups the converse of Theorem 6.1 is also true:

**THEOREM 6.2.** *The derived group of a finitely generated FC-group is finite.*

This follows immediately from Theorems 3.5 and 4.4.

Now we prove the theorem of BAER and NEUMANN [1], [3]:

**THEOREM 6.3.** *If in a group the center has finite index, then the derived group of the group is finite.*

**PROOF.** If the center of the group  $G$  has finite index, then, by Lemma 2.2,  $G$  is an FC-group. Now applying Theorem 3.5, we see that the derived group  $G'$  is a torsion group; moreover,  $G'$ , as a subgroup of the FC-group  $G$ , is itself an FC-group. Thus we have only to show that  $G'$  is finitely generated, for then the finiteness of  $G'$  will immediately follow from Theorem 4.5. It is easy to see that the derived group  $G'$  is generated by a finite number of elements. For if  $r$  is the index of the center of  $G$  in  $G$ , and if  $g_1, g_2, \dots, g_r$  is a complete set of representatives of  $G$  modulo the center, then the  $r^2$  commutators  $[g_i, g_k]$  exhaust all possible commutators of  $G$ , considering that  $[g_i, g_k]$  evidently remains unchanged, if for  $g_i$  or  $g_k$  we substitute an element belonging to the same coset modulo the center. This proves the theorem of BAER and NEUMANN.



The theorem of BAER and NEUMANN can also be reversed in the case of finitely generated groups:

**THEOREM 6.4.** *If the finitely generated group  $G$  has a finite derived group  $G'$ , then the center of  $G$  has a finite index in  $G$ .*

**PROOF.** Let the finitely generated group  $G$  have a finite derived group. Then, by Theorem 6.1,  $G$  is an FC-group. Thus the validity of our statement is an immediate consequence of Lemma 2.3.

### § 7. The theorem of Fjodorov

In the present section we prove the following theorem of JU. G. FJODOROV [5]:

**THEOREM 7.1.** *If in an infinite group  $G$  any subgroup containing at least two elements is of finite index, then  $G$  is a cyclic group.*

**PROOF.** Assume  $G$  satisfies the conditions. Then the normalizer of an arbitrary element has a finite index in  $G$ , and therefore, by Lemma 2.1,  $G$  is an FC-group. On the other hand,  $G$  is torsion-free, since the cyclic group generated by an element of finite order is of infinite index in an infinite group. Thus  $G$  is a torsion-free FC-group, and so by Theorem 4.1 it is abelian. One sees immediately that  $G$  is finitely generated, for any of its elements different from the unity generates a cyclic subgroup of finite index. Thus, by the well-known fundamental theorem on finitely generated torsion-free abelian groups,  $G$  is the direct product of a finite number of infinite cyclic groups. This product contains only a single factor, for each factor must be of finite index. This completes the proof.

### § 8. Concluding remarks

In this section we are going to show that the results obtained in the preceding sections are not capable of further improvement. First of all we give an example of an FC-group which cannot be decomposed into the direct product of an abelian group and a finite group. Consider the group  $G$  with the generators  $a, b$  and with the defining relations

$$(14) \quad a^{-1} b^{-1} a b = c, \quad c^2 = 1, \quad a c = c a, \quad b c = c b.$$

In this group  $G$  the elements of finite order form the cyclic group  $\{c\}$  of order two.  $\{c\}$  evidently coincides with the derived group of  $G$ . As by (14)

$$b^{-1} a b = a c,$$

i. e.

$$b^{-1} a^2 b = (b^{-1} a b)^2 = (a c)^2 = a^2 c^2 = a^2$$

holds, we have found that  $a^2$  is in the center of the group  $G$ . We can show



in an analogous way that  $b^2$  also belongs to the center of  $G$ . Consequently, the center of  $G$  is the direct product of two infinite cyclic groups:

$$\{a^2\} \times \{b^2\};$$

thus the center has the index 4 in  $G$ . We see that  $G$  is an FC-group. However,  $G$  cannot be obtained as the direct product of a finite group by an abelian group, for this would mean that  $G$  itself is abelian, as the elements of finite order of  $G$  form an abelian group.

In connection with this example the problem naturally arises, how we could obtain some sort of survey over the structure of the finitely generated FC-groups. Using the notion of Schreier extension our results enable us to describe the finitely generated FC-groups in two ways, each of which gives a rather deep insight into the structure of these groups. As the group  $A$  in Lemma 2.5 is a finitely generated torsion-free abelian group, i. e. a direct product of a finite number of infinite cyclic groups (or, as we sometimes say, a free abelian group of finite rank), therefore the finitely generated FC-groups can be described as follows: *any finitely generated FC-group is a central Schreier extension of a free abelian group of finite rank by a finite group.* On the other hand, from Theorems 4.2 and 4.4 it follows that *any finitely generated FC-group is the Schreier extension of a finite group by a free abelian group of finite rank.* [That the factor group of a finitely generated FC-group with respect to its characteristic subgroup consisting of the elements of finite order is a free abelian group of finite rank, follows from the fact that this torsion free abelian factor group (see Theorem 4.2) is a homomorphic image of a finitely generated group, i. e. is itself finitely generated.]

As we saw, in order that a group be an FC-group, it is sufficient that its center be of finite index, or that its derived group be finite (Lemma 2.2, Theorem 6.1). Moreover we know that both conditions are also necessary, if we consider only finitely generated groups (Lemma 2.3, Theorem 6.2). Now, by giving a counterexample, we show that these conditions are *necessary in case of finitely generated groups, but not in general*, in order that a group be an FC-group. Let us consider an infinite sequence

$$G_1, G_2, G_3, \dots$$

of non-abelian finite simple groups with strictly increasing orders (e. g. the alternating groups  $A_n$  for  $n = 5, 6, 7, \dots$ ). The center of the direct product  $D$  of this infinite set of groups consists of one element, whereas its derived group coincides with the whole group, the same statements being true for the direct factors  $G_i$ . The center of the group  $D$  is therefore not of finite index, and the derived group  $D'$  is not finite (it contains elements of arbitrarily great order). However,  $D$  is an FC-group. This follows from the fact that any element of  $D$  is contained already in the direct product of a finite number of  $G_i$ , and so its normalizer contains the direct product of the "remaining"  $G_m$ 's and is therefore of finite index.



Finally, we show that Theorem 6.3 of BAER and NEUMANN can be reversed only in the case of finitely generated groups (Theorem 6.4). In order to show this, we construct an infinite group  $G$ , in which the center has not a finite index, but the derived group  $G'$  is finite. In our example the same finite subgroup will be the center and the derived group of the group. Let  $p$  be a fixed prime, and let us consider the group with the infinitely many generators  $b, a_1, a_2, a_3, \dots$  and the defining relations

$$\begin{aligned} b^p &= a_1^p = a_2^p = \dots = a_n^p = \dots = 1; \\ a_i b &= b a_i; & (i = 1, 2, 3, \dots) \\ a_{i+k} a_i &= b a_i a_{i+k}; & (i, k = 1, 2, 3, \dots). \end{aligned}$$

It is immediately clear that the derived group of the group  $G$  is equal to the cyclic group  $\{b\}$  of order  $p$ . It is not hard to see that also the center of  $G$  coincides with  $\{b\}$ , for to any element  $g$  of  $G$  not in  $\{b\}$  one can find an element in  $G$  which does not commute with  $g$ .

(Received 22 November 1953)

### Bibliography

- [1] R. BAER, Representations of groups as quotient groups, II. Minimal central chains of a group, *Trans. Amer. Math. Soc.*, **58** (1945), pp. 348—389.
- [2] R. BAER, Finiteness properties of groups, *Duke Math. J.*, **15** (1948), pp. 1021—1032.
- [3] R. BAER, Endlichkeitskriterien für Kommutatorgruppen, *Math. Ann.*, **124** (1952), pp. 161—177.
- [4] А. П. Дицман, О  $p$ -группах, Докл. Акад. Наук СССР, **15** (1937), pp. 71—76.
- [5] Ю. Г. Федоров, О бесконечных группах, все нетривиальные подгруппы которых имеют конечный индекс, Успехи Мат. Наук, **6:1** (1951), pp. 187—189.
- [6] В. Н. NEUMANN, Groups with finite classes of conjugate elements, *Proc. London Math. Soc.*, (3. Ser.) **1** (1951), pp. 178—187.
- [7] O. SCHREIER, Die Untergruppen der freien Gruppen, *Abh. Math. Sem. Hamburg*, **5** (1927), pp. 161—183.

### ТЕОРИЯ ГРУПП С КОНЕЧНЫМИ КЛАССАМИ СОПРЯЖЕННЫХ ЭЛЕМЕНТОВ

И. ЭРДЕШ (Дебрецен)

(Резюме)

Теория вышеупомянутых групп была разработана Р. Бэром [2] и Б. Г. Нейманом [6]. В настоящей работе автор даёт упрощенное по технике доказательств изложение результатов Неймана, пополняя их некоторыми замечаниями, и указанием разных применений.



# DIE OSKULIERENDEN RIEMANNSCHEN RÄUME REGULÄRER CARTANSCHER RÄUME

Von  
ARTHUR MOÓR (Debrecen)  
(Vorgelegt von O. VARGA)

## Einleitung

Die oskulierenden Räume vereinfachen meistens die geometrische Struktur der allgemeinen Räume. Im Finslerschen Raum, sowie in der allgemeinen affinzusammenhängenden Mannigfaltigkeit von Linienelementen stimmt das invariante Differential eines Vektors  $\xi^i(x, \dot{x})$  längs einer Folge von Linienelementen mit demjenigen invarianten Differential von  $\xi^i$  überein, das von dem längs der Folge der Linienelemente oskulierenden Raum bestimmt wird.<sup>1</sup> Ebenso kann behauptet werden, daß die Riemannschen Krümmungstensoren längs der Folge von Linienelementen übereinstimmen, allerdings wenn noch in den Räumen eine vom Wege unabhängige Parallelübertragung der Linienelemente existiert.

Im folgenden werden wir zeigen, daß ein oskulierender Riemannscher Raum auch in den regulären Cartanschen Räumen konstruiert werden kann, was die Struktur des Raumes ebenso vereinfacht, wie im Falle der Linien-elementenräume. Statt der geodätischen Linien werden wir jetzt zur Konstruktion die geodätischen Hyperflächen, also die Hyperebenen anwenden.

Es scheint uns, daß die Hyperebenen in den Cartanschen Räumen mehr für Fundamentalgebilde betrachtet werden können, als die geodätischen Linien. Allerdings existieren nicht immer zu allen Stellungen Hyperebenen. Es gibt zu jeder Stellung nur dann eine Hyperebene, wenn

$$(0.1) \quad R_{i\alpha jk} \stackrel{\text{def}}{=} R_{rsjk} l^s = 0$$

besteht.<sup>2</sup> Wir wollen annehmen, daß diese Gleichung in den von uns behandelten Räumen immer besteht.

Im § 1 stellen wir die Fundamentalbegriffe und Fundamentaltensoren der Cartanschen Räume zusammen. Wir werden aber nur diejenigen Fundamentalgrößen angeben, die hier benützt werden.<sup>3</sup> Im § 2 geben wir die Kon-

<sup>1</sup> Vgl. [2] und [3]. Literaturverzeichnis am Ende dieser Arbeit.

<sup>2</sup> Vgl. [1], Nr. 29

<sup>3</sup> Für eine vollständige Entwicklung der Theorie der Cartanschen Räume vgl. [1].



struktion des oskulierenden Riemannschen Raumes an. Im § 3 wollen wir einige Anwendungen des oskulierenden Riemannschen Raumes besprechen. Wir zeigen in diesem §, daß mit Hilfe des oskulierenden Riemannschen Raumes das invariante Differential eines Vektors im Cartanschen Raum ebenso eingeführt werden kann, wie bei den Finslerschen Räumen, d. h. sich im wesentlichen durch die Bestimmung des invarianten Differentials im Riemannschen Raum bestimmen läßt. Auch stimmen die Riemannschen Krümmungstensoren des Cartanschen und des oskulierenden Riemannschen Raumes längs einer Mannigfaltigkeit von Hyperflächenelementen überein.

### § 1. Fundamentalgrößen der regulären Cartanschen Räume

Ein Cartanscher Raum ist eine Mannigfaltigkeit der Hyperflächenelemente

$$(x, u) = (x^1, x^2, \dots, x^n, u_1, u_2, \dots, u_n),$$

in der das Oberflächenelement durch die Funktion

$$dO = \frac{F(x, u)}{u_n} dx^1 \cdot dx^2 \dots dx^{n-1}$$

angegeben ist. Die in der Formel des Oberflächenelementes  $dO$  vorkommende Funktion:  $F(x, u)$  ist die Grundfunktion des Cartanschen Raumes. Sie ist in den  $u_i$  positiv homogen von erster Dimension.

Ein Cartanscher Raum ist von  $(2n-1)$  Dimension, denn bei den  $u_i$  kommt nur ihr Verhältnis in Betracht. Die  $u_i$  kann man also immer mit einem willkürlichen Faktor  $q$  multiplizieren;  $(x^i, q u_i)$  bestimmt dasselbe Hyperflächenelement wie  $(x^i, u_i)$ .

Bei einer Koordinatentransformation

$$\bar{x}^i = \bar{x}^i(x^1, x^2, \dots, x^n)$$

— die wir als umkehrbar und eindeutig voraussetzen — transformieren sich die  $u_i$  wie kovariante Vektordichten vom Gewicht  $-1$ . Die  $u_i$  bestimmen die Normalenrichtung des Hyperflächenelementes  $(x^i, u_i)$ .

Alle Fundamentalgrößen eines Cartanschen Raumes sind durch die Grundfunktion  $F(x, u)$  bestimmt. Insbesondere ist die Oberfläche  $O$  einer Hyperfläche

$$(1.1) \quad x^i = x^i(v^\alpha) \quad (\alpha = 1, \dots, n-1)$$

durch das Integral

$$O = \int_{(n-1)} F(x, p) dv^1 \dots dv^{n-1}$$

definiert, wo die  $p_i$  durch die Formel

$$(1.2) \quad p_i = (-1)^{i+1} \det \left| \frac{\partial x^r}{\partial v^\sigma} \right| \quad \left( \begin{array}{l} r = 1, \dots, (i-1), (i+1), \dots, n \\ \sigma = 1, 2, \dots, (n-1) \end{array} \right)$$

bestimmt sind. Die Hyperfläche (1.1) ist die Mannigfaltigkeit der  $(n-1)$



parametrischen Schar der Hyperflächenelemente  $[x^i(v^\alpha), p_i(v^\alpha)]$ ; dabei sind die  $p_i(v^\alpha)$  durch (1.2) festgelegt.

Aus der Fundamentalfunktion  $F(x, u)$  kann man den metrischen Grundtensor  $g^{ik}$  durch die Formel

$$(1.3) \quad g^{ik}(x, u) = \alpha^{-\frac{1}{n-1}} \frac{\partial^2 ({}^1_2 F^2)}{\partial u_i \partial u_k}$$

mit

$$\alpha = \det \left| \frac{1}{2} \frac{\partial^2 F^2}{\partial u_i \partial u_k} \right|$$

bestimmen. Wie gewöhnlich, kann man zu den kontravarianten Komponenten des metrischen Grundtensors  $g^{ik}$  die kovarianten Komponenten  $g_{jk}$  durch die Gleichungen

$$g^{ik} g_{jk} = \delta_j^i = \begin{cases} 1, & \text{wenn } i=j, \\ 0, & \text{wenn } i \neq j \end{cases}$$

bestimmen, und dann mit Hilfe von  $g^{ik}$ , bzw.  $g_{jk}$  die Indizes der Tensoren herauf- oder herabziehen.

Der Normaleinheitsvektor des Hyperflächenelementes  $(x, u)$  hat die kovarianten Komponenten:

$$(1.4) \quad l_i = \frac{\sqrt{g}}{F} u_i, \quad l^i = \frac{1}{\sqrt{g}} \frac{\partial F}{\partial u_i},$$

wo

$$g = \det |g_{ik}|$$

bedeutet.

Die Torsion des Raumes ist durch den Torsionstensor

$$(1.5) \quad A_{ij}^k = \frac{1}{2} \frac{F}{\sqrt{g}} \frac{\partial g_{ij}}{\partial u_k} = \frac{F}{\sqrt{g}} C_{ij}^k$$

bestimmt. Für den Torsionstensor besteht die für das Spätere wichtige Relation:

$$(1.6) \quad A_{ij}^k l^j = A_i^{jk} l_j = l_i A^k,$$

wo

$$(1.7) \quad A^k = A_m^{mk} = -F \frac{\partial}{\partial u_k} \left( \frac{1}{\sqrt{g}} \right)$$

bedeutet.<sup>4</sup> Wenn der Operator  $\|{}^k$  — angewandt auf eine Funktion  $f$  —

$$f \|{}^k = \frac{F}{\sqrt{g}} \frac{\partial f}{\partial u_k}$$

bedeutet, dann kann (1.7) auch in der Form

$$(1.7^*) \quad A^k = -\sqrt{g} \left( \frac{1}{\sqrt{g}} \right) \|{}^k$$

<sup>4</sup> Da wegen (1.5)  $A_m^i{}^k = A_i^{mk}$  besteht, soll  $A_i^{mk} \stackrel{\text{def}}{=} A_i^{mk}$  gesetzt werden.



dargestellt werden. Da  $g_{jk}$  in den  $u_i$  von nulltem Grade homogen ist, erhält man aus den Gleichungen (1.5) und (1.7)

$$(1.8) \quad A_{ij}^k l_k = A^k l_k = 0.$$

Im folgenden beschränken wir uns auf reguläre Cartansche Räume, die dadurch ausgezeichnet sind, daß der Tensor

$$(1.9) \quad H_r^t = \delta_r^t + A^i A_{ir}^t$$

oder, was dasselbe ist,

$$H^{tr} = g^{tr} + A_k A^{ktr}$$

den Rang  $n$  hat. Das zu  $H_r^t$  inverse System pflegt man mit  $K_m^s$  zu bezeichnen. Es bestehen die Relationen

$$(1.10) \quad H_r^t K_s^r = H_s^r K_r^t = \delta_s^t.$$

Das invariante Differential eines Vektors  $\xi^i(x, u)$  längs einer Kurve

$$x^i = x^i(t)$$

ist durch

$$(1.11) \quad D\xi^i = \frac{d\xi^i}{dt} + \Gamma_{jk}^i \xi^j \frac{dx^k}{dt} + C_j^{ik} \xi^j \frac{du_k}{dt}$$

definiert. Dabei ist im Cartanschen Raum dem Punkt  $x^i(t)$  dasjenige Hyperflächenelement  $u_i(t)$  zugeordnet, das von der Kurve senkrecht durchsetzt wird. Selbstverständlich gibt die Formel (1.11) das invariante Differential eines Vektors  $\xi^i$  auch längs einer Folge der Hyperflächenelemente

$$x^i = x^i(t), \quad u_i = u_i(t)$$

an. Die Größe  $C_j^{ik}$  ist durch die Formel

$$(1.12) \quad C_j^{ik} = g^{it} C_{jt}^k = \frac{\sqrt{g}}{F} A_j^{ik}$$

festgelegt. Für die Bestimmung der  $\Gamma_{jk}^i$  aus der Grundfunktion verweisen wir auf die Arbeit von L. BERWALD.<sup>5</sup>

Wir geben noch folgende Zusammenhänge an:

$$(1.13) \quad \Gamma_{jk}^i = g^{it} \Gamma_{jtk},$$

$$(1.13^*) \quad \Gamma_{jtk} = G_{jtk} + A_{tk}^s \Gamma_{soj} - A_{jk}^s \Gamma_{sot},$$

wo

$$G_{jtk} = \frac{1}{2} \left( \frac{\partial g_{jt}}{\partial x^k} + \frac{\partial g_{tk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^t} \right)$$

ist. Der Index „o“ bedeutet eine Überschiebung mit dem Einheitsvektor. Wir werden noch die Größen  $\Gamma_{ijk}^*$  benutzen, die durch

$$(1.14) \quad \Gamma_{ijk} = \Gamma_{ijk}^* - A_{ij}^s \Gamma_{sok}$$

<sup>5</sup> Vgl. [1], § 1, S. 193—207.



definiert sind. Wegen (1.14) und (1.13\*) folgt sofort die Formel

$$(1.15) \quad \Gamma_{ijk}^* = G_{ijk} + A_{ij}^s \Gamma_{sol}^* + A_{jk}^s \Gamma_{soi}^* - A_{ik}^s \Gamma_{soj}^*.$$

Wir benötigen noch den Krümmungstensor

$$(1.16) \quad \bar{R}_{ikm}^j = \frac{\partial \Gamma_{ik}^{*j}}{\partial x^m} + \Gamma_{ik}^{*j} \parallel^t \Gamma_{tom}^* - \frac{\partial \Gamma_{im}^{*j}}{\partial x^k} - \Gamma_{im}^{*j} \parallel^t \Gamma_{tok}^* + \\ + \Gamma_{ik}^{*t} \Gamma_{tm}^{*j} - \Gamma_{im}^{*t} \Gamma_{tk}^{*j}$$

und den Tensor

$$(1.17) \quad \bar{R}_{iojk} = \frac{\partial \Gamma_{ioj}^*}{\partial x^k} + \Gamma_{ioj}^* \parallel^t \Gamma_{tok}^* - \left( \frac{\partial \Gamma_{iok}^*}{\partial x^j} + \Gamma_{iok}^* \parallel^t \Gamma_{toj}^* \right).$$

Die folgenden Untersuchungen beziehen sich ausschließlich auf reguläre Cartansche Räume.

## § 2. Konstruktion des oskulierenden Riemannschen Raumes

Es sei eine einparametrische Mannigfaltigkeit von Hyperflächenelementen

$$(2.1a) \quad x^i = x^i(v),$$

$$(2.1b) \quad u_i = u_i(v)$$

gegeben. Wir setzen voraus, daß die in (2.1) auftretenden Funktionen hinreichend oft stetig differenzierbar sind. Durch jedes Hyperflächenelement aus (2.1) legen wir eine Hyperebene

$$(2.2a) \quad x^i = x^i(t^1, t^2, \dots, t^{n-1}),$$

$$(2.2b) \quad p_i = p_i(x^1(t^\alpha), x^2(t^\alpha), \dots, x^n(t^\alpha)), \quad (t^\alpha = t^1, t^2, \dots, t^{n-1})$$

hindurch, wo die  $(x^i(t^\alpha), p_i(x(t^\alpha)))$  das zu den Parameterwerten  $t^1, t^2, \dots, t^{n-1}$  gehörige Hyperflächenelement der Hyperebene (2.2a) bedeuten. Eine Hyperebene ist dabei in dem von uns betrachteten Falle, in dem

$$R_{iojk} = 0$$

besteht,<sup>6</sup> nach Angabe eines Hyperflächenelementes eindeutig durch das System

$$\frac{\partial^2 x^i}{\partial t^\alpha \partial t^\sigma} + \Gamma_{jk}^{*i} \frac{\partial x^j}{\partial t^\alpha} \frac{\partial x^k}{\partial t^\sigma} = \left\{ \begin{matrix} \lambda \\ \rho \sigma \end{matrix} \right\} \frac{\partial x^i}{\partial t^\lambda}, \\ \frac{\partial l^i}{\partial t^\alpha} + \Gamma_{jk}^{*i} l^j \frac{\partial x^k}{\partial t^\alpha} = 0$$

von partiellen Differentialgleichungen bestimmt.<sup>7</sup> Die  $p_i(t^\alpha)$  in der Gleichung (2.2b) sind durch (1.2) bestimmt, nur sind jetzt die Parameter mit  $t^1, t^2, \dots, t^{n-1}$  bezeichnet.

<sup>6</sup> Vgl. die Einleitung, insbesondere (0.1).

<sup>7</sup> Vgl. [1], S. 236, Formel (29.6) und (29.7). Für  $\left\{ \begin{matrix} \lambda \\ \rho \sigma \end{matrix} \right\}$  vgl. Fußnote 11.



Wir beschränken uns im folgenden auf einen Bereich  $\mathfrak{B}$ , der in einer solchen Umgebung der Kurve (2. 1a) liegt, in der die Hyperebenen den Raum schlicht bedecken. Mit diesem Verfahren haben wir erreicht, daß jedem Punkt  $x^i$  von  $\mathfrak{B}$  eindeutig ein Hyperflächenelement zugeordnet ist. Es ist also

$$(2. 3a) \quad p_i = p_i(x^1, x^2, \dots, x^n).$$

Wir wählen den Parameter  $v$  in (2. 1) gleich so, daß

$$(2. 3b) \quad u_i(v) = p_i(x^1(v), x^2(v), \dots, x^n(v))$$

bestehe.

Wir definieren jetzt einen längs (2. 1) oskulierenden Riemannschen Raum. Das geschieht durch die Angabe des metrischen Grundtensors  $\gamma_{ik}(x)$  des oskulierenden Raumes. Es ist

$$(2. 4a) \quad \gamma^{ik}(x) = g^{ik}(x, u(x)) \quad (u_j(x) = p_j(x)),$$

$$(2. 4b) \quad \gamma_{ik}(x) = g_{ik}(x, u(x)).$$

Offensichtlich ist der Tensor  $\gamma_{ik}(x)$  in  $\mathfrak{B}$  definiert; ebenso ist der normale Einheitsvektor  $l_i$  wegen (2. 3a) und wegen (1. 4) in  $\mathfrak{B}$  als ein vom Orte abhängiger Vektor festgelegt.

Wählen wir einen Punkt  $\bar{x}^i$  aus  $\mathfrak{B}$ , und einen Punkt  $x^i(v)$  aus (2. 1a), so daß

$$(2. 5) \quad |\bar{x}^i - x^i(v)| < \varepsilon \quad (v \text{ fest})$$

bestehe, wo  $\varepsilon$  eine beliebig klein vorgebbare Größe ist.

Wir treffen nun folgende Annahme:

*Der Vektor  $l_i(x)$  soll im Riemannschen Raum in den beiden Punkten  $\bar{x}^i$  und  $x^i(v)$  parallel sein, wenn Größen höherer als erster Ordnung in  $\varepsilon$  vernachlässigt werden.*

Die anschauliche Bedeutung von unserer Annahme betreffs des Vektors  $l_i(x)$  ist die folgende: liegen die Mittelpunkte  $x^i$  der Hyperflächenelemente von (2. 3a) in einer schmalen Umgebung von (2. 1a), so sind die normalen Einheitsvektoren dieser Hyperflächenelemente im oskulierenden Riemannschen Raum in erster Annäherung parallel.

Wir legen nun durch die Punkte  $x^i(v), \bar{x}^i$  eine Kurve

$$x^i = x^i(\tau) \quad (0 \leq \tau \leq 1),$$

$$x^i(0) = x^i(v), \quad x^i(1) = \bar{x}^i$$

hindurch. Für jeden ihrer Punkte  $x^i(\tau)$  soll (2. 5) bestehen. Eine Kurve, die die genannten Punkte enthält und eine sehr einfache Parameterdarstellung hat, ist die folgende:

$$(2. 6) \quad x^i = x^i(v) + \tau(\bar{x}^i - x^i(v)) \quad (v \text{ fest}, 0 \leq \tau \leq 1).$$

Nach unserer Annahme soll der Vektor  $l_i(x)$  längs der Kurve

$$x^i = x^i(\tau)$$



im Riemannschen Raum parallel sein. Das ergibt die Gleichung

$$\left( \frac{\partial l_i}{\partial x^r} - \overset{(0)}{\Gamma}_{ir}^s l_s \right) \frac{dx^r}{d\tau} = 0,$$

wo die  $\overset{(0)}{\Gamma}_{ir}^s$  die aus  $\gamma_{ik}$  gebildeten Christoffelklammern bedeuten. Man erhält aus diesen Gleichungen in Hinblick auf (2.6)

$$(2.7) \quad \left( \frac{\partial l_i}{\partial x^r} - \overset{(0)}{\Gamma}_{ir}^s l_s \right) (\bar{x}^r - x^r(v)) = 0.$$

Die Größen  $l_i$ ,  $\frac{\partial l_i}{\partial x^r}$ ,  $\overset{(0)}{\Gamma}_{ir}^s$  sollen in (2.7) alle längs (2.6) gebildet werden. Nach dem Mittelwertsatz der Differentialrechnung bekommt man aber z. B. für  $l_s$ :

$$l_s(x^i(v) + \tau(\bar{x}^i - x^i(v))) = l_s(x^i(v)) + \tau(\bar{x}^k - x^k(v)) \frac{\partial l_s}{\partial x^k},^8$$

und ebenso für die anderen Größen. Wenn alle diese Werte in (2.7) hineingesetzt werden, und wir die Glieder höherer Ordnung in  $\varepsilon$  vernachlässigen, dann werden wir eine Gleichung von der Form (2.7) erhalten, in der aber die Größen  $l_i$ ,  $\frac{\partial l_i}{\partial x^r}$ ,  $\overset{(0)}{\Gamma}_{ir}^s$  längs  $x^j(v)$  ( $v$  fest) zu bilden sind. Beachten wir jetzt die Willkürlichkeit des Punktes  $\bar{x}^i$ , so folgt aus (2.7) die Gleichung

$$(2.8) \quad \frac{\partial l_i}{\partial x^r} - \overset{(0)}{\Gamma}_{ir}^s l_s = 0$$

längs (2.1a).

Substituieren wir nun die Werte von  $l_i$  aus der Gleichung (1.4) in (2.8), so erhält man, daß längs (2.1a)

$$(2.9) \quad \frac{\partial u_i}{\partial x^r} = u_s \overset{(0)}{\Gamma}_{ir}^s - f_r u_i$$

besteht, mit

$$f_r = - \frac{\partial \log \frac{\sqrt{g}}{F}}{\partial x^r}.$$

Nun sollen die  $\overset{(0)}{\Gamma}_{ir}^s$  mit den Größen des Cartanschen Raumes ausgedrückt werden, was wir jetzt durchführen wollen.

Nach (2.4b) und (1.5) wird

$$(2.10) \quad \overset{(0)}{\Gamma}_{ijk} = G_{ijk} + \frac{\sqrt{g}}{F} \left( A_{ij}^s \frac{\partial u_s}{\partial x^k} + A_{jk}^s \frac{\partial u_s}{\partial x^i} - A_{ik}^s \frac{\partial u_s}{\partial x^j} \right),$$

wo

$$G_{ijk} = \frac{1}{2} \left( \frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} \right)$$

<sup>8</sup> Die Argumente von  $\frac{\partial l_i}{\partial x^k}$  sind:  $x^i(v) + \vartheta \tau(\bar{x}^i - x^i(v))$  ( $0 < \vartheta < 1$ ).



bedeutet. Nach (1.6) wird aus (2.9)

$$(2.11) \quad \frac{\partial u_s}{\partial x^k} = \frac{F}{\sqrt{g}} G_{sok} + l_s A^r \frac{\partial u_r}{\partial x^k} + l_k A^r \frac{\partial u_r}{\partial x^s} - A_{sk}^r \frac{\partial u_r}{\partial x^j} l^j - f_k u_s.$$

Überschieben wir nun diese Gleichung mit  $l^k$ , und dann mit  $A^s$ , so bekommen wir wegen (1.6) und (1.8)

$$(2.12a) \quad \frac{\partial u_s}{\partial x^k} l^k = \frac{F}{\sqrt{g}} G_{soo} + A^r \frac{\partial u_r}{\partial x^s} - f_0 u_s,$$

$$(2.12b) \quad \frac{\partial u_s}{\partial x^k} A^s = \frac{F}{\sqrt{g}} A^s G_{sok} + l_k A^s A^r \frac{\partial u_r}{\partial x^s} - A^s A_{sk}^r \frac{\partial u_r}{\partial x^j} l^j.$$

Im letzten Glied von (2.12b) können wir  $\frac{\partial u_r}{\partial x^j} l^j$  durch den Wert von (2.12a) ersetzen. Bei Beachtung von (1.8) bekommt man nach (1.9)

$$(2.13) \quad (H_k^s - l_k A^s) \frac{\partial u_r}{\partial x^s} A^r = \frac{F}{\sqrt{g}} (A^s G_{sok} - A^s A_{sk}^r G_{roo}).$$

Überschieben wir nun diese Gleichung mit  $K_t^k$  und dann mit  $(\delta_p^t + l_p A^t)$ , so wird wegen (1.10) und auf Grund der Identität

$$l_k K_t^k = l_t$$

nach (2.13)

$$(2.14a) \quad \frac{\partial u_r}{\partial x^p} A^r = \frac{F}{\sqrt{g}} A^r (\delta_p^t + l_p A^t) K_t^h (G_{roh} - A_{rh}^s G_{soo}).$$

Wegen (2.12a) bekommt man

$$(2.14b) \quad \frac{\partial u_s}{\partial x^k} l^k = \frac{F}{\sqrt{g}} \{G_{soo} + A^r (\delta_s^t + l_s A^t) K_t^h (G_{roh} - A_{rh}^p G_{poo})\} - f_0 u_s.$$

Jetzt sind wir imstande, die explizite Form von  $\frac{\partial u_s}{\partial x^k}$  zu bestimmen.

Dazu müssen wir nur die Größen  $\frac{\partial u_r}{\partial x^s} A^r$  und  $\frac{\partial u_s}{\partial x^k} l^k$  aus (2.14b) in die Gleichung (2.11) einsetzen. Nach einer längeren Rechnung erhält man

$$(2.15) \quad \frac{\partial u_s}{\partial x^k} = \frac{F}{\sqrt{g}} \{G_{sok} - l_k G_{soo} - l_s G_{koo} - 2l_s l_k A^t G_{too} + \\ + (l_k \delta_s^t + l_s \delta_k^t + 2l_s l_k A^t - A_{sk}^t) K_t^r (G_{rov} + A^p G_{por})\} - f_k u_s.$$

Die explizite Form der Größen  $\Gamma_{sjk}^*$  hat BERWALD ([1], Gleichung (7.16)) bestimmt. Wenn wir daraus  $\Gamma_{sok}^*$  unter Anwendung der Gleichungen (1.8),

$$A^p A_{pq}^r = H_q^r - \delta_q^r,$$

und (1.10) berechnen, dann kann leicht gezeigt werden, daß nach (2.15)

$$(2.16) \quad \frac{\partial u_s}{\partial x^k} = \frac{F}{\sqrt{g}} \Gamma_{sok}^* - f_k u_s$$



längs (2. 1a) besteht. Aus der Gleichung (1. 13) folgt nun auf Grund von (1. 6)

$$(2. 17) \quad \frac{\partial u_s}{\partial x^k} = \frac{F}{\sqrt{g}} (\Gamma_{sok} + l_s A^t \Gamma_{tok}) - f_k u_s.$$

Substituiert man diese Größen in die Gleichung (2. 10), so bekommt man die für den oskulierenden Riemannschen Raum sehr charakteristische Identität:

$$(2. 18) \quad \overset{(o)}{\Gamma}_{ijk}(x) = I_{ijk}^*(x, u(x)),$$

die längs (2. 1) gültig ist.

In den Formeln (2. 16) und (2. 17) steht ein Glied von der Form

$$h_k u_s,$$

die in (2. 16)

$$h_k = f_k = \frac{\partial \log \frac{\sqrt{g}}{F}}{\partial x^k}$$

und in (2. 17)

$$h_k = A^t \Gamma_{tok} - \frac{\partial \log \frac{\sqrt{g}}{F}}{\partial x^k}$$

bedeutet. Das Glied  $h_k u_s$  besaß schon in der Formel (2. 18) keine besondere Rolle, und wird auch im folgenden keine Bedeutung haben. Für die Funktionen die in  $u_s$  homogen von nullter Dimension sind, besteht nämlich die Eulersche Relation:

$$\frac{\partial \psi}{\partial u_s} u_s = 0;$$

die Glieder  $h_k u_s$  werden im folgenden immer in derartigen Relationen vorkommen, für die die Eulersche Relation gültig ist. Das bedeutet im wesentlichen, daß

$$I_{sok}^* = \Gamma_{sok} + l_s A^t \Gamma_{tok}$$

in den Formeln mit  $\Gamma_{sok}$  gleichberechtigt ist.

### § 3. Anwendungen des oskulierenden Riemannschen Raumes

Wir beweisen den Satz:

SATZ I. *Das invariante Differential eines längs einer Folge von Hyperflächenelementen*

$$(3. 1) \quad x^i = x^i(t), \quad u_i = u_i(t)$$

*definierten Vektors  $\xi^i(x, u)$  im Cartanschen Raum ist mit dem invarianten Differential des Vektors  $\xi^i(x, u(x))$  im längs (3. 1) oskulierenden Riemannschen Raum identisch.*



BEWEIS. Wir bilden das invariante Differential des Vektors  $\xi^i(x, u)$  für den längs (3. 1) oskulierenden Riemannschen Raum. Es wird

$$(3. 2) \quad \frac{D\xi^i}{dt} = \frac{d\xi^i}{dt} + \Gamma_{jk}^{(e)i} \xi^j \frac{dx^k}{dt}.$$

Dabei ist wegen (2. 4a) längs (2. 1)<sup>9</sup>

$$(3. 3) \quad \Gamma_{jk}^{(e)i}(x) = \gamma^{ir} \Gamma_{jrk}^{(e)} = g^{ir}(x, u(x)) \Gamma_{jrk}^{(e)}.$$

Nach der Formel (2. 3b) wird längs (2. 1) die Gleichung

$$(3. 4) \quad \frac{\partial u_s}{\partial x^k} \frac{dx^k}{dt} = \frac{du_s}{dt}$$

bestehen. Setzt man den Wert von  $\Gamma_{ik}^{(e)j}$  aus (3. 3) und (2. 10) in die Relation (3. 2) ein, dann wird wegen (3. 4) das invariante Differential von  $\xi^i$  die Form

$$(3. 5) \quad \frac{D\xi^i}{dt} = \frac{d\xi^i}{dt} + \frac{\sqrt{g}}{F} A_j^{is} \xi^j \frac{du_s}{dt} + \Gamma_{jk}^i \xi^j \frac{dx^k}{dt}$$

haben, wo auf Grund der Gleichung (2. 17)

$$\Gamma_{jk}^i = g^{ir} \Gamma_{jrk}, \\ \Gamma_{jrk} = G_{jrk} + A_{rk}^s \Gamma_{soj} - A_{jk}^s \Gamma_{sor}$$

ist. Man kann sofort feststellen, daß diese beiden letzten Gleichungen mit (1. 13) und (1. 13\*) übereinstimmen. Beachten wir noch (1. 12), so kann sofort verifiziert werden, daß (3. 5) mit (1. 11) übereinstimmt, w. z. b. w.

Der Gedanke ist naheliegend, daß auf Grund dieses Satzes das invariante Differential in die Cartanschen Räume ebenso eingeführt werden kann, wie dies im Finslerschen Raum von O. VARGA<sup>10</sup> durchgeführt wurde. Im Cartanschen Raum ist aber die Anwendung dieses Verfahrens mit weiteren Schwierigkeiten verknüpft. Bei der Konstruktion des oskulierenden Riemannschen Raumes haben wir nämlich die Hyperebenen benutzt, und das bedingt schon die Existenz des Parallelismus der Hyperflächenelemente, also die Existenz des invarianten Differentials.

Bedenken wir aber, daß man in den Cartanschen Räumen die Hyperebenen auch durch gewisse Differentialgleichungen charakterisieren kann, dann werden wir auch diese Schwierigkeiten beseitigen können.

Die charakteristischen Differentialgleichungen der Hyperebenen sind die folgenden:<sup>11</sup>

$$(3. 6a) \quad \frac{\partial^2 x^i}{\partial v^\sigma \partial v^\tau} + \Gamma_{jk}^{*i} \frac{\partial x^j}{\partial v^\sigma} \frac{\partial x^k}{\partial v^\tau} = \begin{Bmatrix} \varrho \\ \sigma\tau \end{Bmatrix} \frac{\partial x^i}{\partial v^\varrho},$$

$$(3. 6b) \quad \frac{\partial l^i}{\partial v^\varrho} + \Gamma_{jk}^{*i} l^j \frac{\partial x^k}{\partial v^\varrho} = 0,$$

<sup>9</sup> Setzt man in der Gleichung (3. 1)  $t = v$ , dann bekommt man offenbar die Gleichung (2. 1). Die Gleichung (3. 1) ist also im wesentlichen mit der Folge (2. 1) identisch.

<sup>10</sup> Vgl. [3].

<sup>11</sup> Vgl. [1], Gleichungen (29. 6) und (29. 7).



wo die  $l^i$  durch

$$l^i = \frac{1}{\sqrt{g}} \frac{\partial F(x, p)}{\partial p_i}$$

und die  $p_i$  durch (1.2) bestimmt sind. Die Größen  $\Gamma_{jk}^{*i}$  sind dabei aus der Grundfunktion  $F$  in der von L. BERWALD angegebenen Weise ableitbar<sup>12</sup> und die  $\left\{ \begin{smallmatrix} \varrho \\ \sigma\tau \end{smallmatrix} \right\}$  bedeuten die zu den  $g_{\alpha\beta}$  gehörigen Christoffelsymbole 2-ter Art. Dabei sind die  $g_{\alpha\beta}$  die durch

$$g_{\alpha\beta} = g_{ik} \frac{\partial x^i}{\partial v^\alpha} \frac{\partial x^k}{\partial v^\beta}$$

bestimmten Flächenkomponenten des Fundamentaltensors. Damit die Konstruktion des oskulierenden Riemannschen Raumes durchführbar sei, müssen wir fordern, daß durch ein Hyperflächenelement des Raumes eben eine Lösung

$$x^i = x^i(v^1, v^2, \dots, v^{n-1})$$

der Gleichungen (3.6a) und (3.6b) hindurchgeht, d. h. diese Gleichungen vollständig integrierbar sind. Dies bedeutet, daß  $R_{ijk} = 0$  und  $B_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta}$  besteht.<sup>13</sup>

Jetzt läßt sich auf Grund des Satzes I das invariante Differential in die Cartanschen Räume mit Hilfe des oskulierenden Riemannschen Raumes ebenso einführen, wie dies in den Finslerschen Räumen schon durchgeführt wurde. Nachher kann selbstverständlich sofort gezeigt werden, daß die Gleichungen (3.6a) und (3.6b) geometrisch den Parallelismus der Hyperflächenelemente bedeuten.

Wir stellen jetzt den Zusammenhang zwischen dem Riemannschen Krümmungstensor des Cartanschen und dem des oskulierenden Riemannschen Raumes fest.

Wir werden dazu die Gleichung (2.10) umformen. Wegen der Gleichung (1.4) wird man auf Grund der Identitäten (1.8) aus (2.10) die Gleichung

$$\overset{(\varrho)}{I}_{ijk} = G_{ijk} + A_{ij}^s \frac{\partial I_s}{\partial x^k} + A_{jk}^s \frac{\partial I_s}{\partial x^i} - A_{ik}^s \frac{\partial I_s}{\partial x^j}$$

erhalten. Wir bemerken, daß diese Gleichung im ganzen Bereich  $\mathfrak{B}$  gültig ist. Wenn wir hiernach diese Gleichung nach  $x^l$  differenzieren, und die Homogenität nullten Grades der Größen in den  $u_i$ , weiter die Gleichungen (1.4), (2.8) und (2.18) beachten, so wird:

$$(3.7) \quad \frac{\partial \overset{(\varrho)}{I}_{ijk}}{\partial x^l} = \frac{\partial I_{ijk}^*}{\partial x^l} + I_{ijk}^* \parallel^l I_{tol}^* + A_{ij}^s (skl) + A_{jk}^s (sil) - A_{ik}^s (sjl),$$

<sup>12</sup> Vgl. [1], Gleichung (7.16).

<sup>13</sup> Vgl. [1], S. 236.



wo

$$(3.7a) \quad (skl) = \frac{\partial^2 l_s}{\partial x^k \partial x^l} - \frac{\partial \Gamma_{sok}^*}{\partial x^l} - \Gamma_{sok}^* \parallel^t \Gamma_{tol}^*$$

bedeutet. Die Gleichung (3.7) besteht schon nur längs der angegebenen Folge (2.1) von Hyperflächenelementen, da die bei der Berechnung der Relation (3.7) benützte Formel (2.8) nur längs (2.1) gültig ist.

Der Krümmungstensor des Riemannschen Raumes ist durch die Formel

$$R_{ijkl}^{(e)} = \frac{\partial \Gamma_{ijk}^{(e)}}{\partial x^l} - \frac{\partial \Gamma_{ijl}^{(e)}}{\partial x^k} + \Gamma_{il}^{(e)} \Gamma_{jtk}^{(e)} - \Gamma_{ik}^{(e)} \Gamma_{jtl}^{(e)}$$

festgelegt. Nach den Gleichungen (2.18) und (3.7) bekommt man aus dieser Gleichung

$$(3.8) \quad R_{ijkl}^{(e)} = \bar{R}_{ijkl} + R_{ijkl}^*,$$

wo  $\bar{R}_{ijkl}$  den Krümmungstensor des Cartanschen Raumes und  $R_{ijkl}^*$  den Tensor

$$(3.9) \quad R_{ijkl}^* = A_{jk}^s (sil) - A_{ik}^s (sjl) - [A_{jl}^s (sik) - A_{il}^s (sjk)]$$

bedeutet, denn es ist

$$(skl) - (slk) = \bar{R}_{solk} = 0.$$

Diese Gleichung folgt sofort aus (3.7a), wenn wir noch die Relation

$$(3.10) \quad \bar{R}_{solk} = (\delta_s^r + l_s A^r) R_{rolk}$$

und unsere Annahme (0.1) beachten.<sup>14</sup>

Der Tensor  $R_{ijkl}^*$  bestimmt die Abweichung des Krümmungstensors des oskulierenden Riemannschen Raumes von dem des Cartanschen Raumes. Aus (3.8) folgt der

SATZ II. *Verswindet längs der Hyperflächenelementfolge (2.1) der Abweichungstensor  $R_{ijkl}^*$ , so wird längs (2.1) der Riemannsche Krümmungstensor  $R_{ijkl}^{(e)}$  mit dem Krümmungstensor  $\bar{R}_{ijkl}$  des Cartanschen Raumes übereinstimmen.*

Wir werden jetzt einen expliziten Fall angeben, in dem

$$R_{ijkl}^* = 0$$

in dem ganzen Bereich  $\mathfrak{B}$  besteht. Dazu schreiben wir die Gleichung (2.8) auf Grund von (2.18) in der Form

$$(3.11) \quad \frac{\partial l_s}{\partial x^k} - \Gamma_{sok}^*(x, l(x)) = 0.$$

Wegen der Homogenität nullten Grades ist offenbar  $\Gamma_{sok}^*(x, u) = \Gamma_{sok}^*(x, l)$ . Wir fordern nun, daß die Relation (3.11) im ganzen Bereich  $\mathfrak{B}$  gültig sei. Das können wir immer erreichen, denn die Integrabilitätsbedingung von

<sup>14</sup> Die Relation (3.10) folgt unmittelbar aus der Formel (12.12) des Aufsatzes von L. BERWALD [1], wenn wir die Gleichung (12.12) mit  $(\delta_s^r + l_s A^r)$  überschieben.



(3. 11) ist

$$\bar{R}_{solk} = 0,$$

und dieser Tensor verschwindet identisch nach (0. 1) und (3. 10) im ganzen Cartanschen Raum.

Es wird somit (3. 11) für den Lösungsvektor  $l_i(x)$  eine Identität, von der wir nach Ableitung nach  $x^l$  in Hinsicht auf (1. 4)

$$(skl) = 0$$

erhalten; dann wird aber nach (3. 9) der Tensor  $R_{ijkl}^*$  verschwinden. Es ist zweckmäßig in diesem Falle statt der Folge (2. 1) von Hyperflächenelementen eine Folge der Einheitsvektoren

$$(3. 12a) \quad x^i = x^i(v),$$

$$(3. 12b) \quad l_i = l_i(x(v))$$

zu wählen, wo  $l_i(x)$  eine Lösung der Gleichung (3. 11) ist.

Die Resultate können wir im folgenden Satz zusammenfassen:

**SATZ III.** *Besteht die Gleichung (3. 11) im Raum identisch, dann kann man immer die Folge der Vektoren (3. 12) in der Weise angeben, daß der Abweichungstensor  $R_{ijkl}^*$  verschwinde.*

(Eingegangen am 4. Januar 1954.)

### Literaturverzeichnis

[1] L. BERWALD, Über die  $n$ -dimensionalen Cartanschen Räume und eine Normalform der zweiten Variation eines  $(n-1)$ -dimensionalen Oberflächenintegrals, *Acta Math.*, **71** (1939), S. 191—248.

[2] ARTHUR MOÓR, Über oskulierende Punkträume von affinzusammenhängenden Linienelementmannigfaltigkeiten, *Annals of Math.*, **56** (1952), S. 397—403.

[3] O. VARGA, Zur Herleitung des invarianten Differentials in Finslerschen Räumen, *Monatshefte für Math. und Phys.*, **50** (1941), S. 165—175.



СОПРИКАСАЮЩИЕСЯ РИМАНОВЫ ПРОСТРАНСТВА РЕГУЛЯРНЫХ  
КАРТАНОВЫХ ПРОСТРАНСТВ

А. МООР (Дебрецен)

(Резюме)

В картановом пространстве через каждую из гиперповерхностей последовательности (2.1) элементов гиперповерхностей располагается гиперплоскость. Если таким образом получается однослойное покрытие некоторой области  $\mathfrak{B}$ , то каждой точке  $x$  области  $\mathfrak{B}$  можно сопоставлять элемент гиперповерхности  $u_j(x)$ . Тензор

$$\gamma_{ik}(x) = g_{ik}(x, u(x))$$

определяет метрический фундаментальный тензор соприкасающегося риманова пространства.

Доказывается, что инвариантный дифференциал картанова пространства и соприкасающегося риманова пространства совпадают вдоль (2.1). Уравнение (3.8) задает связь между тензорами кривизны двух пространств.



# ÜBER GRUPPEN VON AFFINITÄTEN UND BEWEGUNGEN IN FINSLERSCHEN RÄUMEN

Von

GY. SOÓS (Debrecen)

(Vorgelegt von O. VARGA)

## 1. Einleitung

In vorliegender Arbeit definieren wir für eine affinzusammenhängende Mannigfaltigkeit von Linienelementen den Begriff der Affinität. Wir geben ferner — gestützt auf den Begriff der Lieschen Ableitung eines geometrischen Objektes — notwendige und hinreichende Bedingungen für die Existenz einer Gruppe von Affinitäten an. Für den Fall einer allgemeinen affinzusammenhängenden Punktmannigfaltigkeit im Sinne von BERWALD—DOUGLAS ist die entsprechende Fragestellung in mehreren Arbeiten von M. S. KNEBELMANN behandelt worden. Unsere Arbeit stellt aber nicht nur eine Verallgemeinerung der Knebelmannschen Untersuchungen dar, denn wir lösen darüber hinausgehend noch folgendes Problem: Welche sind die notwendigen und hinreichenden Bedingungen dafür, daß eine Gruppe von Affinitäten eines Finslerschen Raumes eine Bewegungsgruppe gestattet?

## 2. Mannigfaltigkeiten von Linienelementen. Transformationsgruppen

Sind  $x^k$  ( $k = 1, \dots, n$ ) die Koordinaten eines Punktes und wird eine Richtung durch diesen Punkt durch das Verhältnis der Parameter  $v^k$  ( $k = 1, \dots, n$ ) charakterisiert, so ist ein Linienelement der  $(2n-1)$ -dimensionalen Mannigfaltigkeit durch  $(x^k, v^k)$  bestimmt. Bei Koordinatentransformationen ändern sich die Bestimmungsstücke des Linienelements nach den Formeln

$$x^{i'} = x^{i'}(x^1, \dots, x^n),$$
$$v^{i'} = \frac{\partial x^{i'}}{\partial x^j} v^j.$$

Es sei nun eine Transformationsschar  $\bar{x}^i = f^i(x^1, \dots, x^n; a^1, \dots, a^r)$ , kurz  $\bar{x}^i = f^i(x; a)$  gegeben, die eine  $r$ -parametrische kontinuierliche Gruppe bildet.



Wir zeigen, daß in diesem Falle die Transformationen

$$(2.1) \quad \bar{x}^i = f^i(x; a), \quad \bar{v}^i = \frac{\partial f^i(x; a)}{\partial x^j} v^j$$

in den  $2n$  Größen  $x^k, v^k$  auch eine  $r$ -parametrische Gruppe bilden. Es genügt die Gruppeneigenschaft für die  $v^k$  nachzuweisen.

Nach Voraussetzung gehört zu je zwei Wertssystemen  $(a_1^1, \dots, a_1^r)$  und  $(a_2^1, \dots, a_2^r)$ , kurz  $(a_1)$  und  $(a_2)$ , ein drittes Wertssystem  $a^3 = \varphi(a_1, a_2)$  der Parameter derart, daß die Gleichungen

$$(2.2) \quad f^i(f(x; a_1); a_2) = f^i(x; \varphi(a_1, a_2))$$

in den  $x$  und  $a$  identisch bestehen.

Wir differenzieren (2.2) nach  $x^k$ :

$$\frac{\partial f^i(\bar{x}; a_2)}{\partial \bar{x}^j} \frac{\partial f^j(x; a_1)}{\partial x^k} = \frac{\partial f^i(x; a_3)}{\partial x^k}$$

und multiplizieren auf beiden Seiten mit  $v^k$ . Man erhält

$$\frac{\partial f^i(\bar{x}; a_2)}{\partial \bar{x}^j} \frac{\partial f^j(x; a_1)}{\partial x^k} v^k = \frac{\partial f^i(x; a_3)}{\partial x^k} v^k$$

oder

$$\bar{v}^i = \frac{\partial f^i(\bar{x}; a_2)}{\partial \bar{x}^j} \bar{v}^j = \frac{\partial f^i(x; a_3)}{\partial x^k} v^k,$$

wodurch die Gruppeneigenschaft nachgewiesen ist.

Ist speziell  $a_0$  das Wertssystem, zu dem die Identität gehört, d. h.

$$x^i = f^i(x; a_0),$$

so sieht man leicht ein, daß dasselbe Wertssystem der Parameter auch die  $v^i$  in sich überführt. Ähnliches ist richtig, wenn es sich um das Wertssystem  $\bar{a}$  handelt, das die inverse Transformation liefert.

Aus der Theorie der Lieschen Gruppen ist es bekannt, daß falls die Schar (2.2) eine Gruppe bildet, die  $\bar{x}^i$  den Differentialgleichungen

$$(2.3) \quad \frac{\partial \bar{x}^i}{\partial a^\alpha} = \frac{\partial f^i(x; a)}{\partial a^\alpha} = \xi_a^i(\bar{x}) A_\alpha^a(a) \quad (a, \alpha = 1, \dots, r)$$

genügen. Ähnliche Differentialgleichungen kann man für die  $\bar{v}^i$  aufstellen. Es ist nämlich

$$\frac{\partial \bar{v}^i}{\partial a^\alpha} = \frac{\partial^2 f^i(x; a)}{\partial x^j \partial a^\alpha} v^j = \frac{\partial}{\partial x^j} \left( \frac{\partial f^i(x; a)}{\partial a^\alpha} \right) v^j.$$

Nach (2.3) ist

$$\frac{\partial \bar{v}^i}{\partial a^\alpha} = \frac{\partial \xi_a^i(\bar{x})}{\partial \bar{x}^s} \frac{\partial \bar{x}^s}{\partial x^j} v^j A_\alpha^a(a) = \frac{\partial \xi_a^i(\bar{x})}{\partial \bar{x}^s} \bar{v}^s A_\alpha^a(a)$$



oder kurz

$$(2.4) \quad \frac{\partial \bar{v}^i}{\partial a^\alpha} = \eta_a^i(\bar{x}, \bar{v}) A_\alpha^a(a), \quad \eta_a^i(\bar{x}, \bar{v}) = \frac{\partial \xi_a^i(\bar{x})}{\partial x^s} \bar{v}^s.$$

Im Falle der Existenz einer Gruppe in den  $x, v$  (erweiterte Gruppe), erweisen sich die Gleichungen (2.3) und (2.4) vollständig integrierbar. Daher gelten außer den Gleichungen

$$\xi_a^j \frac{\partial \xi_b^i}{\partial x^j} - \xi_b^j \frac{\partial \xi_a^i}{\partial x^j} = c_{ab}^e \xi_a^i$$

auch die Gleichungen

$$\xi_a^j \frac{\partial \eta_b^i}{\partial x^j} - \xi_b^j \frac{\partial \eta_a^i}{\partial x^j} + \eta_a^k \frac{\partial \eta_b^i}{\partial v^k} - \eta_b^k \frac{\partial \eta_a^i}{\partial v^k} = c_{ab}^e \eta_a^i$$

mit denselben Konstanten  $c_{ab}^e$ . Diese Konstanten heißen die Zusammensetzungskonstanten der Gruppe.

Das Liesche Symbol der erweiterten Gruppe wird

$$X_a f(x, v) = \xi_a^i \frac{\partial f}{\partial x^i} + \frac{\partial \xi_a^i}{\partial x^j} v^j \frac{\partial f}{\partial v^i} = \xi_a^i \frac{\partial f}{\partial x^i} + \eta_a^i \frac{\partial f}{\partial v^i}.$$

Es ist leicht nachzuweisen, daß der Poissonsche Klammersymbolausdruck der Symbole  $X_a f$  und  $X_b f$  sich linear aus den  $X_e f$  kombinieren läßt:

$$(X_a, X_b) f = c_{ab}^e X_e f.$$

Einen Skalar  $F(x, v)$  nennt man eine absolute Invariante einer  $r$ -parametrischen erweiterten Gruppe, falls jedes Symbol der Gruppe auf die Funktion  $F$  angewandt, verschwindet, d. h.

$$X_a F(x, v) = 0 \quad (a = 1, \dots, r).$$

### 3. Affinzusammenhängende Mannigfaltigkeit von Linienelementen

Die in § 2 betrachtete Mannigfaltigkeit von Linienelementen heißt affinzusammenhängend, wenn für sie Größen  $\Gamma_{jk}^{*i}, C_{jk}^i$  definiert sind, die bei einer Koordinatentransformation dem Transformationsgesetz

$$\bar{\Gamma}_{jk}^{*i} = \Gamma_{bc}^{*a} \frac{\partial \bar{x}^i}{\partial x^a} \frac{\partial x^b}{\partial \bar{x}^j} \cdot \frac{\partial x^c}{\partial \bar{x}^k} + \frac{\partial \bar{x}^i}{\partial x^r} \frac{\partial^2 x^r}{\partial \bar{x}^j \partial \bar{x}^k},$$

$$\bar{C}_{jk}^i = C_{bc}^a \frac{\partial \bar{x}^i}{\partial x^a} \frac{\partial x^b}{\partial \bar{x}^j} \cdot \frac{\partial x^c}{\partial \bar{x}^k}$$

genügen, in den  $v^k$  von  $(-1)$ -ter Ordnung bzw. nullter Ordnung homogen sind, und die Relationen

$$C_{jk}^i v^j = C_{kj}^i v^j = 0, \quad \Gamma_{jk}^{*i} = \Gamma_{kj}^{*i}$$

befriedigen. (Siehe O. VARGA [1].)



Von den Krümmungstensoren des Raumes benötigen wir bloß den sog. Hauptkrümmungstensor:

$$T_{krs}^i = \frac{\partial \Gamma_{kr}^{*i}}{\partial x^s} - \frac{\partial \Gamma_{ks}^{*i}}{\partial x^r} + (\Gamma_{ks;t}^{*i} \Gamma_{rm}^{*t} - \Gamma_{ks;t}^{*i} \Gamma_{sm}^{*t}) v^m + \Gamma_{kr}^{*t} \Gamma_{ts}^{*i} - \Gamma_{ks}^{*t} \Gamma_{tr}^{*i}.$$

(Hier und im folgenden bedeutet das Zeichen „;“ partielle Ableitung nach  $v$ .)  
Dieser Tensor genügt den Relationen:

$$(3.1) \quad T_{jkl}^i + T_{klj}^i + T_{ljk}^i = 0,$$

$$(3.2) \quad T_{jkh|l}^i + T_{jhl|k}^i + T_{jlk|h}^i + (\Gamma_{jk;m}^{*i} T_{shl}^m + \Gamma_{jh;m}^{*i} T_{slk}^m + \Gamma_{jl;m}^{*i} T_{skl}^m) v^s = 0.$$

In unseren Betrachtungen benötigen wir oft die folgende Vertauschungsformel:

$$(3.3) \quad H_{(\alpha)|k|j}^{(\beta)} - H_{(\alpha)|j|k}^{(\beta)} = \sum H_{(\alpha)}^{\beta_1 \dots \beta_{s-1} m \beta_{s+1} \dots \beta_p} T_{mkj}^{\beta_s} - \\ - \sum H_{\alpha_1 \dots \alpha_{s-1} m \alpha_{s+1} \dots \alpha_r} T_{\alpha_s k j}^m - H_{(\alpha);r} T_{mkj}^r v^m,$$

$$(3.4) \quad H_{i^j; m|l} - H_{i^j|l; m} = H_{i^j; r} \Gamma_{sl}^{*r} v^s + H_{r^j} \Gamma_{il}^{*r} v^r - H_i \Gamma_{rl}^{*j} v^r.$$

#### 4. Infinitesimale Transformation. Die Liesche Ableitung eines geometrischen Objektes

Unter einer infinitesimalen Transformation der Mannigfaltigkeit von Linienelementen mit den Zusammenhangsobjekten  $\Gamma_{jk}^{*i}$  und  $C_{jk}^i$  wird die Transformation

$$(4.0) \quad \bar{x}^i = x^i + \xi^i \delta t, \quad (\xi^i = \zeta^i(x)) \\ \bar{v}^i = v^i + \frac{\partial \xi^i}{\partial x^j} v^j \delta t$$

verstanden, wobei  $\delta t$  einen infinitesimalen Parameter bedeutet. Bezüglich dieser infinitesimalen Transformation definieren wir die Liesche Ableitung eines beliebigen geometrischen Objektes, wie üblich, folgenderweise:

$$\Delta \Omega = \lim_{\delta t \rightarrow 0} \frac{\Omega(\bar{x}, \bar{v}) - \overline{\Omega}(x, v)}{\delta t}.$$

Wir stellen das Ergebnis dieses Operators angewandt auf einige Tensoren und die wichtigsten Vertauschungsregeln zusammen:

$$\Delta L(x, v) = \underset{(1)}{\times} L(x, v) = L_{|k} \xi^k + L_{;r} \xi^r v^s,$$

$$\Delta v^i = 0, \quad \Delta \xi^i = 0,$$

$$\Delta \lambda^i = \lambda_{|k}^i \xi^k + \lambda_{;r}^i \xi^r v^s - \lambda^m \xi_{|m}^i,$$

$$\Delta \lambda_i = \lambda_{i|k} + \lambda_{i;r} \xi^r v^s + \lambda_m \xi_{|i}^m.$$



Für einen beliebigen Tensor ist

$$(4.1) \quad \Delta H_{(\alpha)}^{(\beta)} = H_{(\alpha)}^{(\beta)}|_k \xi^k + H_{(\alpha)}^{(\beta)}; r \xi^r v^s - \sum_s H^{\beta_1 \dots \beta_{s-1} m \beta_{s+1} \dots \beta_p} \xi^{\beta_s} + \\ + \sum_s H_{\alpha_1 \dots \alpha_{s-1} m \beta_{s+1} \dots \alpha_q} \xi^m \alpha_s,$$

$$(4.2) \quad \Delta \Gamma_{jk}^{*i} = \xi^i|_j|_k + \Gamma_{jk}; r \xi^r v^s + T_{jkl}^i \xi^l,$$

$$(4.3) \quad \Delta(H_j^i|_k) - (\Delta H_j^i)|_k = H_j^s \Delta \Gamma_{sk}^{*i} - H_s^i \Delta \Gamma_{jk}^{*s} - H_j^i; r v^s \Delta \Gamma_{sk}^{*r},$$

$$(4.4) \quad \Delta(H_j^i; r) - (\Delta H_j^i); r = 0.$$

Der Operator  $\Delta$  ordnet einem beliebigen Tensor einen anderen Tensor vom gleichen Typus zu. Speziell wird  $\Delta \Gamma_{jk}^{*i}$  auch ein Tensor. Der Operator  $\Delta$  und die kovariante Ableitung sind dann und nur dann vertauschbar, falls  $\Delta \Gamma_{jk}^{*i} = 0$  ist.

### 5. Affine Transformationen

Eine infinitesimale Transformation heißt affin, falls sie den Zusammenhang des Raumes unverändert läßt. Dafür ist notwendig und hinreichend, daß bezüglich der infinitesimalen Transformation

$$\bar{x}^i = x^i + \xi^i \delta t, \quad \bar{v}^i = v^i + \frac{\partial \xi^i}{\partial x^j} v^j \delta t$$

für die Zusammenhangsobjekte die Relationen

$$\bar{\Gamma}_{jk}^{*i}(\bar{x}, \bar{v}) = \Gamma_{jk}^{*i}(\bar{x}, \bar{v}), \quad \bar{C}_{jk}^i(\bar{x}, \bar{v}) = C_{jk}^i(\bar{x}, \bar{v})$$

bestehen. Benützen wir Liesche Ableitungen, so bedeutet dies

$$(5.1) \quad \Delta \Gamma_{jk}^{*i} = 0, \quad (5.2) \quad \Delta C_{jk}^i = 0.$$

In ausführlicherer Form:

$$\Delta \Gamma_{jk}^{*i} = \xi^i|_j|_k + \Gamma_{jk}; r \xi^r v^s + T_{jkl}^i \xi^l = 0, \\ \Delta C_{jk}^i = C_{jk}^i|_l \xi^l + C_{jk}; r \xi^r v^s - C_{jk}^m \xi^i_m + C_{mk}^i \xi^m_j + C_{jm}^i \xi^m_k = 0.$$

Durch geeignete Koordinatentransformation können wir erreichen, daß  $\xi^i = \delta_1^i$ . Dann reduzieren sich die Gleichungen (5.1) und (5.2) auf

$$(5.3) \quad \frac{\partial \Gamma_{jk}^{*i}}{\partial x^1} = 0, \quad (5.4) \quad \frac{\partial C_{jk}^i}{\partial x^1} = 0.$$

Sind diese Bedingungen erfüllt, so gestattet der Raum von Linien-elementen eine endliche Gruppe von Affintransformationen der Form:

$$(5.5) \quad \bar{x}^i = x^i + \delta_1^i t, \quad \bar{v}^i = v^i.$$

DEFINITION. Unter einer *affinen Transformationsgruppe der Linien-elementmannigfaltigkeit* verstehen wir eine Transformationsgruppe von der Eigenschaft, daß für jede ihrer Transformation  $\bar{\Gamma}_{jk}^{*i}(\bar{x}, \bar{v}) = \Gamma_{jk}^{*i}(\bar{x}, \bar{v})$  und  $\bar{C}_{jk}^i(\bar{x}, \bar{v}) = C_{jk}^i(\bar{x}, \bar{v})$  ist.



Aus (5.3), (5.4), (5.5) und der vorangehenden Definition folgen:

**SATZ 1.** *Ist eine infinitesimale affine Transformation der Mannigfaltigkeit vorhanden, so existiert eine endliche einparametrische Gruppe von affinen Transformationen.*

**SATZ 2.** *Damit die Mannigfaltigkeit eine einparametrische affine Transformationsgruppe gestattet, ist es notwendig und hinreichend, daß es ein Koordinatensystem gebe, in welchem die Zusammenhangsobjekte  $\Gamma_{jk}^{*i}$  und  $C_{jk}^i$  von einer der Koordinaten nicht abhängen.*

## 6. Die Integrabilitätsbedingungen

In diesem § werden wir die Integrabilitätsbedingungen der Gleichungen

$$(6.1) \quad \Delta \Gamma_{jk}^{*i} = 0, \quad (6.2) \quad \Delta C_{jk}^i = 0$$

aufstellen. Erstens bestimmen wir diejenigen Bedingungen, die aus (6.1) durch kovariante bzw. partielle Ableitung entstehen. Zu diesem Zwecke schreiben wir (6.1) in der Form:

$$\xi_{|j|k}^i = -\Gamma_{jk;r}^{*i} \xi^r v^s - T_{jkr}^i \xi^r.$$

Wir bilden auf beiden Seiten die kovariante Ableitung nach  $x^l$ , dann vertauschen wir die Indizes  $k$  und  $l$ , und nach Subtraktion der entstehenden Gleichungen erhalten wir den Ausdruck:

$$\begin{aligned} \xi_{|j|k|l}^i - \xi_{|j|l|k}^i &= (T_{jlr|k}^i - T_{jkr|l}^i) \xi^r - (\Gamma_{jk;m|l}^{*i} - \Gamma_{jl;m|k}^{*i}) \xi^m v^s + \\ &+ \Gamma_{jl;m}^{*i} \xi^m v^s - \Gamma_{j;m}^{*i} \xi^m v^s + T_{jlr}^i \xi^r - T_{jkr}^i \xi^r. \end{aligned}$$

Nach (3.3) ist

$$\xi_{|j|k|l}^i - \xi_{|j|l|k}^i = \xi_{|j}^r T_{rkl}^i - \xi_{|r}^i T_{jkl}^r - \xi_{|j}^i T_{skl}^r v^s.$$

Durch Vergleichen der beiden Ausdrücke und Anwendung der verallgemeinerten Bianchischen Identitäten (3.2) erhalten wir, die Formel

$$(6.3) \quad T_{jkl|r}^i = \Gamma_{jk;m|l}^{*i} - \Gamma_{jl;m|k}^{*i} + (\Gamma_{jl;r}^{*i} T_{sk}^{*r} - \Gamma_{jk;r}^{*i} T_{sl}^{*r} v^s)$$

in Betracht nehmend, den Ausdruck:

$$T_{jkl|r}^i \xi^r + T_{jkl;m}^i \xi^m v^s - T_{jkl}^r \xi_r^i + T_{rkl}^i \xi_j^r + T_{jrl}^i \xi_k^r + T_{jkr}^i \xi_l^r = 0,$$

oder kurz geschrieben:

$$(6.4) \quad \Delta T_{jkl}^i = 0.$$

Wiederholte kovariante Ableitung auf Grund von (6.1) führt zum Gleichungssystem:

$$(6.5) \quad \Delta T_{jkl|m_1|m_2|\dots|m_s}^i = 0 \quad (s = 1, 2, \dots).$$

Eine andere Gruppe von Bedingungen entsteht durch partielle Ableitung von (6.1). Nach (4.4) ist  $\Delta(\Gamma_{jk}^{*i}; l) = 0$ , und wir bekommen durch Wieder-



holung der partiellen Ableitung das System

$$(6.6) \quad \Delta \Gamma_{jk;l_1;l_2;\dots;l_t}^{*i} = 0 \quad (t = 1, 2, \dots).$$

Aus diesen Gleichungen ist es hinreichend nur die letzte zu betrachten. Es steht nämlich auf der linken Seite von (6.6) eine homogene Funktion  $(-t)$ -ter Ordnung in der  $v^i$ . Die Anwendung der Eulerschen Relationen führt zur Formel:

$$v^s \Delta \Gamma_{jk;l_1;\dots;l_{t-1};s}^{*i} = -t \Delta \Gamma_{jk;l_1;\dots;l_{t-1}}^{*i}.$$

Das Verschwinden von  $\Delta \Gamma_{jk;l_1;\dots;l_t}^{*i}$  ist daher genügend zum Verschwinden aller Gleichungen für kleinere  $t$ . Wir schreiben daher die Gleichungen (6.6) in der Form:

$$(6.7) \quad \Delta \Gamma_{jk;l_1;\dots;l_t}^{*i} = 0.$$

Wir bilden nun die sukzessiven kovarianten Ableitungen von (6.7):

$$(6.8) \quad \Delta \Gamma_{jk;l_1;\dots;l_t|s_1|\dots|s_r}^{*i} = 0 \quad (r = 1, 2, \dots).$$

Die Benützung der Vertauschungsformel (3.4) zeigt, daß weitere partielle Ableitungen von (6.8) keine neueren Bedingungen liefern.

Im folgenden suchen wir die Bedingungen, die aus (6.4) und (6.5) durch partielle Ableitung hervorgehen. Es ist nach (6.3)

$$(6.9) \quad T_{jkl;m}^i = \Gamma_{jk;m|l}^{*i} - \Gamma_{jl;m|k}^{*i} + (\Gamma_{jl;r}^{*i} \Gamma_{sk;m}^{*r} - \Gamma_{jk;r}^{*i} \Gamma_{sl;m}^{*r}) v^s.$$

Wir wenden den Operator  $\Delta$  auf beiden Seiten an. Nach (4.4) steht auf der linken Seite  $(\Delta T_{jkl;m}^i)$ . Diese Größe läßt sich durch die linken Seiten der Gleichungen (6.1), (6.7) und (6.8) ausdrücken, d. h. man bekommt keine weiteren Bedingungen durch partielle Ableitung von (6.4).

Betrachten wir nun die Gleichungen (6.5) für  $s = 1$ . Erstens benützen wir die Vertauschungsformel (3.4) für den Tensor  $T_{jkl}^i$ , und dann wenden wir auf die erhaltene Gleichung den Operator  $\Delta$  an. Es ist leicht zu sehen, daß die Größe  $(\Delta T_{jkl|m}^i)_{;n}$  sich mit Hilfe von a) (6.1) und (6.7), b)  $\Delta T_{jkl;m}^i$  und c)  $(\Delta T_{jkl}^i)_{;n|m}$  ausdrücken läßt. Den Fall b) haben wir schon betrachtet. Den Fall c) betrachtend, wenden wir den Operator  $\Delta$  auf (6.9) an, und bilden die kovariante Ableitung nach  $x^m$ . Die erhaltene Größe können wir mit Hilfe von (6.1), (6.7) und (6.6) ausdrücken. Man bekommt daher keine weitere Bedingung mittels partieller Ableitung aus (6.5). Denselben Gedankengang kann man für jeden Wert von  $s$  anwenden.

Fassen wir die erhaltenen Integrabilitätsbedingungen zusammen:

$$(6.10) \quad \begin{aligned} \Delta \Gamma_{jk;l_1;\dots;l_t}^{*j} &= 0, \\ \Delta \Gamma_{jk;l_1;\dots;l_t|s_1|\dots|s_r}^{*i} &= 0 \quad (r = 1, 2, \dots), \\ \Delta T_{jkl}^i &= 0, \\ \Delta T_{jkl|m_1|\dots|m_s}^i &= 0 \quad (s = 1, 2, \dots). \end{aligned}$$



Durch ähnliches Verfahren lassen sich diejenigen Bedingungen zusammenstellen, die aus (6.2) mittels wiederholter kovarianter bzw. partieller Ableitungen hervorgehen. Diese Bedingungen sind:

$$(6.11) \quad \begin{aligned} \Delta C_{jk;l_1; \dots; l_t}^i &= 0, \\ \Delta C_{jk|m_1| \dots | m_s}^i &= 0 & (s = 1, 2, \dots), \\ \Delta C_{jk;l_1 \dots; l_t | m_1 | \dots | m_r}^i &= 0 & (r = 1, 2, \dots). \end{aligned}$$

Wir haben den

**SATZ 3.** *Eine notwendige und hinreichende Bedingung dafür, daß eine affinzusammenhängende Mannigfaltigkeit von Linienelementen eine affine infinitesimale Transformation und daher — Satz 1 zufolge — mindestens eine einparametrische affine Gruppe gestattet, besteht darin, daß es eine positive ganze Zahl  $N$  gebe derart, daß die ersten  $N$  Gruppen von Gleichungen der Gleichungskette (6.10), (6.11) in den Unbekannten  $\xi^i$  und  $\frac{\partial \xi^i}{\partial x^j}$  algebraisch verträglich sind und die Lösungen  $\xi^i$  und  $\frac{\partial \xi^i}{\partial x^j}$  der ersten  $N$  Gruppen auch der  $(N+1)$ -ten Gruppe genügen.*

## 7. Die Bewegungsgruppe als Untergruppe der Gruppen von affinen Transformationen

Es sei ein Finslerscher Raum vorgelegt. Man sagt, daß die infinitesimale Transformation (4.0) eine infinitesimale Bewegung ist, falls

$$(7.1) \quad \Delta g_{ij} = 0.$$

In ausführlicherer Form erhalten diese verallgemeinerten Killingschen Gleichungen die Gestalt:

$$g_{im} \xi_{|j}^m + g_{mj} \xi_{|i}^m + g_{ij;m} \xi_{|s}^m v^s = \xi_{i|j} + \xi_{j|i} + 2C_{ijm} \xi_{|s}^m v^s = 0,$$

da jetzt  $g_{ij|k} = 0$  und  $C_{ijm} = \frac{1}{2} g_{ij;m}$  ist. Die Integrabilitätsbedingungen von (7.1) sind in den Gleichungen (6.10), (6.11) enthalten.

Wir betrachten eine  $r$ -parametrische Gruppe  $G_r$  von affinen Transformationen. Wir bezeichnen mit  $\xi_{(\sigma)}^i$  ( $\sigma = 1, \dots, r$ ) ein linear unabhängiges Lösungssystem der Gleichungen (6.1), (6.2), (6.10), (6.11). Wird das in Bezug  $\xi_{(\sigma)}^i$  gebildete Liesche Ableitungssymbol  $\Delta$  auf den Tensor  $g_{ij}$  angewandt, so erhalten wir Tensoren, die mit  $h_{ij}^{(\sigma)}$  bezeichnet werden sollen.

Es wurde nun festgestellt, ob es ein System von Konstanten  $c^{(\sigma)}$  so bestimmbar ist, daß die infinitesimale Transformation

$$\bar{x}^i = x^i + c^{(\sigma)} \xi_{(\sigma)}^i \delta t, \quad \bar{v}^i = v^i + c^{(\sigma)} \frac{\partial \xi_{(\sigma)}^i}{\partial x^j} v^j \delta t$$



eine infinitesimale Bewegung darstellt, mit a. W., daß sich die Gleichungen

$$(7.2) \quad c^{(\sigma)}_{(\sigma)} h_{ij} = 0$$

erfüllen.

Für die Existenz eines nichttrivialen Lösungssystems ist es eine notwendige Bedingung, daß der Rang der Matrix  $\|h_{ij}\|_{(\sigma)}$  kleiner als  $r$  ist.

Wir behaupten, daß diese Bedingung für die Existenz eines Systems von Konstanten  $c^{(\sigma)}$  auch hinreichend ist.

Ist nämlich der Rang der Matrix  $\|h_{ij}\|_{(\sigma)}$  gleich  $s$  ( $s < r$ ), dann gibt es  $r-s$  linear unabhängige Funktionen  $\varphi^{(\sigma)}_{(\alpha)}(x, v)$  ( $\alpha = 1, 2, \dots, r-s$ ) derart, daß die Gleichungen

$$(7.3) \quad \varphi^{(\sigma)}_{(\alpha)} h_{ij} = 0$$

Identitäten in den Variablen  $x$  und  $v$  sind. Daraus ergeben sich mittels kovarianter Ableitung die Gleichungen

$$(7.4) \quad \varphi^{(\sigma)}_{(\alpha)|k} h_{ij} + \varphi^{(\sigma)}_{(\alpha)} h_{ij|k} = 0.$$

Nun ist  $h_{ij|k} = 0$ . Es ist nämlich

$$\mathcal{A}(g_{ij|k})_{(\sigma)} - (\mathcal{A}g_{ij})_{|k} = - (g_{ij|m} v^s \mathcal{A}I^*_{sk} + g_{rj} \mathcal{A}I^*_{ik} + g_{ir} \mathcal{A}I^*_{jk}).$$

Da die Vektoren  $\xi^i_{(\sigma)}$  Lösungen des Systems (6.1), (6.2), (6.10), (6.11) sind, verschwinden die Ableitungen  $\mathcal{A}I^*_{jk}$ . Unter Beachtung, daß  $g_{ij|k} = 0$  und  $\mathcal{A}g_{ij} = h_{ij}$  ist, folgt, daß  $h_{ij|k} = 0$  besteht. Setzen wir das Ergebnis in (7.4) ein, so folgt

$$(7.5) \quad \varphi^{(\sigma)}_{(\alpha)|k} h_{ij} = 0,$$

d. h. auch die Funktionen  $\varphi^{(\sigma)}_{(\alpha)|k}$  sind Lösungen von (7.2). Dann existieren Funktionen  $L^{\beta}_{\gamma\delta}$  derart, daß

$$(7.6) \quad \varphi^{(\sigma)}_{(\alpha)|k} = L^{\beta}_{\alpha\gamma} \varphi^{(\sigma)}_{(\beta)}$$

Differenziert man (7.3) partiell nach  $v^l$ , so entstehen die Gleichungen

$$\varphi^{(\sigma)}_{(\alpha);l} h_{ij} + \varphi^{(\sigma)}_{(\alpha)} h_{ij;l} = 0.$$

Es ist  $\mathcal{A}C^i_{jl} = 0$ . Es ist aber

$$h_{ij;l} = 2\mathcal{A}(g_{mi} C^m_{jl}) = 2(C^m_{jl} h_{mi} + g_{mi} \mathcal{A}C^m_{jl}) = 2C^m_{jl} h_{mi}.$$

Nach Einsetzung ist

$$\varphi^{(\sigma)}_{(\alpha);l} + 2\varphi^{(\sigma)}_{(\alpha)} C^m_{jl} h_{mi} = 0$$



oder

$$(\varphi_{(\alpha);l}^{(\sigma)} \delta_j^m + 2\varphi_{(\alpha)}^{(\sigma)} C_{jl}^m) h_{mi} = 0.$$

Die in Klammern stehenden Funktionen sind ebenfalls Lösungen von (7.2). Sie lassen sich daher folgenderweise darstellen:

$$\varphi_{(\alpha);l}^{(\sigma)} \delta_j^m + 2\varphi_{(\alpha)}^{(\sigma)} C_{jl}^m = A_{\alpha l j}^{\beta m} \varphi_{(\beta)}^{(\sigma)}.$$

Setzt man  $l = m$  und summiert über diese Indizes, so bekommt man

$$\varphi_{(\alpha);j}^{(\sigma)} + 2\varphi_{(\alpha)}^{(\sigma)} C_{jm}^m = A_{\alpha m j}^{\beta m} \varphi_{(\beta)}^{(\sigma)}$$

oder

$$(7.7) \quad \varphi_{(\alpha);j}^{(\sigma)} = A_{\alpha j}^{\beta} \varphi_{(\beta)}^{(\sigma)} \quad (A_{\alpha j}^{\beta} = A_{\alpha m j}^{\beta m} - 2C_{jm}^m \delta_{\alpha}^{\beta}).$$

Damit ein aus lauter Konstanten bestehendes Lösungssystem  $c^{(\sigma)}$  existiere, muß es Funktionen  $\lambda^{(\alpha)}$  ( $\alpha = 1, \dots, r-s$ ) mit der Eigenschaft

$$(7.8) \quad c^{(\sigma)} = \lambda^{(\alpha)} \varphi_{(\alpha)}^{(\sigma)}$$

geben. Wir bilden die kovariante Ableitung von (7.8):

$$\lambda_{|k}^{(\beta)} \varphi_{(\beta)}^{(\sigma)} + \lambda^{(\alpha)} L_{\alpha k}^{\beta} \varphi_{(\beta)}^{(\sigma)} = (\lambda_{|k}^{(\beta)} + \lambda^{(\alpha)} L_{\alpha k}^{\beta}) \varphi_{(\beta)}^{(\sigma)} = 0,$$

aus der wegen der linearen Unabhängigkeit der Funktionen  $\varphi_{(\beta)}^{(\sigma)}$  die Gleichungen

$$(7.9) \quad \lambda_{|k}^{(\beta)} + \lambda^{(\alpha)} L_{\alpha k}^{\beta} = 0$$

folgen. Durch partielle Ableitung von (7.8) unter Beachtung von (7.7) gelangt man zu einem analogen Gleichungssystem:

$$(7.10) \quad \lambda_{|k}^{(\beta)} + \lambda^{(\alpha)} A_{\alpha k}^{\beta} = 0.$$

Die Gleichungssysteme (7.9) und (7.10) für die unbekanntenen Funktionen  $\lambda^{(\alpha)}$  sind vollständig integrierbar. Um dies einzusehen, stellen wir die Integrabilitätsbedingungen der Gleichungen (7.9) und (7.10) zusammen.

Es folgt aus (7.9)

$$\lambda_{|k}^{(\beta)} |l = -\lambda_{|l}^{(\alpha)} L_{\alpha k}^{\beta} - \lambda^{\alpha} L_{\alpha k |l}^{\beta} = (L_{\alpha l}^{\alpha} L_{\alpha k}^{\beta} - L_{\alpha k |l}^{\beta}) \lambda^{(\alpha)}.$$

Vertauschen wir jetzt die Indizes  $k$  und  $l$  und subtrahieren die entstehende Gleichung aus der vorangehenden:

$$(7.11) \quad \lambda_{|k}^{(\beta)} |l - \lambda_{|l}^{(\beta)} |k = (L_{\alpha l}^{\alpha} L_{\alpha k}^{\beta} - L_{\alpha k |l}^{\beta} - L_{\alpha k}^{\alpha} L_{\alpha l}^{\beta} + L_{\alpha l |k}^{\beta}) \lambda^{(\alpha)}.$$

Mit Beachtung von (3.3) ist

$$(7.12) \quad \lambda_{|k}^{(\beta)} |l - \lambda_{|l}^{(\beta)} |k = -\lambda_{;r}^{(\beta)} T_{mkl}^r v^m = A_{qr}^{\beta} T_{mkl}^r v^m \lambda^{(q)}.$$

Durch Vergleichen von (7.11) und (7.12) bekommen wir die gesuchten Bedingungen

$$(7.13) \quad (L_{\alpha k |l}^{\beta} - L_{\alpha l |k}^{\beta} + L_{\alpha k}^{\alpha} L_{\alpha l}^{\beta} - L_{\alpha l}^{\alpha} L_{\alpha k}^{\beta} + A_{qr}^{\beta} T_{mkl}^r v^m) \lambda^{(q)} = 0.$$

Aus (7.10) erhalten wir durch analoges Verfahren mittels partieller Ableitung

$$(7.14) \quad (A_{\alpha k |l}^{\beta} - A_{\alpha l |k}^{\beta} + A_{\alpha k}^{\alpha} A_{\alpha l}^{\beta} - A_{\alpha l}^{\alpha} A_{\alpha k}^{\beta}) \lambda^{(q)} = 0.$$



Die Gleichungen (7. 13) und (7. 14) gelten identisch. Es entstehen nämlich aus (7. 5) durch kovariante Ableitung und Vertauschung der Indizes  $k$  und  $l$  nach Subtraktion die Gleichungen

$$(L_{\alpha k ; l}^{\gamma} - L_{\alpha l ; k}^{\gamma} + L_{\alpha k}^{\beta} L_{\beta l}^{\gamma} - L_{\alpha l}^{\beta} L_{\beta k}^{\gamma} + A_{\alpha r}^{\gamma} T_{m k l}^r v^m) q_{(\gamma)}^{(\sigma)} = 0.$$

Wegen der linearen Unabhängigkeit der Funktionen  $q_{(\gamma)}^{(\sigma)}$  folgt

$$(7. 15) \quad L_{\alpha k ; l}^{\gamma} - L_{\alpha l ; k}^{\gamma} - L_{\alpha k}^{\beta} L_{\beta l}^{\gamma} + L_{\alpha l}^{\beta} L_{\beta k}^{\gamma} + A_{\alpha r}^{\gamma} T_{m k l}^r v^m = 0.$$

Analogerweise folgt aus (7. 7) die Gleichung

$$(A_{\alpha k ; l}^{\gamma} - A_{\alpha l ; k}^{\gamma} + A_{\alpha k}^{\beta} A_{\beta l}^{\gamma} - A_{\alpha l}^{\beta} A_{\beta k}^{\gamma}) q_{(\gamma)}^{(\sigma)} = 0,$$

und daher

$$(7. 16) \quad A_{\alpha k ; l}^{\gamma} - A_{\alpha l ; k}^{\gamma} + A_{\alpha k}^{\beta} A_{\beta l}^{\gamma} - A_{\alpha l}^{\beta} A_{\beta k}^{\gamma} = 0.$$

Die Ergebnisse (7. 15) und (7. 16) zeigen, daß die Gleichungen (7. 9) und (7. 10) vollständig integrierbar sind. Es existieren also Funktionen  $z^{(\sigma)}$  und daher auch die Konstanten  $c^{(\sigma)}$ .

SATZ 4. Ist  $G_r$  die Gruppe von affinen Transformationen eines Finslerschen Raumes, so existiert immer eine Bewegungsgruppe von  $G_r$ , falls der Rang der Matrix  $\|h_{ij}^{(\sigma)}\|$  kleiner als  $r$  ist.

Herrn Prof. O. VARGA spreche ich für seine wertvollen Bemerkungen den besten Dank aus.

(Eingegangen am 4. Januar 1954.)

### Literaturverzeichnis

M. S. KNEBELMAN, Collineations and motions in generalized spaces, *Amer. Journal of Math.*, **51** (1929), pp. 527—564.  
 O. VARGA [1], Über affinzusammenhängende Mannigfaltigkeiten von Linienelementen insbesondere deren Äquivalenz, *Publ. Math., Debrecen*, **1** (1949), pp. 7—17.  
 O. VARGA [2], Beiträge zur Theorie der Finslerschen Räume und der affinzusammenhängenden Räume von Linienelementen (Unveröffentlichte Dissertation, Prag, 1934).



ОБ АФФИНИТЕТАХ И ДВИЖЕНИЯХ В ФИНСЛЕРОВЫХ  
ПРОСТРАНСТВАХ

Д. ШООШ (Дебрецен)

(Резюме)

В настоящей работе определяется понятие аффинитета в пространстве линейных элементов с аффинной связностью. В сформулировании основных определений и теорем мы пользуемся дифференциалом Ли различных геометрических объектов. Данное инфинитезимальное преобразование (§ 4) линейных элементов называется инфинитезимальным аффинитетом, если

$$\Delta \Gamma_{jk}^{*i} = 0, \quad \Delta C_{jk}^i = 0,$$

обозначая знаком  $\Delta$  символ дифференциала Ли, образованного относительно инфинитезимального преобразования.

В работе задаются условия, необходимые и достаточные для того, чтобы данное инфинитезимальное преобразование представляло собой инфинитезимальный аффинитет. (См. Теорему 3.)

Для случая финслеровых пространств решается следующая проблема. Пусть дана непрерывная группа  $G_r$  ( $r$ —число параметров) аффинных преобразований финслерова пространства. При каких условиях существует подгруппа аффинной группы, являющаяся группой движений финслерова пространства? Ответ на этот вопрос содержится в Теореме 4.



# ÜBER DIE LÖSUNG DER IM BANACHSCHEN RAUME DEFINIERTEN NICHTLINEAREN GLEICHUNGEN

Von

I. FENYŐ (Budapest)

(Vorgelegt von A. RÉNYI)

1. Es sei  $f(x)$  ein Operator, der den Banachschen Raum  $X$  in den Banachschen Raum  $Y$  abbildet. Betrachten wir die Gleichung

$$(1) \quad f(x) = 0.$$

Es ist bekannt, daß unter gewissen Umständen die durch L. W. KANTOROWITSCH und seinen Mitarbeiter [1, 2] in den Banachschen Raum übertragene Newtonsche und modifizierte Newtonsche Iterationsmethoden zur Lösung der Gleichung (1) geeignet sind. Ist  $f(x)$  ein im üblichen Sinne (GATEAUX, FRÉCHET, WAINBERG) differenzierbarer Operator, so heißt bekanntlich der Algorithmus

$$(2) \quad x_n = x_{n-1} - f'(x_{n-1})^{-1} f(x_{n-1}) \quad (n = 1, 2, \dots)$$

die Newtonsche Iterationsmethode und der Algorithmus

$$(2') \quad \xi_n = \xi_{n-1} - f'(\xi_0)^{-1} f(\xi_{n-1})$$

die modifizierte Newtonsche Methode. Hier bedeutet  $f'(x)^{-1}$  wie üblich den inversen Operator von  $f'(x)$ .

L. W. KANTOROWITSCH [1] und J. P. MISSOWSKICH [2] haben hinreichende Bedingungen für die Konvergenz von (2) und (2') gegeben. Sie haben nämlich bewiesen, daß im Falle der Konvergenz die Elementenfolgen  $\{x_n\}$  bzw.  $\{\xi_n\}$  gegen eine Lösung der Gleichung (1) konvergieren. In den zitierten Abhandlungen wird vorausgesetzt, daß  $f(x)$  im Fréchet'schen Sinne zweimal differenzierbar und seine zweite Ableitung beschränkt ist. Es scheint ganz natürlich, solche Bedingungen zu suchen, welche die Existenz der zweiten Ableitung von  $f(x)$  nicht voraussetzen, da die Verfahren (2) und (2') keinen Gebrauch von der zweiten Ableitung machen. Verf. hat in seiner ungarisch erschienenen Arbeit [3] die im Banachschen Raume definierten Gleichungen von der Form

$$f(x) = y \text{ und } F(x, y) = 0$$

untersucht und unter andern für die Konvergenz des Algorithmus (2') eine genügende Bedingung gegeben. Diese Bedingung setzt nicht die Existenz der



zweiten Derivierten voraus. Neulich ist ein Aufsatz von M. L. STEIN erschienen [4], in dem genügende Bedingungen zur Konvergenz von (2) aufgestellt werden, ohne die Existenz der zweiten Ableitung zu postulieren. M. L. STEIN benützt aber die Ableitung im Sinne von GATEAUX [5], und es ist nach einem Satze von ZORN [6] bekannt, daß ein Operator — falls er in einem Gebiet im Sinne von GATEAUX differenzierbar ist — in diesem Gebiete auch analytisch ist. Das heißt: die Postulierung der Differenzierbarkeit im Sinne von GATEAUX ist eine sehr strenge und weitgehende Forderung. Diese Forderung wurde von STEIN in seinem Beweise stark ausgenützt.

Im folgenden geben wir Bedingungen für die Konvergenz der Prozesse (2) bzw. (2'), ohne Voraussetzung der Existenz der zweiten Ableitung *im Fréchet'schen Sinne*. Es scheint uns, daß unsere Beweise einfacher sind, als der Beweis von M. L. STEIN. Wir wiederholen unser in ungarischer Sprache publiziertes Ergebnis über die Konvergenz des Prozesses (2') und untersuchen mit derselben Idee die Konvergenz von (2).

Es ist bekannt, daß ein Operator, der den Banachschen Raum  $X$  in den Banachschen Raum  $Y$  abbildet, im *Fréchet'schen Sinne* an der Stelle  $x_0$  ( $x_0 \in X$ ) differenzierbar ist, falls ein linearer Operator  $U$  existiert, der den Raum  $X$  in  $Y$  abbildet, so daß die folgende Relation gültig sei:

$$\|f(x+h) - f(x) - Uh\| = o(\|h\|)$$

( $h$  ist ein Element aus  $X$ ). Dieser lineare Operator  $U$  heißt die Ableitung von  $f(x)$  an der Stelle  $x_0$  und wird mit  $f'(x_0)$  bezeichnet [9].

Ein Operator  $f(x)$  wird im Sinne von GATEAUX differenzierbar an der Stelle  $x_0$  genannt, falls für jedes Element  $h \in X$  der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(x_0 + th) - f(x_0)}{t} = V(x_0, h)$$

existiert und  $V(x_0, h)$  ein linearer Operator bezüglich des Elementes  $h$  ist [10].

2. Es sei die folgende Gleichung gegeben:

$$(3) \quad f(x) = y.$$

Im folgenden wird die Bezeichnung  $F(x) = f(x) - y$  oft benützt.

SATZ 1. *Es sei  $f(x)$  ein Operator, der den Banachschen Raum  $X$  in den Banachschen Raum  $Y$  überführt. Falls  $f(x)$  in einer Umgebung des Elementes  $x_0 \in X$  folgende Eigenschaften besitzt:*

1°  $f(x)$  ist im *Fréchet'schen Sinne* differenzierbar und der inverse Operator  $f'(x_0)^{-1}$  von  $f'(x_0)$  existiert und es gilt

$$(4) \quad \|f'(x_0)^{-1}\| \leq B;$$

2° es gilt

$$(5) \quad \|f'(x) - f'(x_0)\| \leq C \|x - x_0\|, \text{ falls } \|x - x_0\| \leq r < \frac{1}{3BC};$$



dann ist für jedes, der Ungleichung

$$(6) \quad \|y - f(x_0)\| \leq \eta = \frac{2 - 3BCr}{11B} r$$

genügende Element  $y \in Y$  die Gleichung (3) in der  $r$ -Umgebung von  $x_0$  eindeutig lösbar. Die Lösung wird als Grenzwert derjenigen Folge  $\{x_n\}$  dargestellt, die durch die Gleichung

$$(7) \quad x_n = x_{n-1} - (F'(x_{n-1}))^{-1} F(x_{n-1}) \quad (n = 1, 2, \dots)$$

definiert ist.

BEWEIS. Der Operator  $F(x)$  besitzt ebenfalls die Eigenschaften 1° und 2°. Nun zeigen wir, daß  $F'(x)^{-1}$  in der ganzen  $r$ -Umgebung von  $x_0$  existiert [1, p. 118]. Bezeichnet  $\mathcal{E}$  den Identitätsoperator, so existiert der inverse Operator  $T(x)$  von

$$\mathcal{E} - f'(x_0)^{-1} [f'(x_0) - f'(x)]$$

wegen (5) und eines klassischen Satzes von S. BANACH [7], und es gilt nach Betrachtung von (5)

$$\|T(x)\| \leq \frac{1}{1 - BCr} < \frac{3}{2}.$$

Da aber

$$(8) \quad T(x) F'(x_0)^{-1} = F'(x)^{-1}$$

gilt, haben wir die Existenz von  $f'(x)^{-1}$  bewiesen.

Aus (8) folgt

$$(9) \quad \|F'(x)^{-1}\| \leq \frac{B}{1 - BCr} = \frac{3}{2} B \quad (\|x - x_0\| \leq r).$$

Nun definieren wir den Operator

$$\tau(x) = F(x) - F'(x)(x - x_0).$$

Es ist evident, daß Gleichung (3) mit der Gleichung

$$(10) \quad x = x_0 - F'(x)^{-1} \tau(x)$$

identisch ist. Diese lösen wir durch die Methode der sukzessiven Approximation:

$$(11) \quad x_n = x_0 - F'(x_{n-1})^{-1} \tau(x_{n-1}) \quad (n = 1, 2, \dots),$$

und beweisen, daß  $\lim_{n \rightarrow \infty} x_n = x^*$  existiert.

Es ist

$$(12) \quad \|x_1 - x_0\| \leq \|F'(x_0)^{-1}\| \cdot \|F(x_0)\| \leq B\eta < r.$$

Wir behaupten, daß die Ungleichung

$$(13) \quad \|x_n - x_{n-1}\| \leq \frac{3}{2} B \frac{3\eta + Cr^2}{r} \|x_{n-1} - x_{n-2}\|$$



für alle ganzen  $n (\geq 2)$  gültig ist. Sie gilt für  $n=2$ , denn es besteht die Identität:

$$x_2 - x_1 = F'(x_0)^{-1} \tau(x_0) - F'(x_1)^{-1} \tau(x_1) = F'(x_0)^{-1} [\tau(x_0) - \tau(x_1)] + \\ + [F'(x_0)^{-1} - F'(x_1)^{-1}] \tau(x_1).$$

Wegen (11) liegt das Element  $x_1$  in der  $r$ -Umgebung von  $x_0$ , weshalb die folgenden Abschätzungen berechtigt sind:

$$\|F'(x_0)^{-1}\| \leq B < \frac{3}{2} B,$$

$$\|\tau(x_1)\| \leq \|F(x_1)\| + \|F'(x_1)\| \cdot \|x_1 - x_0\| \leq \eta + \frac{\eta}{r} r = 2\eta,$$

$$\|F'(x_0)^{-1} - F'(x_1)^{-1}\| = \|F'(x_1)^{-1} [F'(x_0) - F'(x_1)] F'(x_0)^{-1}\| \leq \frac{9}{4} B^2 C \|x_1 - x_0\|.$$

Endlich kann man nach dem Analogon des Lagrangeschen Mittelwertsatzes [8] behaupten:

$$\|\tau(x_0) - \tau(x_1)\| \leq \|F(x_0) - F(x_1)\| + \|F'(x_0)(x_0 - x_0) - F'(x_1)(x_1 - x_0)\| \leq \\ \leq \|F'(x_1)\| \cdot \|x_1 - x_0\| + \|[F'(x_0) - F'(x_1)](x_0 - x_0) + F'(x_1)(x_0 - x_1)\| \leq \\ \leq \frac{\eta}{r} \|x_1 - x_0\| + Cr \|x_1 - x_0\| + \frac{\eta}{r} \|x_1 - x_0\| = \left(\frac{2\eta}{r} + Cr\right) \|x_1 - x_0\|.$$

Hier ist  $\bar{x}_1 = x_0 + \theta_1(x_1 - x_0)$  ( $0 < \theta_1 < 1$ ), also ein zu der  $r$ -Umgebung von  $x_0$  gehörendes Element. Nun haben wir wegen (5)

$$\|x_2 - x_1\| \leq \frac{3}{2} B \frac{3\eta + Cr^2}{r} \|x_1 - x_0\|.$$

Es sei nun vorausgesetzt, daß (13) für  $k=1, 2, \dots, n-1$  gültig ist. Dann gehören die Elemente  $x_1, x_2, \dots, x_{n-1}$  zur  $r$ -Umgebung von  $x_0$ , da die Ungleichung

$$(14) \quad \|x_k - x_0\| \leq \|x_0 - x_1\| + \|x_1 - x_2\| + \dots + \|x_{k-1} - x_k\| \leq \\ \leq B\eta + \frac{3}{2} B^2 \eta^2 \frac{3\eta + Cr^2}{r\eta} + \dots + B^k \eta^k \left(\frac{3}{2} \frac{3\eta + Cr^2}{r\eta}\right)^{k-1} \leq r$$

angesichts (5) und (6) gültig ist. Nun haben wir

$$\|x_n - x_{n-1}\| = \|F'(x_{n-2})^{-1} \tau(x_{n-1}) - F'(x_{n-1})^{-1} \tau(x_{n-1})\| \leq \\ \leq \|F'(x_{n-2})\| \cdot \|\tau(x_{n-2}) - \tau(x_{n-1})\| + \|F'(x_{n-2})^{-1} - F'(x_{n-1})^{-1}\| \cdot \|\tau(x_{n-2})\|.$$

Wegen (9) ist

$$(15) \quad \|F'(x_{n-2})\| \leq \frac{3}{2} B.$$

Infolge (6) und (14) besteht die Ungleichung

$$(16) \quad \|\tau(x_{n-1})\| \leq \|F'(x_{n-1})\| + \|F'(x_1)\| \cdot \|x_{n-1} - x_0\| \leq \eta + \frac{\eta}{r} r = 2\eta,$$



somit kann man unter Benutzung des Mittelwertsatzes folgendes behaupten:

$$\begin{aligned} \| \tau(x_{n-2}) - \tau(x_{n-1}) \| &\leq \| F'(\bar{x}_{n-1}) \| \cdot \| x_{n-1} - x_{n-2} \| + \| [F'(x_{n-2}) - F'(x_{n-1})](x_{n-2} - x_0) + \\ &+ F'(x_{n-1})(x_{n-2} - x_{n-1}) \| \leq \frac{\eta_l}{r} \| x_{n-1} - x_{n-2} \| + Cr \| x_{n-1} - x_{n-2} \| + \\ (17) \quad &+ \frac{\eta_l}{r} \| x_{n-1} - x_{n-2} \| = \left( \frac{2\eta_l}{r} + Cr \right) \| x_{n-1} - x_{n-2} \| \\ &(\bar{x}_{n-1} = x_{n-2} + \theta_{n-1}(x_{n-1} - x_{n-2}); 0 < \theta_{n-1} < 1). \end{aligned}$$

Mit Rücksicht auf (15), (16) und (17) erhalten wir die zu beweisende Ungleichung

$$\begin{aligned} \| x_n - x_{n-1} \| &\leq \frac{3}{2} B \left( \frac{2\eta_l}{r} + Cr \right) \| x_{n-1} - x_{n-2} \| + 2\eta_l \frac{9}{4} B^2 C \| x_{n-1} - x_{n-2} \| \leq \\ &\leq \frac{3}{2} B \eta_l \frac{3r + Cr^2}{r \eta_l} \| x_{n-1} - x_{n-2} \|. \end{aligned}$$

Daraus ergibt sich die Konvergenz von  $\{x_n\}$  unmittelbar. Denn nach (6) ist

$$\eta_l = \frac{2-3BCr}{11B} r < \frac{2-3BCr}{9B} r,$$

und daraus folgt, daß

$$(18) \quad B \eta_l P = B \eta_l \frac{3}{2} \frac{3\eta_l + Cr^2}{r \eta_l} < 1$$

ist. Nun dürfen wir schreiben

$$\begin{aligned} \| x_{k+n} - x_k \| &\leq \| x_k - x_{k+1} \| + \| x_{k+1} - x_{k+2} \| + \dots + \| x_{k+n-1} - x_{k+n} \| \leq \\ &\leq B \eta_l (B \eta_l P)^k [1 + B \eta_l P + \dots + (B \eta_l P)^{n-1}] < \frac{B \eta_l}{1 - B \eta_l P} (B \eta_l P)^k \rightarrow 0. \end{aligned}$$

Dieses folgt aus (18) für alle ganzen  $n$ . Damit ist die Existenz von

$$\lim_{n \rightarrow \infty} x_n = x^*$$

bewiesen.

Daß  $x^*$  eine Lösung von (10) sein muß, kann in üblicher Weise leicht eingesehen werden, da die Operatoren  $F(x)$  und  $F'(x)$  in  $x$  stetig sind. Die Gleichungen (11) sind mit

$$F(x_{n-1}) = F'(x_{n-1})(x_{n-1} - x_n) \quad (n = 1, 2, \dots)$$

identisch und deswegen ist die folgende Beziehung richtig (es muß noch erwähnt werden, daß  $x^*$  auch in die  $r$ -Umgebung von  $x_0$  gehört):

$$\begin{aligned} \lim_{n \rightarrow \infty} \| F(x_{n-1}) \| &\leq \lim_{n \rightarrow \infty} \| F'(x_{n-1}) \| \lim_{n \rightarrow \infty} \| x_{n-1} - x_n \| \leq \\ &\leq \| F'(x^*) \| \lim_{n \rightarrow \infty} \frac{B \eta_l}{1 - B \eta_l P} (B \eta_l P)^n = 0. \end{aligned}$$



Damit ist aber

$$f(x^*) - y = F(x^*) = 0$$

bewiesen.

Da aber (10) mit (3) äquivalent ist, ist die Existenz der Lösung von (3) bewiesen. Die Unizität läßt sich ebenfalls in der üblichen Weise beweisen. Vorausgesetzt, daß es zwei verschiedene Lösungen von (10) in der  $r$ -Umgebung von  $x_0$  gibt, etwa  $x^*$  und  $x^{**}$ , ergibt sich wegen (15), (16), (17) und (18)

$$\begin{aligned} \|x^* - x^{**}\| &= \|F'(x^{**})^{-1} \tau(x^{**}) - F'(x^*)^{-1} \tau(x^*)\| \leq \\ &\leq \|F'(x^{**})^{-1}\| \cdot \|\tau(x^{**}) - \tau(x^*)\| + \|F'(x^*)^{-1} - F'(x^{**})^{-1}\| \cdot \|\tau(x^*)\| \leq \\ &\leq B_1 \frac{3}{2} \frac{3\eta + Cr^2}{r\eta} \|x^* - x^{**}\| < \|x^* - x^{**}\|. \end{aligned}$$

Das ist aber unmöglich, und damit ist die Unizität bewiesen.

Zum Schluß sei noch erwähnt, daß (11) mit (7) identisch ist, und somit sind alle Behauptungen des Satzes völlig bewiesen.

3. Zur Sicherung der Konvergenz des Prozesses (2') genügen schon die schwächeren Bedingungen. Und zwar gilt die folgende Behauptung [3]:

SATZ 2. *Bildet der Operator  $f(x)$  den Banachschen Raum  $X$  in den Banachschen Raum  $Y$  ab und besitzt er in einer Umgebung des Elementes  $x_0$  ( $\in X$ ) eine im Fréchet'schen Sinne stetige Ableitung, existiert außerdem  $f'(x_0)^{-1}$  (der Umkehrungsoperator von  $f'(x_0)$ ), so hat die Gleichung (3) eine einzige, in der Umgebung von  $x_0$  liegende Lösung, vorausgesetzt, daß  $y$  zu  $f(x_0)$  genügend nahe liegt. Genauer: ist  $0 < q < 1$  eine beliebige Zahl und ist die Zahl  $r$  so gewählt, daß die Ungleichungen*

$$(19) \quad \|f'(x_0)^{-1} [f'(x) - f'(x_0)]\| \leq q; \quad \|x - x_0\| \leq r$$

bestehen, dann ist für jedes, der Ungleichung

$$(20) \quad \|f'(x_0)^{-1} [y - f(x_0)]\| \leq r(1 - q)$$

genügende Element des Raumes  $Y$ , die Gleichung (3) in der  $r$ -Umgebung von  $x_0$  eindeutig lösbar. Die Lösung wird als Grenzwert der durch die Gleichungen

$$(21) \quad \xi_n = \xi_{n-1} - F'(x_0)^{-1} F(\xi_{n-1})$$

definierten Elemente von  $X$  dargestellt<sup>1</sup> ( $F(x) \equiv f(x) - y$ ).

BEWEIS. Wir definieren den Operator

$$(22) \quad \varepsilon(x) = f(x) - f'(x_0)(x - x_0).$$

Er ist differenzierbar und somit haben wir

$$\varepsilon'(x) = f'(x) - f'(x_0).$$

<sup>1</sup> Auf diese letztere Behauptung des Satzes wurde ich von A. RÉNYI aufmerksam gemacht.



Mittels  $\varepsilon(x)$  läßt sich die Gleichung (3) ersichtlich in der Form

$$\varepsilon(x) = y - f'(x_0)(x - x_0)$$

darstellen, woraus

$$(23) \quad x = x_0 + f'(x_0)^{-1}[y - \varepsilon(x)]$$

folgt. Nun sei

$$\xi_0 = x_0$$

und

$$(24) \quad \xi_n = x_0 + f'(x_0)^{-1}[y - \varepsilon(\xi_{n-1})] \quad (n = 1, 2, 3, \dots).$$

Nach Anwendung des Mittelwertsatzes erhalten wir

$$(25) \quad \begin{aligned} \xi_n - \xi_{n-1} &= f'(x_0)^{-1} \varepsilon(\xi_{n-2}) - f'(x_0)^{-1} \varepsilon(\xi_{n-1}) = \\ &= f'(x_0)^{-1} [\varepsilon(\xi_{n-2}) - \varepsilon(\xi_{n-1})] = f'(x_0)^{-1} \varepsilon'(\bar{\xi}_{n-1})(\xi_{n-2} - \xi_{n-1}); \end{aligned}$$

hier ist

$$\bar{\xi}_{n-1} = \xi_{n-2} + \theta_{n-1}(\xi_{n-1} - \xi_{n-2}) \quad (0 < \theta_{n-1} < 1).$$

Alle  $\xi_n$  und somit natürlich auch alle  $\bar{\xi}_n$  gehören zur  $r$ -Umgebung von  $x_0$ . Denn definitionsgemäß ist

$$\xi_1 = x_0 + f'(x_0)^{-1}[y - \varepsilon(x_0)],$$

und daher ist wegen (20)

$$\|\xi_1 - x_0\| = \|f'(x_0)^{-1}[y - \varepsilon(x_0)]\| = \|f'(x_0)^{-1}[y - f(x_0)]\| \leq r(1 - q) < r,$$

d. h.  $\xi_1$  und damit natürlich auch  $\bar{\xi}_1$  gehört zur  $r$ -Umgebung von  $x_0$ . Es sei angenommen, daß  $\xi_k$  ( $k \leq n-1$ ) zur  $r$ -Umgebung von  $x_0$  gehören; dann schließen wir aus (25), daß auch  $\xi_n$  in dieser Umgebung von  $x_0$  liegt. Nach Voraussetzung ist

$$(26) \quad \begin{aligned} \|\xi_n - \xi_{n-1}\| &= \|f'(x_0)^{-1} \varepsilon'(\bar{\xi}_{n-1})(\xi_{n-2} - \xi_{n-1})\| \leq \\ &\leq \|f'(x_0)^{-1} \varepsilon'(\bar{\xi}_{n-1})\| \cdot \|\xi_{n-1} - \xi_{n-2}\| = \|f'(x_0)^{-1} [f'(\bar{\xi}_{n-1}) - f'(x_0)]\| \cdot \|\xi_{n-1} - \xi_{n-2}\| \leq \\ &\leq q \|\xi_{n-1} - \xi_{n-2}\| \leq q^n \|\xi_1 - x_0\| \end{aligned}$$

wegen der Relation (19). Es ist nun

$$\begin{aligned} \|\xi_n - x_0\| &\leq \|\xi_n - \xi_{n-1}\| + \|\xi_{n-1} - \xi_{n-2}\| + \dots + \|\xi_1 - x_0\| \leq \\ &\leq (1 + q + q^2 + \dots + q^{n-1}) \|\xi_1 - x_0\| = \frac{1 - q^n}{1 - q} \|\xi_1 - x_0\| < \frac{1}{1 - q} r(1 - q) = r, \end{aligned}$$

womit unsere Behauptung bestätigt ist. Da

$$\begin{aligned} \|\xi_{n+k} - \xi_k\| &\leq \|\xi_{n+k} - \xi_{n+k-1}\| + \|\xi_{n+k-1} - \xi_{n+k-2}\| + \dots + \|\xi_{k+1} - \xi_k\| < \\ &\leq (q^k + q^{k+1} + \dots + q^{n+k}) \|\xi_1 - x_0\| < q^k \frac{1 - q^{n+1}}{1 - q} r(1 - q) < q^k \end{aligned}$$

für alle  $n$  und  $0 < q < 1$  erfüllt ist, ist die Existenz von

$$\lim_{n \rightarrow \infty} \xi_n = x$$

gesichert.



Daß das so erhaltene Element eine Lösung von (3) ist, ist fast trivial. Denn  $f(x)$  ist stetig, und aus (24) folgt, daß

$$f(\xi_{k-1}) - y = f'(x_0) (\xi_k - \xi_{k-1})$$

gültig ist. Nun gilt

$$\lim_{k \rightarrow \infty} \|f(\xi_{k-1}) - y\| \leq \lim_{k \rightarrow \infty} \|f'(x_0)\| \cdot \|\xi_k - \xi_{k-1}\| < \|f'(x_0)\| \lim_{k \rightarrow \infty} q^k = 0.$$

Das bedeutet eben die Behauptung.

Aus dem vorigen ist auch klar, daß  $x$  zu der  $r$ -Umgebung von  $x_0$  gehört.  $x$  ist aber die einzige Lösung in dieser Umgebung, denn falls  $\tilde{x}$  eine andere Lösung im betrachteten Bereiche ist, so wäre

$$f(x) - f(\tilde{x}) = 0$$

gültig. Nach (22) folgt daraus

$$\begin{aligned} \|x - \tilde{x}\| &= \|f'(x_0)^{-1} [f(x) - f(\tilde{x})]\| \leq \|f'(x_0)^{-1}\| \cdot \|f(x) - f(\tilde{x})\| \leq \\ &\leq q \|x - \tilde{x}\| < \|x - \tilde{x}\| \quad (x = x + \theta(\tilde{x} - x); 0 < \theta < 1) \end{aligned}$$

und das ist unmöglich.

Es sei wiederum bemerkt, das wegen (22) die Gleichungen (21) und (24) identisch sind, und somit ist Satz 2 völlig bewiesen.

Mit den Anwendungen der Sätze 1 und 2 beschäftigen wir uns diesmal nicht; die unter [1] und [3] zitierten Arbeiten besprechen zahlreiche Anwendungen.

(Eingegangen am 13. Januar 1954.)

### Literaturverzeichnis

- [1] A. B. Канторович, О методе Ньютона, Труды Матем. Инст. им. Стеклова, **28** (1949), S. 104—144.
- [2] И. П. Мысовских, К вопросу о сходимости метода Ньютона, Труды Матем. Инст. им. Стеклова, **28** (1949), S. 145—147.
- [3] FENYŐ ISTVÁN, Banach-terekben értelmzett nemlineáris egyenletekről, *Magyar Tud. Akad. III. Osztályának Közleményei*, **3** (1953), S. 71—83.
- [4] M. L. STEIN, Sufficient conditions for the convergence of Newton's method in complex Banach spaces, *Proc. Amer. Math. Soc.*, **3** (1952), S. 858—863.
- [5] E. HILLE, *Functional analysis and semigroups*, Amer. Math. Soc. Colloquium Publ. 31 (New York, 1948), S. 71.
- [6] M. ZORN, Gateaux differentiability and essential boundedness, *Duke Math. Journal*, **12** (1945), S. 579—583.
- [7] S. BANACH, *Théorie des opérations linéaires*. (Warszawa, 1932).
- [8] Л. А. Люстерник, Основные идеи функционального анализа, *Успехи Математических Наук*, **1** (1936), S. 77—114.
- [9] M. FRÉCHET, La notion de différentielle dans l'Analyse générale, *Ann. Sc. de l'École Norm-Supér.*, **42** (1925), S. 293—323.
- [10] R. GATEAUX, Sur les fonctionnelles continues et les fonctionnelles analytiques, *Comptes Rendus Acad. Sci., Paris*, **157** (1913), S. 325—327.
- [11] М. М. Вайнберг, Некоторые вопросы дифференциального исчисления в линейных пространствах, *Успехи Математических Наук*, **7** (1952), S. 55—102.



## О НЕЛИНЕЙНЫХ УРАВНЕНИЯХ, ОПРЕДЕЛЕННЫХ В ПРОСТРАНСТВЕ БАНАХА

И. ФЕНЬЁ (Будапешт)

(Резюме)

Пусть  $f(x)$  нелинейный оператор, определенный в пространстве Банаха  $X$ , со значениями, принадлежащими пространству Банаха  $Y$ , и дифференцируемый в смысле Фреше. Л. В. Канторовичем и другими авторами [1, 2] исследовался вопрос, каким именно способом можно применять алгоритмы

$$(1) \quad \begin{aligned} x_n &= x_{n-1} - [f'(x_{n-1})]^{-1} f(x_{n-1}) \text{ соотв.} \\ \xi_n &= \xi_{n-1} - [f'(\xi_{n-1})]^{-1} f(\xi_{n-1}) \end{aligned}$$

к решению уравнения

$$f(x) = 0.$$

Им были найдены условия, достаточные для того, чтобы эти два итерационных процесса сходились и дали корень рассматриваемого нелинейного уравнения. При формулировке и при доказательстве этих теорем  $f(x)$  предполагалась дважды дифференцируемой в смысле Фреше, и предполагалась также органиченность второй производной в окрестности корня уравнения. Этим же вопросом занимался также М. Л. Штейн; Штейном было найдено условие, также достаточное для сходимости вышеуказанных алгоритмов, предполагающее только существование первой производной  $f(x)$ . Однако Штейн предполагает  $f(x)$  дифференцируемой в смысле Гато, что равносильно аналитичности  $f(x)$  [6], так что его предположения в конечном счёте более сильны, чем те, которые сделал Л. В. Канторович.

В настоящей работе сообщается достаточное условие для сходимости алгоритмов (1), при предположении только однократной дифференцируемости в смысле Фреше  $f(x)$ . В случае первого из алгоритмов (1) мы должны также предполагать  $f'(x)$  удовлетворяющей условию Липшица. Доказательство сходимости второго из алгоритмов (1) удается провести, требуя от  $f'(x)$  только непрерывности.

Работа является дополненным переводом статьи, написанной автором ранее на венгерском языке [3].







# ON THE FUNDAMENTAL THEOREM OF COMMUTATIVE IDEAL THEORY

By

L. FUCHS (Budapest)

(Presented by L. RÉDEI)

1. Let  $R$  be a commutative ring with unit element. One says that in  $R$  the fundamental theorem of ideal theory holds if each non-trivial ideal  $A$  of  $R$  may be represented as the product of a finite number of prime ideals,

$$(1) \quad A = P_1^{k_1} P_2^{k_2} \dots P_r^{k_r}$$

where the different prime ideals<sup>1</sup>  $P_i$  ( $i = 1, 2, \dots, r$ ) and the natural integral exponents  $k_i$  are uniquely determined by the ideal  $A$ .

Since the appearance of the fundamental works of M. SONO and E. NOETHER, several authors have discussed necessary and sufficient conditions under which the fundamental theorem of ideal theory holds in  $R$ .<sup>2</sup> All of these investigations are concerned with the set of *all* non-trivial (or all regular) ideals of  $R$ . Our purpose in the present note is to discuss the same problem for a *single* ideal  $A$  of  $R$ , i. e. to give a necessary and sufficient condition for an arbitrary but fixed ideal  $A$  to admit a unique representation (1).<sup>3</sup> We emphasize that we do not want to assume anything on the totality of the ideals of  $R$ . In order to get a result for a class of rings as general as possible, the uniqueness of (1) will be taken in a somewhat less strict sense: we shall not exclude the possibility  $P^{k+1} = P^k$ , i. e. only the uniqueness of the prime ideal powers  $P^k$  will be required; then the exponents  $k_i$  will not be necessarily uniquely determined unless minimality is assumed. On the other hand, it turns out that it is not enough to require the representability in the unique form (1), but it is necessary to suppose — what is automatically satisfied if the fundamental theorem holds for all ideals of  $R$  — that each divisor  $B$  of  $A$  has a product representation which is a part of (1). More precisely:

<sup>1</sup> The ring  $R$  itself will not be considered as a prime ideal.

<sup>2</sup> SONO [7], NOETHER [6]. For a brief survey of various results see COHEN [1].

<sup>3</sup> Let us mention that in a previous paper [2] we have given a sufficient condition, but this was not necessary and besides, the maximum condition was supposed to hold for the ideals of  $R$ .



we shall say that for the ideal  $A$  of  $R$  the fundamental theorem of ideal theory holds, if (1) is true with uniquely determined different prime ideals  $P_i$  and uniquely determined minimal exponents  $k_i$ , further,<sup>4</sup>  $A \subseteq B \subset R$  implies  $B = P_1^{l_1} P_2^{l_2} \dots P_r^{l_r}$  with  $0 \leq l_i \leq k_i$ .

In what follows we establish a necessary and sufficient condition ensuring that for an ideal  $A$  the fundamental theorem of ideal theory holds. The ring  $R$  is an arbitrary commutative ring, we only assume the presence of a unit element  $e$  in  $R$ . As regards our discussion we remark that it is elementary and self-contained, we shall make use only of the simplest fundamental results of ideal theory; we shall avoid the use of transfinite methods.

The only not quite familiar notion which we need is the following. Let  $P$  be a prime ideal in  $R$ . By the principal component<sup>5</sup>  $A(P)$  of  $A$ , associated with  $P$ , we mean the set of all  $x \in R$  satisfying  $cx \in A$  for some  $c \notin P$ . It is readily seen that  $A(P)$  is an overideal of  $A$  and, if  $A \subseteq P$ , then also a subideal of  $P$ . Evidently,  $A \subseteq B$  implies  $A(P) \subseteq B(P)$ .

2. The theorem which we intend to prove here is what follows.

THEOREM. For the ideal  $A$  of the commutative ring  $R$  with unit element  $e$  the fundamental theorem of ideal theory holds if and only if

(i) the minimum condition holds modulo  $A$ ,<sup>6</sup>

(ii) for each prime divisor  $P$  of  $A$  with  $A(P) \subset P$ ,  $P^2$  is an immediate multiple of  $P$ .<sup>7</sup>

3. NECESSITY. If the fundamental theorem of ideal theory holds for  $A = P_1^{k_1} \dots P_r^{k_r}$  (where  $k_i$  are as minimal as possible), then  $A$  has but a finite number of overideals, so that (i) is fulfilled. But, as it is well known, (i) implies that the prime ideals  $P_i$  ( $i = 1, 2, \dots, r$ ) are maximal in  $R$ . Hence  $P_1^{k_1} \dots P_{i-1}^{k_{i-1}} P_{i+1}^{k_{i+1}} \dots P_r^{k_r}$  certainly contains an element  $x$  not in  $P_i$ , and thus  $x P_i^{k_i} \subseteq A$  implies  $P_i^{k_i} \subseteq A(P_i)$ . This inclusion relation can not be a proper one, for in the contrary case there would exist an element  $y$  with  $y \in A(P_i)$ ,  $y \in P_i^{k_i-1}$ ,  $y \notin P_i^{k_i}$ , and if  $c \notin P_i$ ,  $cy \in A \subseteq P_i^{k_i}$ , then  $P_i y \subseteq P_i^{k_i}$  implies

$$y \in Ry = (P_i, c)y = (P_i y, cy) \subseteq P_i^{k_i},$$

a contradiction. Now, if  $A(P_i) = P_i^{k_i} \subset P_i$ , then  $k_i \geq 2$ , and there is no ideal  $C$  between  $P_i$  and  $P_i^2$ ,  $P_i \supset C \supset P_i^2$ , for then  $C$  would be an overideal of  $A$ , not representable as the product of prime ideal powers.

<sup>4</sup>  $\subseteq$  means inclusion, and  $\subset$  is reserved to denote a proper one.

<sup>5</sup> This concept is due to KRULL [3]. (For the basic notions and results of commutative ideal theory see KRULL [4], VAN DER WAERDEN [8] or MCCOY [5].)

<sup>6</sup> We say that the minimum condition holds modulo  $A$  if each set of overideals of  $A$  contains a minimal ideal, i. e. one not containing properly any ideal of the same set.

<sup>7</sup> We observe that if we should like to have also the exponents  $k_i$  in (1) to be uniquely determined by  $A$ , then it would be necessary to add a third postulate: (iii)  $A(P)$ ,  $P$  is different from  $A(P)$ . The proof is immediate and may be left to the reader.



4. The sufficiency part of the proof rests on the following simple lemmas.

LEMMA 1.<sup>8</sup> *If modulo A the minimum condition holds and  $A \subseteq B \subset R$ , then there is a prime ideal P with  $B \subseteq P \subset R$  and  $B : P \supset B$ .*

The set of all ideals  $C(\supset B)$  with  $B : C \neq R$  is not vacuous, for it contains  $R$ . If  $C_0$  is a minimal ideal in this set, then  $P = B : C_0$  is a prime ideal, since  $a \notin P, b \notin P, ab \in P$  would imply  $P \subset P : a \subset R$  (for  $b \in P : a$  and  $Ra \subseteq P$ ) and hence

$$P : a = (B : C_0) : a = B : (aC_0) = B : (aC_0, B)$$

implies

$$B : C_0 \subset B : (aC_0, B) \subset R,$$

in contradiction to the minimality of  $C_0$ . Thus  $P$  is prime;  $B : P = B$  is impossible considering that  $B : P \supseteq C_0$ .

LEMMA 2. *If modulo A the minimum condition holds and  $A \subseteq P$ , then for some natural integer n we have  $P^n \subseteq A(P)$ .*

$A \subseteq P$  implies  $A(P) \neq R$ ; consequently, by Lemma 1 there is a prime ideal  $P^*$  such that  $A(P) : P^* \supset A(P)$ . Here  $P^* \neq P$  is impossible, for  $x \in A(P) : P^*, c \in P^*, c \notin P$  imply  $cx \in A(P)$ . By definition, there is some  $d \notin P$  such that  $cdx \in A$ , whence  $cd \notin P$  implies  $x \in A(P)$ , i. e.  $A(P) : P^* \subseteq A(P)$ , a contradiction. Hence  $A(P) : P \supset A(P)$ .

If  $A(P) : P^m \neq R$ , then let  $P^*$  be a prime divisor of  $A(P) : P^m$  with  $A(P) : P^m \subset (A(P) : P^m) : P^* = (A(P) : P^*) : P^m$ . Therefore  $A(P) \subset A(P) : P^*$  and the preceding paragraph shows that  $P^* = P$ . If  $(A(P), P^n)$  is a minimal ideal in the set of ideals  $(A(P), P^m)$  ( $m = 1, 2, \dots$ ), all dividing  $A$ , then

$$A(P) : P^n = A(P) : (A(P), P^n) = A(P) : (A(P), P^{n+1}) = A(P) : P^{n+1},$$

and what has been said implies  $A(P) : P^n = R$ , i. e.  $P^n \subseteq A(P)$ .

LEMMA 3. *If the overideals of A satisfy the minimum condition, then A may be represented as the product of a finite number of its principal components.*

First of all we show that  $A$  is the intersection of its principal components,  $A = \bigcap_v A(P_v)$ .<sup>9</sup> Clearly, it is sufficient to show that  $x \in \bigcap_v A(P_v)$  implies  $x \in A$ .

The ideal  $A : x$  is no subideal of  $P_v$ , for from  $x \in A(P_v)$  we conclude to the existence of a  $c_v \notin P_v$  with  $c_v x \in A$ , whence  $c_v \in A : x$ . The last fact, together with  $A \subseteq A : x$  and Lemma 1, shows that  $A : x = R$ , i. e.  $x \in A$ .

Now observe that from the minimum condition modulo  $A$  it follows at once that a finite number of  $A(P_v)$  ( $A \subseteq P_v$ ) suffices to represent  $A$ :

$$A = A(P_1) \cap A(P_2) \cap \dots \cap A(P_r).$$

<sup>8</sup> This lemma, together with its proof, is due to COHEN [1].

<sup>9</sup> This is true without any assumption, but the proof makes use of some form of the well-ordering hypothesis (see KRULL [3]).



The product representation will follow from

$$(A(P_1) \dots A(P_{i-1}), A(P_i)) = R \quad (i=2, \dots, r).$$

This relation actually holds, considering that the ideal in the left member can not have any prime divisor  $P$ , for in the contrary case by Lemma 2 we should have  $P_i^{n_i} \subseteq A(P_i) \subseteq P$  and  $P_1^{n_1} \dots P_{i-1}^{n_{i-1}} \subseteq A(P_1) \dots A(P_{i-1}) \subseteq P$ , that is,  $P_i \subseteq P$  and  $P_j \subseteq P$  for some  $j=1, 2, \dots, i-1$ , which is absurd. Consequently,  $A = A(P_1)A(P_2) \dots A(P_r)$  as we wished to prove.

LEMMA 4.<sup>10</sup> *If the maximal ideal  $P$  is an immediate divisor of its square  $P^2$ , then  $P^n \subseteq B \subseteq P$  implies  $B = P^k$  for some  $k$ .*

Hypothesis implies that for  $c \in P, c \notin P^2$  one has  $(c, P^2) = P$ . By induction we get

$$(2) \quad P^n = (c^n, P^{n+1}).$$

In fact, if (2) is true for some  $n$ , then

$$P^{n+1} = (c^n P, P^{n+2}) = (c^n (c, P^2), P^{n+2}) = (c^{n+1}, c^n P^2, P^{n+2}) = (c^{n+1}, P^{n+2}).$$

Let now  $s$  and  $t$  ( $s \geq t$ ) be the least resp. the greatest exponent with

$$P^s \subseteq B \subseteq P^t.$$

If  $s=t$ , we are ready. We are going to show that  $s > t$  is impossible. If  $s > t$ , then there is a  $b \in B$  with  $b \notin P^{t+1}$ , whence by  $b \in P^t$  and (2) we conclude that  $b - rc^t \in P^{t+1}$  for some  $r \in R$ . Here  $r \notin P$ , since  $r \in P, rc^t \in P^{t+1}, b \in P^{t+1}$  is absurd. We obtain  $rc^t \in (B, P^{t+1})$  and hence  $Pc^t \subseteq P^{t+1}$  implies  $c^t \in Rc^t = (P, r)c^t = (Pc^t, rc^t) \subseteq (B, P^{t+1})$ . Therefore we infer

$$c^{s-1} = c^t c^{s-t-1} \in (B, P^{t+1}) P^{s-t-1} \subseteq (B, P^t) = B$$

and by (2)

$$P^{s-1} = (c^{s-1}, P^s) \subseteq B$$

which contradicts the minimality of  $s$ .

5. SUFFICIENCY will be a simple consequence of the above lemmas. Assuming (i) and (ii) to hold, by Lemma 3 we obtain that  $A$  is the product of a finite number of its principal components  $A(P_i)$ . In case  $A(P_i) \subset P_i$  from Lemmas 2 and 4 it follows  $A(P_i) = P_i^{k_i}$  ( $k_i \geq 2$ ), i. e. (1) holds. Since  $A$  can not have prime divisors other than  $P_i$  ( $i=1, 2, \dots, r$ ), further  $A(P_i)$  is evidently uniquely determined by  $P_i$  and  $k_i$  may be assumed to be chosen as minimal as possible, the unicity of (1) has been established.

There remains still to show the statement concerning the divisors of  $A$ . If  $A \subseteq B \subset R$ , then (i) and (ii) hold for  $B$  too (in place of  $A$ ); consequently, we have  $B = P_1^{l_1} P_2^{l_2} \dots P_r^{l_r}$ . Here we may take without loss of generality  $0 \leq l_i \leq k_i$ , considering that  $P_i^{k_i} = A(P_i) \subseteq B(P_i) = P_i^{l_i}$ .

This completes the proof of our theorem.

(Received 1 February 1954)

<sup>10</sup> See NOETHER [6].



## Bibliography

- [1] I. S. COHEN, Commutative rings with restricted minimum condition, *Duke Math. Journal*, **17** (1950), pp. 27—41.  
 [2] L. FUCHS, A note on the idealizer of a subring, *Publicationes Math. Debrecen*, **1** (1950), pp. 160—161.  
 [3] W. KRULL, Idealtheorie in Ringen ohne Endlichkeitsbedingung, *Math. Annalen*, **101** (1929), pp. 729—744.  
 [4] W. KRULL, *Idealtheorie*, Ergebnisse d. Math. u. Grenzgeb., **4** (Berlin, 1935).  
 [5] N. H. MCCOY, *Rings and ideals*, Carus Math. Monographs, **8** (1948).  
 [6] E. NOETHER, Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionskörpern, *Math. Annalen*, **96** (1927), pp. 26—61.  
 [7] M. SONO, On congruences, I—IV, *Mem. Coll. Sci. Kyoto*, **2** (1917), pp. 203—226; **3** (1918), pp. 113—149, 189—197 and 229—308.  
 [8] B. L. VAN DER WAERDEN, *Moderne Algebra*, vol. II (Berlin, 1940).

## ОБ ОСНОВНОЙ ТЕОРЕМЕ ТЕОРИИ КОММУТАТИВНЫХ ИДЕАЛОВ

Л. ФУКС (Будапешт)

(Резюме)

Мы говорим, что идеал  $A$  коммутативного кольца  $R$  удовлетворяет основной теореме теории идеалов, если  $A$  допускает представление в виде произведения простых идеалов  $P_i$ ,

$$A = P_1^{k_1} \dots P_r^{k_r} \quad (P_i \neq P_j, \text{ если } i \neq j),$$

где степени простых идеалов  $P_i^{k_i}$  однозначно определены, и любой идеал  $B$  ( $B \supseteq A$ ) разлагается в произведение

$$B = P_1^{l_1} \dots P_r^{l_r} \quad (0 \leq l_i \leq k_i)$$

(итак, мы не требуем однозначности показателей. Показатели будут однозначно определены только, если мы требуем от них также минимальности.) Автором доказывается, что необходимое и достаточное условие для того, чтобы основная теорема теории идеалов выполнялась для фиксированного идеала  $A$  коммутативного кольца  $R$ , дается следующими двумя требованиями:

(I) Выполняется условие минимальности mod  $A$  для идеалов;

(II) Для любого главного компонента<sup>1</sup>  $A(P_i)$  идеала  $A$ , являющегося истинным подидеалом  $P_i$ , справедливо то, что между  $P_i$  и  $P_i^2$  нет других идеалов.

Если мы потребуем и однозначности показателей  $k_i$ , то мы должны также предполагать, что

(III)  $A(P_i) \cdot P_i$  отлично от  $A(P_i)$ .

<sup>1</sup> Под главным компонентом  $A(P_i)$  мы понимаем множество всех элементов  $x$  кольца  $R$ , для которых можно найти элемент  $b \notin P_i$ , такой, что  $b x \in A$ .







# SIMPLE PROOF OF THE WEDDERBURN—ARTIN STRUCTURE THEOREM

By

T. SZELE (Debrecen)

(Presented by L. RÉDEI)

## § 1. Introduction

The two fundamental Wedderburn—Artin structure theorems which determine the structure of simple resp. semisimple rings with minimum condition, play a very important role in the development of modern ring theory. This can be seen even of the great many proofs of these theorems which arose in the first half of our century. [See e. g. the papers in the Bibliography at the end of this note]. The consecutive proofs grew more and more simple, and they gave a deeper and deeper insight into the essence of the problem. The most important three stages of this development were the following: the discovery due to E. ARTIN according to which the natural ground of this problem is the theory of rings with minimum condition, and not only the less extensive class of hypercomplex systems [1]; the introducing of the important concept of a representation space due to E. NOETHER [10]; and finally, the general density theorem of C. CHEVALLEY and N. JACOBSON [see T. NAKAYAMA, *Über einfache distributive Systeme unendlicher Ränge*, *Proc. Imp. Acad. Tokyo*, **20** (1944), pp. 61—66; furthermore [8] Theorem 6, and N. JACOBSON, *Lectures in abstract algebra*, II, Linear algebra (New York, 1953), p. 274]. The last gives also a far-reaching generalization of previous results, in particular a general theory of simple rings without finiteness assumptions [8], and, on the other hand, it makes possible a treatment of the classical theory of simple rings with minimum condition, which, by its clarity and transparency, can be regarded as absolutely perfect. This general density theorem of CHEVALLEY and JACOBSON is one of the most precious pearls of modern mathematics. We remark that independently of JACOBSON, J. DIEUDONNÉ [6] also succeeded in obtaining a far-reaching generalization of the classical theory of simple rings and in deriving results essentially similar to those of JACOBSON.

In his paper [4], following the method in [8], E. ARTIN obtains a very brief and elegant proof of the classical structure theorem of simple rings with



minimum condition from the general density theorem of CHEVALLEY and JACOBSON, no idempotents, no Peirce decompositions, and, in general, no non-invariant systems and methods being required for the proof. It is also remarkable that this new method of proofs makes no use of the fact that the ring under consideration contains an identity, whereas in the classical proofs it was necessary to show this previously. In this note we shall show, with an argument of essentially the same simplicity and using as little technicalities, that from the general density theorem the classical structure theorem of semisimple rings can also be deduced. We believe that, in view of the importance of the problem, such a little remark can also be perhaps of some interest, e. g. for lecture purposes. Of course, the extensive theories worked out by JACOBSON and DIEUDONNÉ yield the mentioned classical theorem as a simple special case (see [9], pp. 315—316, and [6], p. 59), this way however is not the simplest possible, for it requires deeper results and not fully elementary concepts. One can consider also an elementary proof obtained by combining the proof in [4] with the classical method of reducing the case of semisimple rings to that of simple rings (see e. g. [3], pp. 27—31); such a treatment, however, would be methodically inhomogeneous. Thus it may be perhaps not superfluous to publish the following proof in § 3 which is thoroughly elementary and methodically conformable to the general density theorem, making use only of the most well-known basic concepts of the theory of rings and operator modules. In § 2 we give, for the sake of completeness, a proof of the general density theorem of CHEVALLEY and JACOBSON.

## § 2. The general density theorem of Chevalley and Jacobson

In what follows we consider a module (i. e. an additive abelian group)  $V$  which admits an arbitrary ring  $R$  as right operator domain. (We write all occurring operators on the right in order to avoid anti-homomorphisms.) We suppose that  $VR \neq 0$  (i. e.  $R$  operates non-trivially on  $V$ ), and that  $V$  is an irreducible  $R$ -module, i. e.  $0$  and  $V$  are the only  $R$ -submodules of  $V$ . We denote the elements of  $R$  by  $a, b$ , and the elements of  $V$  by  $x, y$ .

With each element  $a \in R$  we associate the endomorphism

$$(1) \quad x \rightarrow xa \quad (x \in V)$$

of  $V$ . This correspondence is obviously a homomorphism

$$(2) \quad R \sim L$$

of the ring  $R$  onto a subring  $L$  of the endomorphism ring of  $V$ . Now, we shall see that the ring  $\mathcal{A}$  consisting of all  $R$ -endomorphisms of  $V$  — i. e. the ring of all endomorphisms which commute with every endomorphism (1) — is a skewfield, relative to which  $V$  can be regarded as a right vector space.<sup>1</sup>

<sup>1</sup> An operator module is called a vector space if the operator domain is a skewfield.



The typical element of  $\mathcal{A}$  will be denoted by  $\alpha$ . Since, by the definition of  $\mathcal{A}$ , every endomorphism (1) commutes with any  $\alpha \in \mathcal{A}$ , i. e.

$$(x_1\alpha_1 + x_2\alpha_2)a = (x_1a)\alpha_1 + (x_2a)\alpha_2 \quad (x_i \in V, \alpha_i \in \mathcal{A}, a \in R),$$

we conclude that each element  $a$  of  $R$  induces a linear transformation (1) of the vector space  $V$  over  $\mathcal{A}$ . The general density theorem of CHEVALLEY and JACOBSON states the very important and surprising fact that — even under the above weak hypotheses (see<sup>3</sup> below) — the ring  $L$  consisting of the linear transformations (1) of the  $\mathcal{A}$ -space  $V$  is *dense*<sup>2</sup> in the following sense: for any two finite ordered sets of vectors

$$(3) \quad x_1, x_2, \dots, x_n; \quad y_1, y_2, \dots, y_n$$

such that the  $x$ 's are linearly independent over  $\mathcal{A}$ , there exists an  $a \in R$  such that  $x_i a = y_i$  ( $x_i, y_i \in V$ ;  $i = 1, 2, \dots, n$ ). In fuller detail:

THE GENERAL DENSITY THEOREM OF CHEVALLEY AND JACOBSON. *Let  $V$  be an irreducible  $R$ -module over an arbitrary right operator ring  $R$  such that  $VR \neq 0$ .<sup>3</sup> Then*

1)  *$V$  is a right vector space over a skewfield  $\mathcal{A}$  which consists of all endomorphisms of  $V$  commuting with each endomorphism (1) (i. e. of all  $R$ -endomorphisms);*

2) *the ring  $L$  consisting of the linear transformations (1) of the vector space  $V$  over  $\mathcal{A}$  is dense;*

3) *if  $R$  satisfies the minimum condition on right ideals, then the vector space  $V$  is finite-dimensional over  $\mathcal{A}$ , and so  $L$  consists of all linear transformations of  $V$ .*

PROOF. The content of statement 1) is essentially the same as that of SCHUR's lemma which can be proved in the usual way. (Here there is made no use of the condition  $VR \neq 0$ .) It is obvious that the set  $\mathcal{A}$  of all  $R$ -endomorphisms of  $V$  is a ring with identity. Now let  $\alpha \neq 0$  be an arbitrary element of  $\mathcal{A}$ . It is sufficient to show that  $\alpha$  is an automorphism of  $V$ , for then the inverse mapping  $\alpha^{-1}$  is clearly also an  $R$ -endomorphism of  $V$ , i. e.,  $\alpha^{-1} \in \mathcal{A}$ . Now, by the commutativity of  $\alpha$  with the elements of  $R$ , the kernel of the endomorphism  $\alpha$  is an  $R$ -submodule of  $V$  which is different from  $V$ ; hence this kernel is zero. Again from the mentioned commutativity we infer that the image module  $V\alpha$  is an  $R$ -submodule of  $V$  which is not zero; thus  $V\alpha = V$ , completing the proof of the statement that  $\alpha$  is an automorphism and so  $\mathcal{A}$  is a skewfield.

The validity of the statements 2) and 3) will follow immediately from the following

<sup>2</sup>  $L$  is actually dense in the sense of the finite topology for linear transformations.

<sup>3</sup> It may be noted that the only essential hypothesis is the irreducibility of  $V$ , for this implies  $VR \neq 0$  apart from the uninteresting trivial case in which the order of the module  $V$  is a prime number.



LEMMA. Let  $x_1, x_2, \dots, x_n$  be linearly independent vectors of  $V$  over  $\mathcal{A}$ . Then there exists an element  $a \in R$  such that

$$(4) \quad x_1 a = \dots = x_{n-1} a = 0,$$

$$(5) \quad x_n a \neq 0.$$

Before proving this lemma, we show that this implies statements 2) and 3) of the theorem. Let  $J$  be the set of all elements  $a \in R$  satisfying (4). Then  $J$  is a right ideal in  $R$ , and, by the Lemma,  $x_n J \neq 0$ . On the other hand,  $x_n J$  is an  $R$ -submodule of  $V$ , so that  $x_n J = V$ . Consequently there exists an element  $a_n \in J$  for which

$$x_1 a_n = \dots = x_{n-1} a_n = 0, \quad x_n a_n = y_n$$

holds. Similarly we can determine elements  $a_i$  with analogous properties for each index  $i = 1, 2, \dots, n-1$  (instead of  $n$ ), and thus we obtain that the element  $a = a_1 + a_2 + \dots + a_n$  satisfies the requirements  $x_i a = y_i$  ( $i = 1, 2, \dots, n$ ) for the prescribed vector systems (3). So we have shown that the linear transformation ring  $L$  is dense.

In order to prove statement 3) we show that if  $V$  is infinite dimensional over  $\mathcal{A}$ , then  $R$  does not satisfy the minimum condition (on right ideals). Indeed, let  $x_1, x_2, \dots, x_n, \dots$  be an infinite set of linearly independent vectors of  $V$  over  $\mathcal{A}$ . We denote by  $J_{n-1}$  the right ideal of  $R$  consisting of all elements  $a \in R$  which satisfy (4). Then, by the Lemma,  $J_n$  is a proper subset of  $J_{n-1}$ , so that  $J_1 \supset J_2 \supset \dots \supset J_n \supset \dots$  is an infinite strictly descending chain of right ideals in  $R$ . Hence  $R$  does not satisfy the minimum condition.

Thus, there remains only the Lemma to be proved. First let  $n = 1$ . In this most simple case we have to show that  $x_1 R \neq 0$ . Suppose the contrary:  $x_1 R = 0$ . Since, by hypothesis,  $x_1 \neq 0$ , the set of all vectors  $x \in V$  satisfying  $xR = 0$  is a non-zero  $R$ -submodule of  $V$  and hence equal to  $V$ . Then, however,  $VR = 0$ , in contradiction to the hypothesis  $VR \neq 0$  (and this is the only place in which we make use of this hypothesis).

Now let  $n > 1$ , and suppose, by induction, that the Lemma is true for  $n-1$  instead of  $n$ . Let  $J$  be the set of all elements  $a \in R$  such that

$$(6) \quad x_1 a = \dots = x_{n-2} a = 0.$$

(In case  $n = 2$  we take  $J = R$ .) Clearly,  $J$  is a right ideal in  $R$  and so  $x_{n-1} J$  is an  $R$ -submodule of  $V$ . Now, by our induction hypothesis, there exists an element  $a \in J$  such that  $x_{n-1} a \neq 0$ . Therefore  $x_{n-1} J \neq 0$ , i. e., by the irreducibility of  $V$ ,

$$(7) \quad x_{n-1} J = V.$$

Owing to the definition of  $J$  (see (6)) we have only to show that there exists an element  $a$  such that

$$(8) \quad a \in J, \quad x_{n-1} a = 0, \quad x_n a \neq 0.$$



Suppose this is not true, i. e. for every element  $a \in J$  the validity of  $x_{n-1}a = 0$  implies  $x_n a = 0$ . Then it is easy to show that the mapping

$$(9) \quad x = x_{n-1}a \rightarrow x_n a \quad (a \in J, x \in V)$$

is an  $R$ -endomorphism of  $V$ . In fact, (7) assures that (9) is a mapping of the whole module  $V$  into itself. This mapping is, moreover, single-valued. For if  $x = x_{n-1}a = x_{n-1}a'$  ( $a, a' \in J$ ), then  $x_{n-1}(a-a') = 0$  which implies by hypothesis  $x_n(a-a') = 0$ , i. e.,  $x_n a = x_n a'$ . It is also clear that our mapping is an endomorphism; that it is an  $R$ -endomorphism follows from the fact that  $J$  is a right ideal in  $R$ . Now, since the mapping (9) is an  $R$ -endomorphism of  $V$ , it is (by the definition of  $\mathcal{A}$ ) a scalar multiplication by a suitable element  $\alpha \in \mathcal{A}$ . This means that

$$x_n a = (x_{n-1} a) \alpha = (x_{n-1} \alpha) a,$$

i. e.

$$(10) \quad (x_n - x_{n-1} \alpha) a = 0$$

holds for every element  $a \in J$ . From this fact we infer that the  $n-1$  vectors

$$(11) \quad x_1, x_2, \dots, x_{n-2}, x_n - x_{n-1} \alpha$$

cannot be linearly independent. In fact, if the vectors in (11) were linearly independent over  $\mathcal{A}$ , then our induction hypothesis would imply the existence of an element  $a \in R$  satisfying (6), i. e.  $a \in J$  but not satisfying (10). Thus we have obtained that the vectors in (11) are linearly dependent which is impossible since  $x_1, x_2, \dots, x_n$  are linearly independent. This contradiction establishes the existence of an element  $a$  with properties (8), thus completing the proof.

### § 3. The Wedderburn—Artin structure theorem

In what follows the minimum condition (briefly m. c.) refers always to the right ideals of the ring mentioned.

Now we are going to prove on basis of the density theorem the following

**WEDDERBURN—ARTIN STRUCTURE THEOREM.** *Any ring containing no nilpotent right ideals and satisfying the minimum condition is isomorphic to a direct sum of a finite number of rings (two-sided ideals) each of which is isomorphic to the complete ring of linear transformations in a suitable finite-dimensional vector space over a skewfield.*

**REMARK.** From this theorem the second Wedderburn—Artin structure theorem follows easily: *Any simple ring with m. c. is isomorphic to the ring of all linear transformations in a finite dimensional vector space over a skewfield or is a zeroing with  $p$  elements ( $p$  a prime).* — Indeed, this is an immediate consequence of the preceding theorem if we shall show that a simple ring  $R$  containing a nilpotent right ideal  $J$  is necessarily a zeroing. But this is clear since the two-sided ideal  $RJ$  cannot coincide with  $R$  (for  $RJ = R$



implies  $RJ^n = R$ ,  $n = 1, 2, 3, \dots$  in contradiction to the hypothesis that  $J$  is nilpotent). Therefore  $RJ = 0$  and so  $R$  contains right annihilators  $\neq 0$ . But the set of all right annihilators is a two-sided ideal in  $R$  which must coincide with  $R$ . Hence  $R^2 = 0$ , i. e.,  $R$  is a zeroing.

PROOF. Let  $R$  be an arbitrary ring satisfying the m. c. and containing no nilpotent right ideal. Let  $V \neq 0$  be a minimal right ideal in  $R$ . Then  $V$  is an irreducible right  $R$ -module. Also it is clear that  $VR \neq 0$ , for  $VR = 0$  implies  $V^2 = 0$  which is impossible by the non-nilpotency of  $V$ . Therefore it follows from the general density theorem that

*$V$  is an  $n$ -dimensional vector space over a skewfield  $A$  and the ring  $L$  of linear transformations (1) of this space induced by the elements  $a \in R$  coincides with the complete ring  $A_n$  of all linear transformations in this space.*

So we have by the homomorphism (2)

$$(12) \quad R/K \cong A_n$$

where the kernel  $K$  of this homomorphism is a two-sided ideal in  $R$  consisting of all right annihilators of  $V$  (i. e. of all elements  $b \in R$  such that  $Vb = 0$ ).

In what follows we have only to show that a direct decomposition

$$(13) \quad R = R_1 + K$$

holds with a suitable two-sided ideal  $R_1$  in  $R$ . Namely this implies that  $R$  is a direct sum of a ring  $R_1$  isomorphic to  $R/K \cong A_n$  and of a ring  $K$  which again satisfies the m. c. and contains no nilpotent right ideals. Thus a similar direct decomposition holds also for  $K$  and this process leads, by the m. c., in a finite number of steps to a decomposition of  $R$  stated in the Wedderburn—Artin theorem.

The validity of the decomposition (13) can be proved for the two-sided ideal  $RV = R_1$  of  $R$ . First it is clear that the intersection  $D = RV \cap K$  is zero. For  $D$  is a two-sided ideal in  $R$  and we have, by the definition of  $K$ ,

$$D^2 = D \cdot D \subseteq (RV) \cdot K = R \cdot (VK) = R \cdot 0 = 0$$

and thus  $D = 0$ . On the other hand, the direct sum

$$(14) \quad RV + K$$

cannot be a proper two-sided ideal in  $R$ . Indeed, in the contrary case also the residue class ring  $(RV + K)/K$  would be a proper two-sided ideal of  $R/K$ . This, however, is impossible since, by (12),  $R/K$  is a simple ring, and  $(RV + K)/K = 0$  would imply  $RV \subseteq K$ , i. e. (since  $RV \cap K = 0$ )  $RV = 0$  and so  $V^2 \subseteq RV = 0$ . This contradiction completes the proof.

(Received 14 February 1954)



## Bibliography

- [1] E. ARTIN, Zur Theorie der hyperkomplexen Zahlen, *Abh. Hamburg*, **5** (1927), pp. 251—260.
- [2] E. ARTIN and G. WHAPLES, The theory of simple rings, *Amer. J. Math.*, **65** (1943), pp. 87—107.
- [3] E. ARTIN, C. J. NESBITT and R. M. THRALL, *Rings with minimum condition* (Ann. Arbor, Mich., 1948).
- [4] E. ARTIN, The influence of J. H. M. Wedderburn on the development of modern algebra, *Bull. Amer. Math. Soc.*, **56** (1950), pp. 65—72.
- [5] M. DEURING, *Algebren*, Ergebnisse der Math. u. ihrer Grenzgebiete, IV. 1 (Berlin, 1935).
- [6] J. DIEUDONNÉ, Sur le socle d'un anneau et les anneaux simples infinis, *Bull. Soc. Math. de France*, **70** (1942), pp. 46—75.
- [7] N. JACOBSON, *The theory of rings*, Math. Surveys, **2** (New York, 1943).
- [8] N. JACOBSON, Structure theory of simple rings without finiteness assumptions, *Trans. Amer. Math. Soc.*, **57** (1945), pp. 228—245.
- [9] N. JACOBSON, The radical and semi-simplicity for arbitrary rings, *Amer. J. Math.*, **67** (1945), pp. 300—320.
- [10] E. NOETHER, Hyperkomplexe Größen und Darstellungstheorie, *Math. Zeitschrift*, **30** (1929), pp. 641—692.
- [11] B. L. VAN DER WAERDEN, *Moderne Algebra*, II (Berlin, 1940).
- [12] J. H. M. WEDDERBURN, On hypercomplex numbers, *Proc. London Math. Soc.*, (2), **6** (1908), pp. 77—118.

ПРОСТОЕ ДОКАЗАТЕЛЬСТВО СТРУКТУРНОЙ ТЕОРЕМЫ  
ВЕДДЕРБЭРНА—АРТИНА

Т. СЕЛЕ (Дебрецен)

(Резюме)

Автор очень простым способом выводит из общей теоремы плотности Джексона [8] классическую структурную теорему Веддербэрна—Артина, соответствующей которой кольцо не содержащее нильпотентных правых идеалов и удовлетворяющее условию минимума для правых идеалов изоморфно прямой сумме конечного числа колец, каждое из которых изоморфно полному кольцу линейных преобразований некоторого векторного пространства конечной размерности над косым полем.







# ÜBER DIE KONVERGENZ DES HERMITE—FEJÉRSCHEN INTERPOLATIONSVERFAHRENS

Von

GÉZA FREUD (Budapest)  
(Vorgelegt von P. TURÁN)

## Einleitung

Es sei  $w(x) \in L$  eine in  $(-1, +1)$  definierte nichtnegative Gewichtsfunktion, und  $\{p_n(x)\}$  sei die Folge der zu dieser Gewichtsfunktion gehörigen normierten orthogonalen Polynome. Die Wurzeln von  $p_n(x)$  seien  $x_{1n}, x_{2n}, \dots, x_{nn}$ . Die zu diesen Grundpunkten gehörigen Hermite—Fejérschen interpolatorischen Grundpolynome seien

$$(1) \quad h_{kn}(x) = v_{kn}(x) l_{kn}^2(x)$$

und

$$(2) \quad \eta_{kn}(x) = (x - x_{kn}) l_{kn}^2(x),$$

wobei

$$(3) \quad l_{kn}(x) = \frac{p_n(x)}{p'_n(x_{kn})(x - x_{kn})}$$

die zum Punkte  $x_{kn}$  des Grundpunktsystems  $(x_{1n}, x_{2n}, \dots, x_{nn})$  gehörige Lagrangesche Grundparabel bedeutet, es ist ferner

$$(4) \quad v_{kn}(x) = 1 - \frac{p''_n(x_{kn})}{p'_n(x_{kn})} (x - x_{kn}).$$

Infolge der Definition der interpolatorischen Grundpolynome besteht für ein beliebiges Polynom höchstens  $2n-1$ -ten Grades  $\tau_{2n-1}(x)$  die Relation

$$(5) \quad \tau_{2n-1}(x) \equiv \sum_{k=1}^n \tau_{2n-1}(x_{kn}) h_{kn}(x) + \sum_{k=1}^n \tau'_{2n-1}(x_{kn}) \eta_{kn}(x).$$

Im folgenden sei  $f(x)$  eine in  $(-1, +1)$  definierte Funktion, und es soll die zugehörige Interpolationsfolge

$$(6) \quad H_n(f; x) = \sum_{k=1}^n f(x_{kn}) h_{kn}(x) + \sum_{k=1}^n d_{kn} \eta_{kn}(x)$$

betrachtet werden. Hierbei sind  $\{d_{kn}\}$  vorgeschriebene Zahlenwerte, die bestimmten Beschränkungen bezüglich ihrer Größenordnung unterworfen sind,



sonst aber beliebig wählbar sind. Somit ist  $H_n(f; x)$  das Polynom höchstens  $2n-1$ -ten Grades, dessen Wert an den Stellen  $x_{kn}$  mit  $f(x)$  übereinstimmt, und dessen Differentialquotient an denselben Stellen die vorgeschriebenen Werte  $d_{kn}$  ( $k=1, 2, \dots, n$ ) annimmt. Die ersten Untersuchungen über Interpolationsfolgen dieser Art stammen von L. FEJÉR ([4], [5], [6], [9]).

Nach FEJÉR nennen wir eine Grundpunktfolge normal, falls für jedes  $k$  und  $n$  im ganzen Interpolationsintervall  $(-1, +1)$

$$(7) \quad v_{kn}(x) \geq 0$$

besteht, und streng normal, falls sogar

$$(8) \quad v_{kn}(x) \geq \rho > 0$$

besteht, wobei  $\rho$  von  $k$  und  $n$  unabhängig ist. L. FEJÉR [7] zeigte, daß die Jacobischen Polynome, die zur Gewichtsfunktion  $w(x) = (1-x)^\alpha (1+x)^\beta$  orthogonal sind, eine normale Punktgruppenfolge bilden, falls  $-1 < \alpha, \beta \leq 0$  gültig ist, und eine streng normale Punktgruppenfolge, falls sogar  $-1 < \alpha, \beta < 0$  gilt. FEJÉR selbst untersuchte hauptsächlich die Fälle, wo die Grundpunkte als Nullstellen der Tschebyscheffschen ( $\alpha = \beta = -1/2$ ), bzw. Legendreschen Polynome ( $\alpha = \beta = 0$ ) gewählt werden. Im Tschebyscheffschen Fall zeigte er, daß die Folge  $H_n(f; x)$  in  $(-1, +1)$  gleichmäßig gegen die beliebige stetige Funktion  $f(x)$  konvergiert, falls die Folge der Deriviertenwerte in  $k$  gleichmäßig der Ungleichung

$$(9) \quad d_{kn} = o\left(\frac{n}{\log n}\right) \frac{1}{\sqrt{1-x_{kn}^2}}$$

genügt. Er zeigte ferner, daß die Konvergenz in jedem inneren Teilintervall von  $(-1, +1)$  auch noch dann besteht, falls als Grundpunkte die Nullstellen der Legendreschen Polynome gewählt werden, und die Zahlenfolge  $\{d_{kn}\}$  beschränkt ist. Später zeigte L. FEJÉR [9], daß auch für ein beliebiges normales Grundpunktsystem  $H_n(f; x) \rightarrow f(x)$  gilt, falls die Folge  $d_{kn}$  von  $f(x)$  abhängig passend gewählt wird. An die Arbeiten von L. FEJÉR anschließend, zeigt G. GRÜNWARD [13], daß die  $H_n(f; x)$  in jedem inneren Teilintervall von  $(-1, +1)$  gleichmäßig gegen die stetige Funktion  $f(x)$  konvergieren, falls die Grundpunktfolge normal ist und die Zahlen  $\{d_{kn}\}$  beschränkt sind. Ist die Grundpunktfolge streng normal, dann ist die Konvergenz im ganzen abgeschlossenen Intervall  $[-1, +1]$  gleichmäßig. Er zeigte sogar, daß in diesem Falle die Beschränktheit der  $d_{kn}$  durch  $|d_{kn}| < An^{\rho-\epsilon}$  ersetzt werden kann, wo  $\rho$  die in (8) auftretende positive Zahl ist. Für Jacobische Polynome, aber auch für die Fälle  $\alpha > 0, \beta > 0$ , wurde die Konvergenz der Hermite—Fejérschen Interpolationsfolge von G. SZEGÖ [16] und J. SHOHAT [15] bewiesen. Prof. P. TURÁN hat in persönlichen Besprechungen die Frage aufgeworfen, ob man nicht auf die Konvergenz von  $H_n(f; x)$  schließen kann, falls man die Grundpunkte  $\{x_{kn}\}$  in der schon auseinandergesetzten Weise als Nullstellen der Orthogonalpolynome  $p_n(x)$  wählt. Im folgenden bezeichnen wir die Interpolations-



tionsfolge in diesem Fall kurz als die „zu  $w(x)$  zugeordnete“. Zur Orientierung bemerken wir, daß L. FEJÉR [7] das Bestehen von (7), bzw. (8) für bestimmte Jacobische Polynome mittels der linearen Differenzialgleichung zweiter Ordnung bewiesen hat, welcher diese Polynome genügen. Andererseits ist die Frage, unter welchen Bedingungen die  $\{p_n(x)\}$  einer solchen Differentialgleichung genügen, zur Zeit unentschieden. Es muß also eine wesentlich verschiedene Beweismethode aufgebaut werden. Für eine allgemeine Gewichtsfunktion war bis jetzt auch die einfachere Frage unbeantwortet, ob  $H_n(f; x)$  konvergiert, falls  $f(x)$  stetig differenzierbar ist und  $d_{kn} = f'(x_{kn})$  gewählt wird.

Nach einer persönlichen Mitteilung von Prof. P. TURÁN, konvergieren die Hermite—Fejérschen Treppenparabeln gegen jede stetige Funktion  $f(x)$ , falls die Funktion  $w(\cos \theta) \sin \theta = g(\theta)$  in  $0 \leq \theta \leq \pi$  gleichmäßig einer logarithmischen Lipschitz-Bedingung mit einem Exponenten  $\alpha > 1$  genügt. Der Beweis beruht auf S. N. BERNSTEIN'S [1] asymptotischer Formel für die Orthogonalpolynome  $p_n(x)$ .

Im weiteren sei vorausgesetzt, daß die Folge der Orthogonalpolynome in einem inneren Teilintervall  $(\alpha, \beta)$  von  $(-1, +1)$  beschränkt ist:

$$(10) \quad |p_n(x)| \leq K \quad (\alpha \leq x \leq \beta).$$

Ferner soll die Folge  $\{d_{kn}\}$  immer so gewählt werden, daß die Abschätzung

$$(11) \quad d_{kn} = \begin{cases} o\left(\frac{n}{\log n}\right) & \text{für } \alpha \leq x_{kn} \leq \beta, \\ o\left\{\text{Min}\left(\frac{n}{\sqrt{1-x_{kn}^2}}, n^2\right)\right\} & \text{für alle } x_{kn} \end{cases}$$

gleichmäßig in den  $x_{kn}$  besteht.

Wir werden die folgenden zwei Sätze beweisen:

SATZ I. Die Gewichtsfunktion  $w(x) \in L$  sei im ganzen abgeschlossenen Orthogonalitätsintervall  $[-1, +1]$  stetig und positiv, ferner genüge  $w(x)$  in  $(\alpha, \beta)$  gleichmäßig der logarithmischen Lipschitz-Bedingung

$$(12) \quad w(x_1) - w(x_2) = o\left(\log^{-1} \frac{1}{|x_2 - x_1|}\right).$$

Dann konvergiert die Hermite—Fejérsche Interpolationsfolge (6) an der Stelle  $\xi$  ( $\alpha < \xi < \beta$ ) gegen  $f(\xi)$ , falls  $f(x)$  eine an den drei Stellen  $\xi, -1, +1$  stetige, in  $(-1, +1)$  beschränkte Funktion ist. Ist  $f(x)$  im ganzen Intervall  $(\alpha, \beta)$  stetig, dann ist die Konvergenz der Interpolationsfolge in jedem inneren Teilintervall von  $(\alpha, \beta)$  gleichmäßig.

SATZ II. Es sei

$$(13) \quad 0 < m \leq w(x) \leq M \quad (-1 \leq x \leq +1).$$

Die Funktion  $f(x)$  sei im abgeschlossenen Intervall  $[-1, +1]$  stetig, genüge



in  $(\alpha, \beta)$  gleichmäßig der logarithmischen Lipschitz-Bedingung

$$(14) \quad f(x_1) - f(x_2) = o\left(\log^{-1} \frac{1}{|x_1 - x_2|}\right) \quad (\alpha \leq x_1, x_2 \leq \beta).$$

Dann konvergiert  $H_n(f; x)$  in jedem inneren Teilintervalle von  $(\alpha, \beta)$  gleichmäßig gegen  $f(x)$ .

Es sei bemerkt, daß eine Ungleichung der Form (13) wegen der Stetigkeit und Positivität von  $w(x)$  auch aus den Voraussetzungen des Satzes I folgt.

### I. Eine Verallgemeinerung der Cotesschen Zahlen

Wir bilden aus den zu  $w(x)$  gehörigen orthogonalen Polynomen die Linearkombination

$$p_n^*(x, \xi) = p_{n-1}(\xi) p_n(x) - p_n(\xi) p_{n-1}(x).$$

Vorläufig sei  $p_{n-1}(\xi) \neq 0$ , und die  $n$  Wurzeln in  $x$  von  $p_n^*(x, \xi)$  seien  $x_1^*, x_2^*, \dots, x_n^*$ . Eine dieser Nullstellen ist  $x_v^* = \xi$ . Die zum Grundpunktsystem  $\{x_i^*\}$  gehörigen Lagrange-Parabeln seien  $l_{1n}^*(x), l_{2n}^*(x), \dots, l_{nn}^*(x)$ , und es seien die Bezeichnungen

$$l_{vn}^*(x) = l_n(x, \xi), \quad \int_{-1}^{+1} l_n(x, \xi) w(x) dx = \lambda_n(\xi)$$

eingeführt. Im folgenden nennen wir  $l_n(x, \xi)$  die zur Stelle  $\xi$  gehörende Lagrange-Parabel, und  $\lambda_n(\xi)$  die zur Stelle  $\xi$  gehörende Cotessche Zahl. Es sei  $\pi_{2n-2}(x)$  ein beliebiges Polynom  $2n-2$ -ten Grades. Nach einer klassischen Schlußweise von C. G. J. JACOBI gilt als Verallgemeinerung der mechanischen Quadraturformel von GAUSS und JACOBI

$$(15) \quad \int_{-1}^{+1} \pi_{2n-2}(x) w(x) dx = \sum_{k=1}^n \pi_{2n-2}(x_k^*) \int_{-1}^{+1} l_{kn}^*(x) w(x) dx,$$

wobei

$$(16) \quad \int_{-1}^{+1} l_{kn}^*(x) w(x) dx = \int_{-1}^{+1} l_{kn}^{*2}(x) w(x) dx > 0$$

gültig ist.<sup>1</sup> Die Positivität von (16) wurde von L. FEJÉR [8] gezeigt.

Also ist

$$(17) \quad \lambda_n(\xi) = \int_{-1}^{+1} l_n(x, \xi) w(x) dx = \int_{-1}^{+1} l_n^2(x, \xi) w(x) dx > 0.$$

Ferner erhalten wir unmittelbar aus (15) und (16), daß  $\lambda_n(\xi)$  das genaue Minimum des Ausdruckes

<sup>1</sup> Vgl. etwa meine Arbeit [12], Hilfssatz I.



$$\int_{-1}^{+1} [\pi_{n-1}(x)]^2 w(x) dx$$

angibt, falls  $\pi_{n-1}(x)$  alle Polynome  $n-1$ -ten Grades durchläuft, die der Nebenbedingung  $\pi_{n-1}(\xi) \geq 1$  genügen.

Es sei sogleich bemerkt, daß man aus (15) und (16) auch die etwas schärfere Behauptung abliest, daß  $\lambda_n(\xi)$  das genaue Minimum von

$$\int_{-1}^{+1} \pi_{2n-2}(x) w(x) dx$$

bedeutet, falls  $\pi_{2n-2}(x)$  alle Polynome höchstens  $2n-2$ -ten Grades durchläuft, die im ganzen Intervall  $[-1, +1]$  nichtnegativ sind und der Forderung  $\pi_{2n-2}(\xi) \geq 1$  genügen.

Dieses Minimumproblem wurde von G. SZEGŐ [17], ferner von P. ERDŐS und P. TURÁN [3] eingehend untersucht. Nach G. SZEGŐ<sup>2</sup> gilt

$$(18) \quad \lambda_n(\xi) = \frac{1}{\sum_{k=0}^{n-1} p_k^2(\xi)} = \frac{\alpha_n}{\alpha_{n-1}} \frac{1}{p_n'(\xi) p_{n-1}(\xi) - p_{n-1}'(\xi) p_n(\xi)}.$$

Die rechtsstehende Identität folgt aus der Summenformel von CHRISTOFFEL—DARBOUX;  $\alpha_n$  bedeutet den Koeffizienten von  $x^n$  in  $p_n(x)$ . Somit haben wir die Cotesschen Zahlen  $n$ -ter Ordnung der Gauß—Jacobischen mechanischen Quadratur zu einer stetigen „Cotesschen Funktion“  $n$ -ter Ordnung  $\lambda_n(\xi)$  verallgemeinert, die nicht nur an den Nullstellen von  $p_n(x)$ , sondern auch an jeder reellen Stelle  $\xi$  definiert ist. Für  $p_{n-1}(\xi) = 0$  sei wegen (18)  $\lambda_n(\xi) = \lambda_{n-1}(\xi)$ . Falls  $w(x) \leq W(x)$  ist und  $A_n(\xi)$  die zu  $W(x)$  zugeordnete Cotessche Funktion  $n$ -ter Ordnung bedeutet, folgt aus der Minimumeigenschaft

$$\lambda_n(\xi) \leq A_n(\xi).$$

Diese Ungleichung wurde zuerst von P. ERDŐS und P. TURÁN [3] erkannt und auf interpolationstheoretische Fragen angewandt. Infolge (18) ist  $\lambda_n(\xi)$  differenzierbar, und es gilt

$$\frac{\lambda_n'(\xi)}{\lambda_n(\xi)} = - \frac{p_n''(\xi) p_{n-1}(\xi) - p_{n-1}''(\xi) p_n(\xi)}{p_n'(\xi) p_{n-1}(\xi) - p_{n-1}'(\xi) p_n(\xi)}.$$

Ist nun  $x_{kn}$  eine Nullstelle von  $p_n(x)$ , so besteht die Beziehung

$$(19) \quad \frac{\lambda_n'(x_{kn})}{\lambda_n(x_{kn})} = - \frac{p_n''(x_{kn})}{p_n'(x_{kn})}.$$

Die Formel (19) ist die Grundlage unserer Überlegungen; mit ihrer Hilfe erhalten wir eine Abschätzung von  $v_{kn}(x)$ .

<sup>2</sup> [17], Theorem 3.1.3, S. 38.



## II. Über einige Sätze von P. Erdős und P. Turán

Der Ausgangspunkt unserer Überlegungen ist die interpolationstheoretische Arbeit von P. ERDŐS und P. TURÁN [3]. Wir zitieren kurz die Resultate dieser Arbeit, und werden sie in einigen Punkten ergänzen. Mit Hinsicht auf die Minimumeigenschaft von  $\lambda_n(\xi)$  gilt

$$(20) \quad \lambda_n(\xi) = \frac{1}{\sum_{k=1}^n \frac{l_{kn}^2(\xi)}{\lambda_{kn}}}.$$

Infolge der Minimumeigenschaft von  $\lambda_n(\xi)$  und  $w(x) \geq m > 0$  besteht gleichmäßig in jedem inneren Teilintervall von  $(-1, +1)$

$$(21) \quad \sum_{k=1}^n \frac{l_{kn}^2(x)}{\lambda_{kn}} = O(n).$$

Aus  $w(x) \leq M$  folgt, daß es eine nur von  $h$  abhängige positive Zahl  $c_0(h)$  gibt, für welche

$$(22) \quad \lambda_{kn} \geq \frac{c_0(h)}{n} \quad (-1+h \leq x_{kn} \leq 1-h)$$

erfüllt ist. Aus (21) und (22) folgt

HILFSSATZ I. *Es gilt*

$$(23) \quad \sum_{|x_{kn}| < 1-h} l_{kn}^2(x) = O(1),$$

wobei  $(-1+h, 1-h)$  ein festes inneres Teilintervall von  $(-1, +1)$  ist, und (23) gleichmäßig in  $x$  für jedes innere Teilintervall von  $(-1, +1)$  besteht.

P. ERDŐS und P. TURÁN [3] zeigen ferner, daß es infolge (13) positive Zahlen  $c_1(h)$  und  $c_2(h)$  gibt, so daß für zwei benachbarte, in  $(-1+h, 1-h)$  fallende Nullstellen  $x_{kn}$  und  $x_{k+1,n}$  die Ungleichung

$$(24) \quad \frac{c_1(h)}{n} \leq x_{k+1,n} - x_{kn} \leq \frac{c_2(h)}{n}$$

befriedigt ist.

HILFSSATZ II. *Es gilt*

$$(25) \quad \lambda_n(\xi) \leq 100 M \left( \frac{\sqrt{1-\xi^2}}{n} + \frac{1}{n^2} \right).$$

BEWEIS. Infolge der Minimumeigenschaft von  $\lambda_n(\xi)$  gilt

$$(26) \quad \lambda_n(\xi) \leq M A_n(\xi),$$

wobei  $A_n(\xi)$  die zur Gewichtsfunktion  $W(x) \equiv 1$  gehörige Cotessche Funktion  $n$ -ter Ordnung bedeutet.

Es sei

$$(27) \quad \psi_{2n-2}(\cos \vartheta) = g(\Theta - \vartheta) + g(\Theta + \vartheta),$$



wobei  $\xi = \cos \vartheta$ ,  $0 \leq \vartheta \leq \pi$  ist, und

$$g(t) = \frac{3}{2\pi n(2n^2 + 1)} \left( \frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right)^4$$

die Jacksonsche Kernfunktion ist. (Vgl. D. JACKSON [12].) Nach dem Approximationsatz von D. JACKSON erhalten wir, indem wir es auf die Funktion

$$h(\Theta) = \begin{cases} \sin \Theta & \text{für } 0 \leq \Theta \leq \pi, \\ 0 & \text{für } -\pi \leq \Theta \leq 0 \end{cases}$$

anwenden, die einer Lipschitz-Bedingung mit dem Exponenten 1 genügt:

$$(28) \quad \left| \int_0^\pi [g(\Theta - \vartheta) + g(\Theta + \vartheta)] \sin \Theta d\Theta - \sin \vartheta \right| < \frac{12}{n}.$$

Es gilt infolge (27)

$$(29) \quad \psi_{2n-2}(\xi) \geq g(0) > \frac{1}{7} n.$$

Das Polynom  $\pi_{2n-2}(x) = \frac{7}{n} \psi_{2n-2}(x)$  ist in  $[-1, +1]$  nichtnegativ und an der Stelle  $\xi$  ist  $\pi_{2n-2}(\xi) \geq 1$ ; also ist infolge des Minimumprinzips und (28)

$$(30) \quad \begin{aligned} A_n(\xi) &\leq \int_{-1}^{+1} \pi_{2n-2}(x) dx = \frac{7}{n} \int_{-1}^{+1} [g(\Theta - \vartheta) + g(\Theta + \vartheta)] \sin \vartheta d\vartheta < \\ &< \frac{7}{n} \left( \sin \vartheta + \frac{12}{n} \right) = \frac{7\sqrt{1-\xi^2}}{n} + \frac{84}{n^2}. \end{aligned}$$

Wegen (26) und (30) ist (25) erfüllt.

### III. Ein Prinzip zur Abschätzung von Polynomen vorgeschriebenen Grades

Der folgende Satz stammt von S. N. BERNSTEIN [1]: Es sei  $\pi_n(x)$  ein Polynom  $n$ -ten Grades, in  $(-1, +1)$  sei

$$|\pi_n(x)| \leq \frac{1}{\sqrt{1-x^2}},$$

dann besteht auch die Abschätzung

$$|\pi_n(x)| \leq n + 1.$$

Wir wollen ähnliche Sätze für eine recht allgemeine Klasse von Majorantenfunktionen beweisen. Diese verallgemeinerten Abschätzungen werden sich im weiteren als äußerst nützlich erweisen. Der Betrag des Polynoms  $\pi_n(x)$   $n$ -ten



Grades habe in  $[-1, +1]$  eine Majorante  $\mu(x)$ :

$$(31) \quad |\tau_n(x)| \leq \mu(x) \quad (-1 \leq x \leq +1).$$

Hierbei soll  $\mu(x)$  nicht notwendigerweise an jeder Stelle von  $[-1, +1]$  endlich sein, doch seien die Unendlichkeitsstellen von  $\mu(x)$  in endlicher Anzahl, und  $\mu(x)$  sei beschränkt auf jeder Punktmenge, die aus  $[-1, +1]$  entsteht, indem man von diesem Intervall beliebig kleine Umgebungen der Unendlichkeitsstellen ausschließt. Unter diesen Bedingungen kann für den Betrag aller Polynome, die (31) befriedigen, eine einheitliche, nur vom Grade  $n$  und der Beschaffenheit von  $\mu(x)$  abhängige Schranke angegeben werden. Das Maximum des Betrages von  $\tau_n(x)$  in  $[-1, +1]$  sei  $\mu_0 = |\tau_n(x_0)|$ . Nach dem Satze von MARKOFF gilt  $|\tau'_n(x)| \leq n^2 \mu$  in  $[-1, +1]$ . An den Stellen  $x_1$ , die in  $[-1, +1]$  und sogleich in  $\left[x_0 - \frac{1}{2n^2}, x_0 + \frac{1}{2n^2}\right]$  fallen, ist also infolge des ersten Mittelwertsatzes

$$|\tau_n(x_0) - \tau_n(x_1)| \leq \frac{1}{2n^2} n^2 \mu_0 = \frac{\mu_0}{2},$$

also  $|\tau_n(x_1)| \geq \frac{\mu_0}{2}$  und endlich

$$\mu_0 \leq 2|\tau_n(x_1)|.$$

Hieraus folgt folgender Satz:

HILFSSATZ IIIa. *Es sei  $I_n \in [-1, +1]$  ein Intervall von der Länge  $\frac{1}{2n^2}$ , ferner sei  $\mu^*(I_n)$  die untere Grenze von  $\mu(x)$  in  $I_n$  und  $\mu_0^*$  die obere Grenze aller  $\mu^*(I_n)$ . Dann besteht für jedes Polynom höchstens  $n$ -ten Grades, das (31) befriedigt, die Abschätzung*

$$|\tau_n(x)| \leq 2\mu_0^*.$$

Eine ähnliche Abschätzung erhalten wir, wenn wir statt des Satzes von MARKOFF den Satz von BERNSTEIN über den Betrag der Derivierten eines trigonometrischen Polynoms anwenden. Es sei  $\tau_n(x)$  ein trigonometrisches Polynom höchstens  $n$ -ten Grades, und es sei

$$(31a) \quad |\tau_n(x)| \leq \nu(x),$$

wobei  $\nu(x)$  eine nach  $2\pi$  periodische Funktion ist von derselben Beschaffenheit, wie  $\mu(x)$ .

HILFSSATZ IIIb.  *$J_n$  sei ein beliebiges Intervall von der Länge  $\frac{1}{n}$  und  $\nu^*(J_n) = \inf_{x \in J_n} \nu(x)$ , ferner sei  $\nu_0^* = \sup \nu^*(J_n)$ . Dann folgt aus (31a)*

$$|\tau_n(x)| \leq 2\nu_0^*.$$

Da  $\tau_n(\cos\theta)$  ein trigonometrisches Polynom höchstens  $n$ -ten Grades ist, folgt aus Hilfssatz IIIb folgender Satz:



HILFSSATZ IIIc. Es sei

$$\mu^{**}(\xi) = \inf_{|\xi - \xi'| \leq \frac{\sqrt{1-\xi^2}}{4n}} \mu(\xi')$$

und

$$\mu_0^{**} = \sup \mu^{**}(\xi),$$

falls  $\xi$  alle Werte durchläuft, für die  $|\xi| + \frac{\sqrt{1-\xi^2}}{4n} \leq 1$  besteht. Dann folgt aus (31)

$$|\pi_n(x)| \leq 2\mu_0^{**} \quad (-1 \leq x \leq 1).$$

Man muß nur beachten, daß aus  $\xi = \cos \vartheta$ ,  $\xi' = \cos \vartheta'$  und  $|\xi - \xi'| \leq \frac{\sqrt{1-\xi^2}}{4n} = \frac{\sin \vartheta}{4n}$  ferner  $0 \leq \vartheta, \vartheta' \leq \pi$  folgt, daß  $|\vartheta - \vartheta'| \leq \frac{1}{2n}$  besteht.

Als erste Anwendung dieses Satzes betrachten wir eine Funktion  $f(x)$ , die den Forderungen des Satzes II genügt, also in  $[-1, +1]$  stetig ist, und in  $(\alpha, \beta)$  der Lipschitz—Dini Bedingung (14) genügt. Es gibt eine Polynomfolge  $\{q_n(x)\}$  (z. B. die Jacksonschen Approximationspolynome von  $f(x)$ ), wobei  $q_n(x)$  höchstens den Grad  $n$  hat, so daß die stetige Funktion  $f(x)$  in  $[-1, +1]$  durch die  $q_n(x)$  gleichmäßig approximiert wird, und in jedem inneren Teilintervall von  $(\alpha, \beta)$

$$(32) \quad f(x) - q_n(x) = o\left(\frac{1}{\log n}\right) \quad (\alpha < x < \beta)$$

besteht. Unter diesen Voraussetzungen werden wir  $q'_n(x)$  abschätzen. Zuerst sei  $x \in (\alpha, \beta)$ , dann folgt aus (14) und (32)

$$\begin{aligned} \left| \frac{q_n(x) - q_n(x_1)}{x - x_1} \right| &= \left| \frac{f(x) - f(x_1)}{x - x_1} \right| + \frac{1}{|x - x_1|} o\left(\frac{1}{\log n}\right) = \\ &= \frac{1}{|x - x_1|} \left\{ o\left(\log^{-1} \frac{1}{|x - x_1|}\right) + o\left(\frac{1}{\log n}\right) \right\}. \end{aligned}$$

Die Größenordnung der rechten Seite ist für ein beliebiges festes positives  $c_3$  und  $|x - x_1| > \frac{c_3}{n}$  gleich  $o\left(\frac{n}{\log n}\right)$ , also gilt infolge Satz IIIc, da

$\frac{q_n(x) - q_n(x_1)}{x - x_1}$  ein Polynom  $n-1$ -ten Grades in  $x_1$  ist,

$$(33) \quad q'_n(x) = o\left(\frac{n}{\log n}\right) \quad (\alpha < x < \beta)$$

gleichmäßig in jedem inneren Teilintervall von  $(\alpha, \beta)$ . Andererseits gilt für jedes  $x$  in  $[-1, +1]$

$$\left| \frac{q_n(x) - q_n(x_1)}{x - x_1} \right| = \left| \frac{f(x) - f(x_1)}{x - x_1} \right| + \frac{1}{|x - x_1|} o(1),$$



und somit erhalten wir aus den Sätzen IIIa und IIIc

$$(34) \quad q'_n(x) = \begin{cases} o\left(\frac{n}{\sqrt{1-x^2}}\right), \\ o(n^2) \end{cases}$$

gleichmäßig in  $[-1, +1]$ . (34) kann auch mit einer einfacheren eleganten Schlußweise erhalten werden, die zuerst von L. FEJÉR [10] in seiner Untersuchung über die Konjugierte einer trigonometrischen Reihe angewandt wurde. Doch kann in dieser Weise (33) nicht bewiesen werden.

#### IV. Abschätzungen für die Wellenparabel

HILFSSATZ IV. *Es gilt gleichmäßig in  $x$*

$$(35) \quad \sum_{|x-x_{kn}| < \delta} |x-x_{kn}| l_{kn}^2(x) = O\left(\frac{\log n}{n}\right)$$

für jeden festen inneren Teilintervall von  $(\alpha + 2\delta, \beta - 2\delta)$ .

BEWEIS. Wie ich in einer vorigen Arbeit zeigte [11], besteht die Identität

$$(36) \quad l_{kn}(x) = \frac{\alpha_{n-1}}{\alpha_n} \lambda_{kn} \frac{p_{n-1}(x_{kn}) p_n(x)}{x - x_{kn}},$$

wobei

$$(37) \quad 0 < \frac{\alpha_{n-1}}{\alpha_n} < 1$$

gilt.  $x_{jn}$  und  $x_{j+1, n}$  seien die zu  $x$  nächstliegenden Nullstellen von  $p_n(x)$ ; dann folgt aus (25) und (10)

$$(38) \quad \sum_{|x-x_{kn}| < \delta} |x-x_{kn}| l_{kn}^2(x) < O\left(\frac{1}{n^2}\right) \sum''_{|x-x_{kn}| < \delta} \frac{1}{|x-x_{kn}|} + \\ + |x-x_{jn}| l_{jn}^2(x) + |x-x_{j+1, n}| l_{j+1, n}^2(x).$$

Hier müssen bei der Summation  $\sum''$  die Glieder mit  $k=j$  und  $k=j+1$  weggelassen werden. Aus der linken Seite von (24) erhalten wir

$$\sum'' \frac{1}{|x-x_{kn}|} < 2 \sum_{v=1}^n \frac{1}{vc_1/n} = O(n \log n)$$

und aus (23) und der rechten Seite von (24)

$$|x-x_{jn}| l_{jn}^2(x) + |x-x_{j+1, n}| l_{j+1, n}^2(x) = O\left(\frac{1}{n}\right),$$

und hieraus folgt (35).

HILFSSATZ V. *Es sei*

$$(39) \quad |\gamma_{kn}| \leq c_4 \frac{n}{\sqrt{1-x_{kn}^2}} \quad \text{und} \quad |\gamma_{kn}| \leq c_4 n^2;$$



dann gibt es eine von  $n$  unabhängige Konstante  $c_5$ , so daß die Ungleichung

$$(40) \quad \left| \sum_{|x-x_{kn}|>\delta} \gamma_{kn} l_{kn}^2(x) \right| < \frac{c_4 c_5}{\delta^2}$$

besteht.

BEWEIS. Infolge (10), (36) und (37) gilt

$$\sum_{|x-x_{kn}|>\delta} |\gamma_{kn}| l_{kn}^2(x) \leq \frac{K^2}{\delta^2} \sum_{k=1}^n |\gamma_{kn}| \lambda_{kn}^2 p_{n-1}^2(x_{kn});$$

wegen (39) und Hilfssatz II ist  $|\gamma_{kn}| \lambda_{kn} \leq 200 M c_4$ ; somit erhalten wir

$$\begin{aligned} \sum_{|x-x_{kn}|>\delta} |\gamma_{kn}| l_{kn}^2(x) &\leq \frac{200 M K^2 c_4}{\delta^2} \sum_{k=1}^n \lambda_{kn} p_{n-1}^2(x_{kn}) = \\ &= 200 M K^2 \frac{c_4}{\delta^2} \int_{-1}^{+1} p_{n-1}^2(x) w(x) dx = \frac{200 M K^2 c_4}{\delta^2}. \end{aligned}$$

## V. Abschätzungen für die Treppenparabel

HILFSSATZ VI. Es gilt

$$(41) \quad \frac{|\lambda'_n(x_{kn})|}{\lambda_{kn}^2} \leq c_6 \frac{n^2}{1-x_{kn}^2}$$

und

$$(42) \quad \frac{|\lambda'_n(x_{kn})|}{\lambda_{kn}^2} \leq c_6 n^4,$$

wobei  $c_6$  von  $k$  und  $n$  unabhängig ist.

BEWEIS. Nach dem Lemma von P. ERDÖS und P. TURÁN ist

$$(43) \quad \frac{1}{\lambda_n(\xi)} = \sum_{k=0}^{n-1} p_k^2(\xi) \leq \frac{1}{m} \sum_{k=0}^{n-1} \frac{2k+1}{2} P_k^2(\xi),$$

wobei  $P_k(x)$  das  $k$ -te Legendresche Polynom bedeutet. Nach LAPLACE besteht die Abschätzung

$$(44) \quad |P_k(\xi)| \leq \frac{c_7}{k^2 (1-\xi^2)^{\frac{1}{4}}},$$

wobei  $c_7$  von  $k$  und  $\xi$  unabhängig ist. Also ist

$$(45) \quad 0 < \frac{1}{\lambda_n(\xi)} \leq \frac{c_8 n}{\sqrt{1-\xi^2}}.$$

Da  $\lambda_n^{-1}(\xi)$  ein Polynom  $2n-2$ -ten Grades ist, folgt aus dem am Anfang von Teil III erwähnten Bernsteinschen Satze

$$(46) \quad 0 < \frac{1}{\lambda_n(\xi)} \leq 2c_8 n^2$$



und durch Differenzieren, infolge des Satzes von MARKOFF

$$\frac{|\lambda'_n(\xi)|}{\lambda_n^2(\xi)} < 8c_8 n^4.$$

Damit ist (42) bewiesen. Andererseits erhalten wir aus (45) mit der Substitution  $\xi = \cos \Theta$

$$0 < \frac{\sin \Theta}{\lambda_n(\cos \Theta)} \leq c_8 n;$$

hier steht ein trigonometrisches Polynom  $2n-1$ -ten Grades, somit erhält man durch Differenzieren, infolge des Satzes von BERNSTEIN,

$$(47) \quad \left| -\frac{\lambda'_n(\cos \Theta)}{\lambda_n^2(\cos \Theta)} \sin^2 \Theta + \frac{\cos \Theta}{\lambda_n(\cos \Theta)} \right| \leq 2c_8 n^2.$$

Aus (46) und (47) folgt (41).

Es sei bemerkt, daß nach Hilfssatz IIIa (42) eine Folgerung von (41) ist.

Aus den Hilfssätzen II und VI folgt

$$(48) \quad \frac{|\lambda'_n(x_{kn})|}{\lambda_{kn}} \leq c_9 \text{Min} \left( \frac{n}{\sqrt{1-x_{kn}^2}}, n^2 \right).$$

HILFSSATZ VII. *Es gilt*

$$\sum_{|x-x_{kn}| \leq \delta} |v_{kn}(x)| l_{kn}^2(x) = O(\log n)$$

*gleichmäßig in x für  $(\alpha + 2\delta, \beta - 2\delta)$ .*

BEWEIS. Infolge (4) und (19) ist

$$(49) \quad v_{kn}(x) = 1 + \frac{\lambda'_n(x_{kn})}{\lambda_{kn}} (x - x_{kn}).$$

Wegen Hilfssatz I genügt es zu zeigen, daß

$$\sum_{|x-x_{kn}| < \delta} \frac{|\lambda'_n(x_{kn})|}{\lambda_{kn}} |x - x_{kn}| l_{kn}^2(x) = O(\log n)$$

besteht. Dies folgt aber aus (48) und Hilfssatz IV.

HILFSSATZ VIII. *Es gilt*

$$(50) \quad \sum_{|x-x_{kn}| > \delta} |v_{kn}(x)| l_{kn}^2(x) = O(1).$$

BEWEIS. Aus (48), (49) und Hilfssatz V folgt unmittelbar die Behauptung.



## VI. Beweis des Satzes II

Betrachten wir die Folge der approximierenden Polynome  $q_n(x)$ , die nach Teil III (32), (33) und (34) befriedigen. Es ist

$$(51) \quad H_n(f, x) - q_n(x) = \sum_{k=1}^n [f(x_{kn}) - q_n(x_{kn})] v_{kn}(x) l_{kn}^2(x) + \\ + \sum_{k=1}^n [d_{kn} - q'_n(x_{kn})] (x - x_{kn}) l_{kn}^2(x) = A_{1n} + A_{2n} + A_{3n} + A_{4n},$$

wobei

$$(52) \quad A_{1n} = \sum_{|x - x_{kn}| \leq \delta} [f(x_{kn}) - q_n(x_{kn})] v_{kn}(x) l_{kn}^2(x),$$

$$(53) \quad A_{2n} = \sum_{|x - x_{kn}| > \delta} [f(x_{kn}) - q_n(x_{kn})] v_{kn}(x) l_{kn}^2(x),$$

$$(54) \quad A_{3n} = \sum_{|x - x_{kn}| \leq \delta} [d_{kn} - q'_n(x_{kn})] (x - x_{kn}) l_{kn}^2(x),$$

$$(55) \quad A_{4n} = \sum_{|x - x_{kn}| > \delta} [d_{kn} - q'_n(x_{kn})] (x - x_{kn}) l_{kn}^2(x).$$

Nach Hilfssatz VII und (32) erhalten wir:  $A_{1n} = o(1)$ . Wegen Hilfssatz VIII und  $q_n(x) \Rightarrow f(x)$  erhalten wir  $A_{2n} = o(1)$ .

Wegen Hilfssatz IV, (11) und (33) gilt  $A_{3n} = o(1)$ . Endlich besteht  $A_{4n} = o(1)$ , wegen (11), (34) und Hilfssatz V. Somit besteht

$$H_n(f; x) - q_n(x) = A_{1n} + A_{2n} + A_{3n} + A_{4n} = o(1)$$

und hieraus folgt

$$H_n(f, x) \rightarrow f(x),$$

w. z. b. w.

VII. Genauere Abschätzung von  $\lambda'_n(\xi)$ 

Im folgenden sollen die Voraussetzungen des Satzes I befriedigt sein. Es sei  $-1 < \xi_1 < \xi_2 < 1$  und  $\varphi(x)$  die lineare Funktion, für welche  $\varphi(\xi_1) = \xi_2$  und  $\varphi(1) = 1$  ist:

$$\varphi(x) = 1 + \frac{1 - \xi_2}{1 - \xi_1} (x - 1).$$

In  $[-1, +1]$  ist  $\varphi(x) \leq 1$ , und der kleinste Wert von  $\varphi(x)$  ist

$$\eta = \varphi(-1) = -1 + \frac{2(\xi_2 - \xi_1)}{1 - \xi_1} > -1.$$

Es gilt weiter

$$(56) \quad 0 \leq \varphi(x) - x \leq \eta + 1 = \frac{2}{1 - \xi_1} (\xi_2 - \xi_1), \quad (-1 \leq x \leq +1).$$



Das Polynom  $l_n(\varphi(x), \xi_2)$ , dessen Grad gleich  $n-1$  ist, nimmt an der Stelle  $x = \xi_1$  den Wert  $l_n(\xi_2, \xi_2) = 1$  an. Hieraus erhalten wir infolge der Minimumeigenschaft von  $\lambda_n(\xi)$ :

$$(57) \quad \lambda_n(\xi_1) \leq \int_{-1}^{+1} l_n^2(\varphi(t), \xi_2) w(t) dt.$$

Es sei  $\Phi(x)$  die inverse Funktion von  $\varphi(x)$ :

$$(58) \quad \Phi(x) = 1 + \frac{1 - \xi_1}{1 - \xi_2} (x - 1).$$

Wegen (56) besteht

$$(59) \quad 0 \leq x - \Phi(x) \leq \frac{2}{1 - \xi_1} (\xi_2 - \xi_1) \quad (\eta \leq x \leq 1).$$

Mit der Substitution  $t = \Phi(x)$  folgt aus (57)

$$\lambda_n(\xi_1) \leq \frac{1 - \xi_1}{1 - \xi_2} \int_{\eta}^1 l_n^2(x, \xi_2) w(\Phi(x)) dx.$$

Hieraus folgt wegen (17)

$$(60) \quad \begin{aligned} \lambda_n(\xi_1) - \lambda_n(\xi_2) &\leq \int_{\eta}^1 l_n^2(x, \xi_2) [w(\Phi(x)) - w(x)] dx + \\ &+ \frac{\xi_2 - \xi_1}{1 - \xi_2} \int_{\eta}^1 l_n^2(x, \xi_2) w(\Phi(x)) dx. \end{aligned}$$

Ebenso sei  $\varphi^*(x)$  die lineare Funktion mit  $\varphi^*(-1) = -1$  und  $\varphi^*(\xi_2) = \xi_1$ , ferner sei  $\Phi^*(x)$  die inverse von  $\varphi^*(x)$ . Dann gilt in  $[-1, +1]$   $\varphi^*(x) \geq -1$ ,

$$\varphi^*(x) \leq \eta^* = \varphi^*(1) = 1 - \frac{2(\xi_2 - \xi_1)}{1 + \xi_2} < 1$$

und

$$(56a) \quad 0 \leq \Phi^*(x) - x \leq \frac{2}{1 + \xi_2} (\xi_2 - \xi_1).$$

Somit folgt mit derselben Schlußweise, wie vorher:

$$(60a) \quad \begin{aligned} \lambda_n(\xi_1) - \lambda_n(\xi_2) &\geq - \int_{-1}^{\eta^*} l_n^2(x, \xi_1) [w(\Phi^*(x)) - w(x)] dx - \\ &- \frac{\xi_2 - \xi_1}{1 + \xi_1} \int_{-1}^{\eta^*} l_n^2(x, \xi_1) w(\Phi^*(x)) dx. \end{aligned}$$

Man wähle  $\delta^*$  so klein, daß für  $|\xi_2 - x| < \delta^*$  auch  $|\Phi(x) - x| < \delta$  besteht. Dann gilt nach (12) und (56)

$$(61) \quad \max_{|x - \xi_2| < \delta^*} |w(\Phi(x)) - w(x)| = o\left(\log^{-1} \frac{1}{\xi_2 - \xi_1}\right).$$



Infolge der Stetigkeit von  $w(x)$  gilt weiter

$$(62) \quad w(\Phi(x)) - w(x) = o(1) \text{ für } \xi_2 - \xi_1 \rightarrow 0$$

gleichmäßig in  $x$  ( $\eta \leq x \leq 1$ ).

Wegen (60), (61) und (62) gilt für  $|\xi_2| < 1 - h$

$$(63) \quad \begin{aligned} \lambda_n(\xi_1) - \lambda_n(\xi_2) &\leq o\left(\log^{-1} \frac{1}{|\xi_2 - \xi_1|}\right) \int_{\xi_2 - \delta^*}^{\xi_2 + \delta^*} l_n^2(x, \xi_2) dx + \\ &+ o(1) \left\{ \int_{-1}^{\xi_2 - \delta^*} l_n^2(x, \xi_2) dx + \int_{\xi_2 + \delta^*}^1 l_n^2(x, \xi_2) dx \right\} + \\ &+ \frac{M(\xi_2 - \xi_1)}{h} \int_{-1}^{+1} l_n^2(x, \xi_2) dx. \end{aligned}$$

Aus  $w(x) \geq m$  und Hilfssatz II ergibt sich

$$(64) \quad \int_{-1}^{+1} l_n^2(x, \xi_2) dx \leq \frac{1}{m} \int_{-1}^{+1} l_n^2(x, \xi_2) w(x) dx = \frac{1}{m} \lambda_n(\xi_2) \leq \frac{c_{10}}{n}.$$

Um den zweiten Teil von (63) abzuschätzen, benötigen wir eine Verallgemeinerung von (36):

$$(65) \quad \begin{aligned} L_n(x, \xi) &= \lambda_n(\xi) \sum_{k=0}^{n-1} p_k(\xi) p_k(x) = \\ &= \frac{\alpha_{n-1} \lambda_n(\xi) p_{n-1}(\xi) p_n(x) - p_n(\xi) p_{n-1}(x)}{\alpha_n x - \xi} \end{aligned}$$

Der Beweis kann aus meiner Arbeit [11] wörtlich übernommen werden.

Angesichts (65) erhalten wir wegen (37) und (10)

$$(66) \quad \begin{aligned} \int_{-1}^{\xi_2 - \delta^*} l_n^2(x, \xi_2) dx &\leq \frac{2}{\delta^{*2} m} \lambda_n^2(\xi_2) \left[ p_{n-1}^2(\xi_2) \int_{-1}^{+1} p_n^2(x) w(x) dx + \right. \\ &\left. + p_n^2(\xi_2) \int_{-1}^{+1} p_{n-1}^2(x) w(x) dx \right] \leq \frac{4K^2}{\delta^{*2} m} \lambda_n^2(\xi_2) \leq \frac{c_{11}}{n^2}, \end{aligned}$$

und ebenso kann das Integral  $\int_{\xi_2 + \delta^*}^1 l_n^2(x, \xi_2) dx$  abgeschätzt werden. Also besteht auf Grund von (63) und (64)

$$(67) \quad \lambda_n(\xi_1) - \lambda_n(\xi_2) \leq \frac{1}{n} o\left(\log^{-1} \frac{1}{|\xi_2 - \xi_1|}\right) + \frac{1}{n^2} o(1).$$

Aus (60a) folgt eine ähnliche untere Abschätzung für  $\lambda_n(\xi_1) - \lambda_n(\xi_2)$ , und in diesen beiden Abschätzungen kann auch die Bedingung  $\xi_1 < \xi_2$  fallen gelassen werden, wenn man neben  $\xi_2 \in (\alpha + \delta, \beta - \delta)$  auch  $\xi_1 \in (\alpha + \delta, \beta - \delta)$  voraussetzt.



Somit erhalten wir wegen (22)

$$(68) \quad \frac{1}{\lambda_n(\xi_1)} - \frac{1}{\lambda_n(\xi_2)} = n o \left( \frac{\log^{-1} \frac{1}{|\xi_2 - \xi_1|}}{|\xi_2 - \xi_1|} \right).$$

HILFSATZ IX. Es besteht

$$(69) \quad \frac{\lambda'_n(\xi)}{\lambda_n^2(\xi)} = o \left( \frac{n^2}{\log n} \right)$$

gleichmäßig in  $\xi$  in jedem inneren Teilintervall von  $(\alpha, \beta)$ .

BEWEIS. Die linke Seite von (68) ist infolge (18) ein Polynom in  $\xi_2$ , dessen Grad niedriger als  $2n$  ist. Somit erhalten wir aus Hilfssatz IIIc, (68) und (22):

$$\frac{1}{\lambda_n(\xi_1)} - \frac{1}{\lambda_n(\xi_2)} = o \left( \frac{n^2}{\log n} \right).$$

Setzt man hier  $\xi_1 = \xi_2 = \xi$ , dann bekommt man die Abschätzung (69).

HILFSATZ X. Es gilt

$$(70) \quad \frac{\lambda'_n(\xi)}{\lambda_n^2(\xi)} = o(n^2)$$

gleichmäßig in jedem inneren Teilintervall von  $-1, +1$ .

BEWEIS. Aus (60) und (60a) folgt, indem man (62) und (64) beachtet,

$$(71) \quad |\lambda_n(\xi) - \lambda_n(\xi_2)| \leq \frac{c_{12}}{n} [|\xi_2 - \xi_1| + o(1)]$$

gleichmäßig für  $-1 + h \leq \xi_1 < \xi_2 \leq 1 - h$ . In dieser Abschätzung kann wieder die Bedingung  $\xi_1 < \xi_2$  fallen gelassen werden. Also besteht wegen (22)

$$(72) \quad \frac{1}{\lambda_n(\xi_2)} - \frac{1}{\lambda_n(\xi_1)} = \left[ 1 + o \left( \frac{1}{|\xi_2 - \xi_1|} \right) \right] O(n).$$

Hier steht an der linken Seite wieder ein Polynom vom Grade  $< 2n$ . Man erhält also (70) aus Hilfssatz IIIc.

### VIII. Genauere Abschätzungen für die Treppenparabel

HILFSATZ XI. Es gilt gleichmäßig in  $x$  für jedes innere Teilintervall von  $(\alpha, \beta)$

$$(73) \quad \sum_{\substack{|x - x_{kn}| > \delta \\ |x_{kn}| < 1 - h}} |v_{kn}(x)| l_n^2(x) = o(1).$$



BEWEIS. Wegen Hilfssatz X und II ist

$$\frac{\lambda'_n(\xi)}{\lambda_n(\xi)} = o(n) \quad (-1+h \leq \xi \leq 1-h),$$

also nach (49)

$$(74) \quad v_{kn}(x) = o(n) \quad (-1+h \leq x_{kn} \leq 1-h).$$

Aus (36), (37), (10) erhalten wir, falls  $|x - x_{kn}| > \delta$  ist, daß

$$(75) \quad |l_{kn}(x)| < \frac{K|p_{n-1}(x_{kn})|}{|x - x_{kn}|} < \frac{K}{\delta} \lambda_{kn} |p_{n-1}(x_{kn})|.$$

Aus (74) und (75) erhält man also

$$(76) \quad \sum_{\substack{|x-x_{kn}| > \delta \\ |x_{kn}| < 1-h}} |v_{kn}(x)| l_{kn}^2(x) = o(n) \frac{K^2}{\delta^2} \text{Max } \lambda_{kn} \sum_{k=1}^n \lambda_{kn} p_{n-1}^2(x_{kn}).$$

Aus Hilfssatz II schließt man  $\lambda_{kn} = O\left(\frac{1}{n}\right)$ , ferner ist

$$\sum_{k=1}^n \lambda_{kn} p_{n-1}^2(x_{kn}) = \int_{-1}^{+1} p_{n-1}^2(x) w(x) dx = 1,$$

somit folgt (73) aus (76).

HILFSSATZ XII. *Es gilt gleichmäßig für jedes innere Teilintervall von  $(\alpha + 2\delta, \beta - 2\delta)$*

$$(77) \quad \sum_{|x-x_{kn}| < \delta} |v_{kn}(x)| l_{kn}^2(x) = O(1).$$

BEWEIS. Wegen Hilfssatz I und (49) genügt es zu zeigen, daß

$$(78) \quad \sum_{|x-x_{kn}| < \delta} \frac{|\lambda'_n(x_{kn})|}{\lambda_{kn}} |x - x_{kn}| l_{kn}^2(x) = O(1)$$

besteht. Aus den Hilfssätzen IX und II folgt

$$\frac{\lambda'_n(x_{kn})}{\lambda_{kn}} = o\left(\frac{n}{\log n}\right),$$

also folgt (78) aus Hilfssatz IV.

### IX. Beweis des Satzes I

Die Funktion  $f(x)$  genüge den Bedingungen des Satzes I. Es sei  $\chi(x)$  das Polynom zweiten Grades, das an den drei Stellen  $\xi, -1, +1$  mit  $f(x)$  übereinstimmt.

Infolge der Stetigkeit von  $f(x)$  und  $\chi(x)$  an den Stellen  $\xi, -1, +1$  kann man zu einem beliebigen positiven  $\varepsilon$  positive Zahlen  $\delta$  und  $h$  finden derart, daß



$$(79) \quad |f(x) - \chi(x)| \leq \varepsilon \quad \text{für } |x - \xi| \leq \delta,$$

$$(80) \quad |f(x) - \chi(x)| \leq \varepsilon \quad \text{für } 1 - h \leq |x| \leq 1.$$

Es kann vorausgesetzt werden, daß  $\alpha \leq \xi - \delta < \xi + \delta \leq \beta$  ist.

Das  $n$ -te Hermite—Fejérsche Interpolationspolynom zerlegen wir folgenderweise:

$$(81) \quad H_n(f, \xi) = \sum_{k=1}^n f(x_{kn}) v_{kn}(\xi) l_{kn}^2(\xi) + \sum_{k=1}^n d_{kn} (\xi - x_{kn}) l_{kn}^2(\xi) = \chi(\xi) + \\ + \sum_{k=1}^n [f(x_{kn}) - \chi(x_{kn})] v_{kn}(\xi) l_{kn}^2(\xi) = f(\xi) + \sum_{1n} + \sum_{2n} + \sum_{3n} + \sum_{4n},$$

wobei

$$(82) \quad \sum_{1n} = \sum_{|\xi - x_{kn}| \leq \delta} [f(x_{kn}) - \chi(x_{kn})] v_{kn}(\xi) l_{kn}^2(\xi),$$

$$(83) \quad \sum_{2n} = \sum_{\substack{|\xi - x_{kn}| > \delta \\ |x_{kn}| \leq 1-h}} [f(x_{kn}) - \chi(x_{kn})] v_{kn}(\xi) l_{kn}^2(\xi),$$

$$(84) \quad \sum_{3n} = \sum_{|x_{kn}| > 1-h} [f(x_{kn}) - \chi(x_{kn})] v_{kn}(\xi) l_{kn}^2(\xi)$$

und endlich

$$(85) \quad \sum_{4n} = \sum_{k=1}^n [d_{kn} - \chi'(x_{kn})] (\xi - x_{kn}) l_{kn}^2(\xi).$$

Aus Hilfssatz XII und (79) erhalten wir

$$(86) \quad \left| \sum_{1n} \right| < \varepsilon O(1).$$

Hilfssatz XI zieht

$$(87) \quad \sum_{2n} = o(1)$$

nach sich; weiter folgt aus Hilfssatz VIII und (80)

$$(88) \quad \left| \sum_{3n} \right| < \varepsilon O(1);$$

zuletzt erhalten wir aus Hilfssatz IV und V, infolge (11),

$$(89) \quad \sum_{4n} = o(1).$$

Somit besteht wegen (81), (86), (87), (88) und (89)

$$H_n(f; \xi) \rightarrow f(\xi),$$

w. z. b. w.

(Eingegangen am 18. Februar 1954.)



## Literaturverzeichnis

- [1] S. N. BERNSTEIN, Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné, *Mémoires Ac. de Belgique*, II. 4 (1912), S. 1—103.
- [2] S. N. BERNSTEIN, Sur les polynomes orthogonaux relatifs à un segment fini, *Journal de Mathématiques*, 9(1930), S. 127—177, und 10 (1921), S. 219—286.
- [3] P. ERDŐS—P. TURÁN, On interpolation. III, *Annals of Math.*, 41 (1940), S. 510—553.
- [4] L. FEJÉR, Über Interpolation, *Nachrichten der Ges. der Wiss. Göttingen*, Math.—phys. Kl., 1916, S. 66—91.
- [5] L. FEJÉR, Über Weierstraßsche Approximation, besonders durch Hermitesche Interpolation, *Math. Annalen*, 102 (1930), S. 707—725.
- [6] L. FEJÉR, Die Abschätzung eines Polynoms, *Math. Zeitschr.*, 32 (1930), S. 426—457.
- [7] L. FEJÉR, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Annalen*, 106 (1932), S. 1—55.
- [8] L. FEJÉR, Mechanische Quadraturen mit positiven Cotesschen Zahlen, *Math. Zeitschr.*, 37 (1933), S. 287—309.
- [9] L. FEJÉR, On the characterization of some remarkable systems of points of interpolation by means of conjugate points, *American Math. Monthly*, 41 (1934), S. 1—14.
- [10] L. FEJÉR, Über konjugierte trigonometrische Reihen, *Journal f. reine u. angew. Math.*, 144 (1914), S. 48—56.
- [11] G. FREUD, Über die Lebesgueschen Funktionen der Lagrangeschen Interpolation *Acta Math. Acad. Sci. Hung.*, 4 (1953), S. 137—142.
- [12] G. FREUD, Über einen Satz von P. Erdős und P. Turán, *Acta. Math. Acad. Sci. Hung.*, 4 (1953), S. 255—266.
- [13] G. GRÜNWARD, On the theory of interpolation, *Acta. Math.*, 75 (1943), S. 219—245.
- [14] D. JACKSON, *The theory of approximation*, Amer. Math. Soc. Coll. Publ., Band, XI (New York, 1930).
- [15] J. SHOHAT, On interpolation, *Annals of Math. (2)*, 34 (1933), S. 130—146.
- [16] G. SZEGŐ, Über gewisse Interpolationspolynome, die zu den Jacobischen und Laguerreschen Abszissen gehören, *Math. Zeitschr.*, 35 (1932), S. 579—602.
- [17] G. SZEGŐ, *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ., Band XXIII (New York, 1939).



## О СХОДИМОСТИ ИНТЕРПОЛЯЦИОННОГО ПРОЦЕССА ЭРМИТА—ФЕЙЕРА

Г. ФРАЙД (Будапешт)

(Резюме)

Пусть  $\{p_n(x)\}$  — последовательность ортогональных полиномов, соответствующих весовой функции  $w(x)$ . Предположим, что в подинтервале  $[\alpha, \beta]$  интервала ортогональности  $(-1, +1)$  эти ортогональные полиномы равномерно ограничены. Пусть корни полиномов  $p_n(x)$  суть  $x_{1n}, x_{2n}, \dots, x_{nn}$ . Рассмотрим интерполяционный процесс Эрмита—Фейера, соответствующий системе основных точек  $\{x_{kn}\}$ :

$$H_n(f; x) = \sum_{k=1}^n [f(x_{kn}) V_{kn}(x) l_{kn}^2(x) + d_{kn}(x - x_{kn}) \cdot l_{kn}^2(x)],$$

где

$$d_{kn} = \begin{cases} o\left(\frac{n}{\log n}\right) & \text{если } \alpha \leq x_{kn} < \beta, \\ o\left[\text{Min}\left(\frac{n}{\sqrt{1-x_{kn}^2}}, n^2\right)\right] & \text{для всех } x_{kn}. \end{cases}$$

**Т е о р е м а 1.** Пусть  $w(x)$  — положительная и непрерывная в  $[-1, +1]$  функция, и пусть она внутри интервала  $(\alpha, \beta)$  удовлетворяет условию Липшица—Дини

$$w(x_1) - w(x_2) = o\left(\log^{-1} \frac{1}{|x_2 - x_1|}\right).$$

Если  $f(x)$  непрерывна в концах  $(-1, +1)$  интервала, то функция  $H_n(f; x)$  стремится к  $f(x)$  во всякой точке непрерывности этой функции, лежащей внутри  $(\alpha, \beta)$ . Если  $f(x)$  непрерывна в  $(\alpha, \beta)$ , то сходимость равномерна во всяком внутреннем подинтервале  $(\alpha, \beta)$ .

**Т е о р е м а 2.** Пусть в интервале  $[-1, +1]$

$$0 < m \leq w(x) \leq M.$$

Пусть функция  $f(x)$  непрерывна в  $(-1, +1)$ , и пусть она удовлетворяет условию Липшица—Дини

$$f(x_1) - f(x_2) = o\left(\log^{-1} \frac{1}{|x_2 - x_1|}\right)$$

равномерно в интервале  $(\alpha, \beta)$ . Тогда  $H_n(f, x)$  стремится к  $f(x)$  равномерно во всяком подинтервале интервала  $(\alpha, \beta)$ .



# ON THE BASIC SUBGROUPS OF ABELIAN $p$ -GROUPS

By

T. SZELE (Debrecen)

(Presented by L. RÉDEI)

## § 1. Introduction

It is well known that the PRÜFER—ULM—ZIPPIN theory furnishes a complete solution of the structure problem of countable abelian  $p$ -groups or, what is essentially the same, that of countable abelian torsion groups. Since, however, this theory cannot be extended to the case of abelian  $p$ -groups of arbitrary cardinality [1], [5], [6], [7], [8], [9], [10],<sup>1</sup> recently several investigations were carried out in other directions, in order to obtain certain structural invariants of arbitrary abelian  $p$ -groups. One of the most important such structural invariants is the *basic subgroup* discovered independently by L. KULIKOV and L. KALOUJNINE [6], [3]. The basic subgroup  $B$  of an abelian  $p$ -group  $G$  is a direct sum of cyclic groups and it is uniquely determined by  $G$  up to isomorphism. Thus the basic subgroup is a structural invariant of the group which, although discovered not long ago, has already proved very fruitful in the theory of abelian  $p$ -groups. As important examples we mention the classification of all abelian  $p$ -groups without elements of infinite height due to L. KULIKOV and L. KALOUJNINE [6], [3] (see also L. FUCHS [1]) which is “almost complete” in the sense that the only gap in it is the lack of the solution of the isomorphism problem; moreover another important application of the theory of the basic subgroup is the generalization of ZIPPIN’s famous existence theorem from the countable case to the case of arbitrary  $p$ -groups due to L. FUCHS and L. KULIKOV [1], [7], [8].

The major goal of the present paper is to prove that a basic subgroup of the group  $G$  is always a homomorphic image of  $G$ . This result settles a problem stated by L. FUCHS [2] and, more generally, answers the question whether or not a prescribed group  $H$ , a direct sum of cyclic  $p$ -groups, is a homomorphic image of a given abelian  $p$ -group  $G$ . In particular, we obtain that every countable (or finite) abelian  $p$ -group is a homomorphic image of any infinite abelian  $p$ -group  $G$  unless  $G$  is a direct sum of cyclic  $p$ -groups of bounded order and of quasicyclic groups of type  $(p^\infty)$ .

<sup>1</sup> The numbers in brackets refer to the Bibliography given at the end of this paper.



The definition and fundamental properties of the basic subgroup can be found in [1], [3], [6], [9]. Although all these treatments are very clear and simple, we give here a new method of introducing the concept of basic subgroup; this method seems to be still simpler inasmuch as it is based on the quite elementary concept of "regular" group, but makes no use of the notion of serving subgroups. Namely, I have proved recently a theorem on direct decomposibility of abelian groups [11] which furnishes several important results as immediate corollaries. Now, it will turn out that the theory of the basic subgroup can also be obtained as a simple corollary of this theorem. In this treatment a basic subgroup  $B$  of an abelian  $p$ -group  $G$  appears as a direct sum

$$(1) \quad B = B_1 + B_2 + \cdots + B_n + \cdots$$

the "partial sums"  $B_1 + B_2 + \cdots + B_n$  of which are maximal  $p^n$ -bounded direct summands of  $G$ ; here  $B_n$  denotes a regular  $p^n$ -bounded group, i. e. a direct sum of cyclic groups of order  $p^n$ . However, the group  $B$  itself is no direct summand of  $G$  in general. It is a direct summand if and only if  $G = A + B$  holds with an algebraically closed (or, in another terminology, complete) group  $A$  which is therefore a direct sum of groups of type  $(p^\infty)$ . Thus  $G$  possesses a basic subgroup, which is a direct summand of  $G$ , if and only if  $G$  is fully decomposable. (A group is said to be fully decomposable if it can be represented as a direct sum of directly indecomposable groups.) Another criterion for this statement was found recently by A. KERTÉSZ [4].

## § 2. Preliminaries

By a group we shall always mean an additively written abelian  $p$ -group. Groups will be denoted by Latin capital letters and their elements by  $x, a, b, c, d$ . The other small Latin letters serve to denote rational integers (in particular  $p$  a prime number). By  $nG$  we denote the set of elements of the form  $na$  ( $a \in G$ ), by  $O(a)$  the order of an element  $a$  of a group. The symbol  $\{a, \dots, A, \dots\}$  denotes the group generated by the elements  $a, \dots$  and the groups  $A, \dots$ .  $C(p^k)$  denotes the cyclic group of order  $p^k$  for a natural number  $k$ , while  $C(p^\infty)$  is PRÜFER's group of type  $(p^\infty)$  (i. e. the additive group mod 1 of all rational numbers with  $p$ -power denominators).  $A + B$  denotes the direct sum of the groups  $A$  and  $B$ . An abelian group  $A$  is called *algebraically closed* (or in another terminology, complete) if  $nA = A$  holds for each natural number  $n$ . It is well known that the algebraically closed abelian  $p$ -groups coincide with those groups which admit a decomposition into the direct sum of groups  $C(p^\infty)$ . (See e. g. [9].) The maximal algebraically closed subgroup  $A$  of a group  $G$  is uniquely determined as the union of all algebraically closed subgroups in  $G$ . Since, by a well-known theorem of R. BAER, every algebrai-



cally closed subgroup of  $G$  is a direct summand of  $G$ , we have a direct decomposition

$$(2) \quad G = A + U$$

where the subgroup  $U$  contains no algebraically closed subgroup  $\neq 0$ . Such a group is called a *reduced group*. While  $A$  is invariantly defined by  $G$ , the reduced subgroup  $U$  in (2) is determined merely up to an isomorphism. However, we may refer to  $U$  as *the reduced part of  $G$* .

A subset  $(a_i)$  of  $G$ , not containing 0, is said to be *independent* if for any finite subset  $a_1, \dots, a_j$  of the system a relation  $m_1 a_1 + \dots + m_j a_j = 0$  implies  $m_1 a_1 = \dots = m_j a_j = 0$ . Since the independence so defined is a property of finite character, it follows that every group  $G$  contains a maximal independent system of elements. If  $G$  is an abelian  $p$ -group, then the cardinality of a maximal independent system in  $G$  is called the *rank* of  $G$ . The rank is an invariant of  $G$  and it can be defined also as the dimensionality of  $G_p$ , regarded as a vector space over the prime field of characteristic  $p$ ,  $G_p$  being the subgroup of elements of order  $\leq p$  in  $G$ . If  $G$  is a direct sum of cyclic  $p$ -groups, then the rank of  $G$  obviously coincides with the cardinal number of the cyclic direct components. On the other hand, it is easy to see that the rank of  $G$  is equal to the power of the set  $G$  provided that the rank is infinite. We denote in the sequel the rank of  $G$  by the symbol:  $\text{rank}(G)$ .

The *height* in  $G$  of an element  $a \neq 0$  is defined as follows. If there exists a maximal non-negative integer  $n = h$  for which the equation  $p^n x = a$  is solvable in  $G$ , then  $a$  is said to have the height  $h$ ; in the contrary case, i. e., if  $p^n x = a$  can be solved in  $G$  for each natural number  $n$ , we say that  $a$  is an element of infinite height. We emphasize that the height of an element depends on the group relative to which it is understood; if  $H$  is a subgroup of  $G$  containing the element  $a$ , then the height of  $a$  relative to  $H$  is smaller than or equal to that relative to  $G$ . On the other hand, if  $G \sim \bar{G}$ , then the homomorphic image  $\bar{a}$  of  $a$  under this homomorphism possesses a height relative to  $\bar{G}$  greater than or equal to the height of  $a$  relative to  $G$ . In what follows, we make use also of the simple but important fact that if  $G = \Sigma C(p^n)$  (read:  $G$  is a direct sum of cyclic groups of order  $p^n$ ,  $n$  fixed), then for each non-zero element  $a \in G$  of order  $p^k$  the sum  $k + h$  of  $k$  and of the height  $h$  (in  $G$ ) of  $a$  is equal to  $n$ .

In order to formulate the following Lemma, let  $G$  be for the moment an arbitrary abelian group. If  $G$  contains only elements of finite order and there exists an element of maximal order  $r$  in  $G$ , then we say that  $G$  is an  *$r$ -bounded group*. The  $r$ -bounded group  $G$  we call *regular* if

$$(3) \quad a \in \frac{r}{O(a)} G$$

for each element  $a \in G$ . It is obvious that a regular  $r$ -bounded group is a



direct sum of regular  $p$ -groups and our following Lemma shows immediately that the regular  $p^k$ -bounded groups coincide with the groups admitting a decomposition into a direct sum of cyclic groups of order  $p^k$ . [Indeed, choose  $H$  as a subgroup generated by a maximal independent set of elements of order  $p^k$  in  $G$ ,  $G$  being an arbitrary regular  $p^k$ -bounded group; then the Lemma implies  $G=H$ .] Now we can formulate our Lemma which will serve as a basis for the following development of the theory of the basic subgroup:

LEMMA. *A regular  $r$ -bounded subgroup  $H$  of an arbitrary abelian group  $G$  is a direct summand of  $G$  if and only if*

$$(4) \quad H \cap rG = 0.$$

A short and simple proof of this Lemma can be found in [11].

### § 3. Definition and fundamental properties of the basic subgroup

Let  $G$  be an arbitrary abelian  $p$ -group (in the sequel briefly: "a group"). We shall show that  $G$  contains a maximal direct summand admitting a decomposition into a direct sum of cyclic groups only in the case when  $G$  is fully decomposable, i. e.,

$$G = \sum C(p^{m_r}) \quad (1 \leq m_r \leq \infty).$$

Now we find a quite different situation concerning the existence of a maximal  $p^n$ -regular direct summand. Namely, it follows immediately from the above Lemma that  $G$  contains a maximal  $p^n$ -regular direct summand for each fixed natural number  $n$ , since the  $p^n$ -regularity as well as the property (4) of a subgroup  $H$  in  $G$  are properties of inductive character. (ZORN'S lemma!) Making use of this statement for  $n=1, 2, 3, \dots$  we have successively:  $G = B_1 + D_1 = B_1 + B_2 + D_2 = \dots$ ,

$$(5) \quad G = B_1 + B_2 + \dots + B_n + D_n$$

where  $B_1$  is a maximal  $p$ -regular direct summand of  $G$ ,  $B_2$  is a maximal  $p^2$ -regular direct summand of  $D_1$ , and so on. Thus the subgroup (1) of  $G$  so obtained is a direct sum of cyclic groups, in particular, the component  $B_n$  is a direct sum of groups  $C(p^m)$ , and  $B$  is, by definition, a maximal subgroup of  $G$  such that it is "partial-wise" direct summand of  $G$ . This means that the "partial sums"

$$(6) \quad S_n = B_1 + B_2 + \dots + B_n$$

are maximal  $p^n$ -bounded direct summands of  $G$ , i. e.,  $G$  contains no direct summand of type  $S_n + C(p^m)$  with  $1 \leq m \leq n$ . Any subgroup  $B$  of  $G$  with the above properties will be called a *basic subgroup* of  $G$ . Thus we have the following

THEOREM 1. *Every abelian  $p$ -group  $G$  contains a basic subgroup, i. e. a subgroup (1) where  $B_n$  denotes an (eventually empty) direct sum of cyclic*



groups of order  $p^n$ , such that (6) is a maximal  $p^n$ -bounded direct summand of  $G$  for each  $n$ . Any two basic subgroups of  $G$  are isomorphic.

We have only to prove the isomorphism of the basic subgroups. This can be shown easily using the following statement:

(7) Each element  $d \in D_n$  of order  $p$  has a height<sup>2</sup>  $h \geq n$ .

For let  $d \in D_n$  be an element of order  $p$  and of height  $h$ , and let  $d = p^h d'$  ( $d' \in D_n$ ). Then  $\{d'\} \cap p^{h+1} D_n = 0$ , so that we have by the above Lemma:  $D_n = \{d'\} + F_n = C(p^{h+1}) + F_n$ . But since (6) is a maximal  $p^n$ -bounded direct summand of  $G$ , this is possible only if  $h+1 > n$ , i. e.,  $h \geq n$ .

Now let us suppose that  $B' = B'_1 + B'_2 + \dots + B'_n + \dots$  is another basic subgroup of  $G$  with

$$(8) \quad G = B'_1 + B'_2 + \dots + B'_n + D'_n.$$

We infer to the isomorphism of the  $p^n$ -regular groups  $B_n$  and  $B'_n$  from the isomorphism of the elementary  $p$ -groups  $p^{n-1} B_n$ ,  $p^{n-1} B'_n$ . Now we have by (5) and (8)

$$(9) \quad p^{n-1} B_n + p^{n-1} D_n = p^{n-1} B'_n + p^{n-1} D'_n.$$

By the symmetry it is sufficient to show that  $p^{n-1} B_n$  contains at least as many independent elements as the cardinality of a basis  $(a'_1, a'_2, \dots)$  of  $p^{n-1} B'_n$  in a representation  $p^{n-1} B'_n = \{a'_1\} + \{a'_2\} + \dots$ . (This implies obviously the equality of the ranks of  $p^{n-1} B_n$  and  $p^{n-1} B'_n$ , i. e., the isomorphism of these groups.) Now according to (9) let

$$a_v + d_v = a'_v \quad (a_v \in p^{n-1} B_n, d_v \in p^{n-1} D_n).$$

The independence of the system  $(a_1, a_2, \dots)$  in  $p^{n-1} B_n$  follows from the fact that a relation

$$(10) \quad 0 = r_1 a_1 + r_2 a_2 + \dots = r_1 (a'_1 - d_1) + r_2 (a'_2 - d_2) + \dots$$

would imply

$$(11) \quad r_1 a'_1 + r_2 a'_2 + \dots = r_1 d_1 + r_2 d_2 + \dots = d$$

with  $d \in p^{n-1} D_n$ ,  $O(d) = p$ , i. e., by (7),

$$d = p^n d' = p(p^{n-1} d') = p d''$$

with  $d'' \in p^{n-1} D_n$ . But this means  $d = 0$ , i. e., (see (11))  $p|r_1, p|r_2, \dots$  in (10), since the elementary  $p$ -group  $p^{n-1} B'_n$  in (9) contains no element  $\neq 0$  of positive height. This completes the proof of Theorem 1.

EXAMPLE. Let  $\{c_n\}$  be a cyclic group of order  $p^n$ . We consider the complete direct sum of the groups  $\{c_1\}, \{c_2\}, \dots, \{c_n\}, \dots$ , i. e. the set of all "vectors"

$$(12) \quad a = \langle m_1 c_1, m_2 c_2, \dots, m_n c_n, \dots \rangle \quad (m_n \text{ rational integers})$$

which are added component-wise. The elements of finite order of this group

<sup>2</sup> Equation (5) shows that the height of an element  $d \in D_n$  relative to  $G$  is equal to that relative to  $D_n$ .



form an abelian  $p$ -group  $G^*$  of the power of the continuum. By the natural correspondence

$$c_n \longleftrightarrow \langle 0, 0, \dots, c_n, 0, \dots \rangle$$

the (discrete) direct sum  $B^* = \sum_{n=1}^{\infty} \{c_n\}$  can be embedded in  $G^*$ , and we can easily verify that  $B^*$  is a basic subgroup of  $G^*$ . For this purpose we take into account the direct representation

$$(13) \quad G^* = \{c_1\} + \{c_2\} + \dots + \{c_n\} + D_n^*$$

(where  $D_n^*$  is the set of all vectors (12) in  $G^*$  with  $m_1 = \dots = m_n = 0$ ) and show that  $G^*$  has no direct summand of type

$$(14) \quad \{c_1\} + \dots + \{c_n\} + \{c\}$$

with  $O(c) \leq p^n$ . As a matter of fact, by (13) we have

$$c = m_1 c_1 + \dots + m_n c_n + d_n$$

with  $d_n \in D_n^*$ . But  $d_n \in D_n^*$  means that for the order  $p^k$  of  $d_n$  resp. the height  $h$  of  $d_n$  (in  $G^*$ ) we have  $k + h > n$ , and an element  $d_n$  with this property cannot exist in a  $p^n$ -bounded direct summand (14) of  $G^*$ . Since any two basic subgroups of  $G^*$  are isomorphic, we have obtained also the result that  $G^*$  cannot be decomposed into a direct sum of cyclic groups.

An important property of the basic subgroup is expressed by

**THEOREM 2.** *If  $B$  is a basic subgroup of the abelian  $p$ -group  $G$ , then the factorgroup  $G/B$  is algebraically closed.*

As a matter of fact, let the coset  $B+d$  be an arbitrary element of order  $p$  in  $G/B$ . Then  $pd \in B$ , i. e.  $pd \in S_n$  for some  $n$ , and so we have, by (5), (6),  $pd = pb$  ( $b \in S_n \subseteq B$ ). Hence  $d$  can be replaced by the element  $d-b$  of order  $p$  and therefore we can assume that  $O(d) = p$ . But then we get from (5) the representation

$$d = b_1 + b_2 + \dots + b_n + d_n \quad (b_i \in B_i, d_n \in D_n)$$

with  $O(d_n) = p$  or  $d_n = 0$  whence it follows by (7) that the height of the element  $B+d$  in  $G/B$  is  $\geq n$ . Since this is true for  $n = 1, 2, 3, \dots$ , we have proved that each element of order  $p$  in  $G/B$  has infinite height, i. e.,  $G/B$  is an algebraically closed group.<sup>3</sup>

Now let us suppose that the basic subgroup  $B$  of  $G$  is a direct summand of  $G$ , i. e.,  $G = A + B$ . Then by Theorem 2  $A \cong G/B$  is an algebraically closed group, i. e., a direct sum of groups  $C(p^\infty)$ . So we have proved

<sup>3</sup>) In fact, if every element of order  $p$  in a group  $H$  has infinite height, then  $pH = H$  holds. Suppose, namely, this is not true and let  $a$  be an element of minimal order  $p^k$  such that the equation  $px = a$  cannot be solved in  $H$ . Since the element  $p^{k-1}a$  is of infinite height we have  $p^{k-1}a = p^k x'$  for some  $x' \in H$ , i. e.,  $p^{k-1}(a - px') = 0$ . But then  $a' = a - px'$  is an element of order  $< p^k$  not divisible by  $p$  in  $H$ , a contradiction.



that an abelian  $p$ -group  $G$  possesses a basic subgroup, which is a direct summand of  $G$ , if and only if  $G$  is fully decomposable, i. e.,  $G$  is a direct sum of cyclic and quasicyclic groups.

If we split  $G$ , according to (2), into a direct sum of an algebraically closed group  $A$  and of a reduced group  $U$ , then it is obvious that a basic subgroup of  $U$  is at the same time a basic subgroup of  $G$ . Hence, in constructing a basic subgroup of  $G$ , we can restrict ourselves to the reduced part  $U$  of  $G$ . We make, however, no use of this fact in the sequel.

Finally, we remark that  $G$  is a basic subgroup of  $G$  if and only if  $G$  is a direct sum of cyclic groups, and  $0$  is a basic subgroup of  $G$  if and only if  $G$  is a direct sum of quasicyclic groups. Also it is easy to show that  $G$  possesses a unique basic subgroup if and only if  $G$  is a direct sum of cyclic groups of bounded order.

#### § 4. The definitions of the basic subgroup by Kulikov and Kaloujnine

The present section serves merely the purpose of showing the equivalence of the definition of basic subgroup given in § 3 with that due to KULIKOV resp. KALOUJNINE. For this purpose we must introduce the notion of serving subgroup. A subgroup  $H$  of an abelian  $p$ -group  $G$  is called a *serving subgroup* in  $G$  if the height of any element  $a \in H$  relative to  $H$  coincides with that relative to  $G$ . This requirement can be expressed also as follows:  $p^n H = H \cap p^n G$  for  $n = 1, 2, 3, \dots$ . Obviously every direct summand of  $G$  is a serving subgroup in  $G$  and, more generally, the union of any ascending chain of direct summands is a serving subgroup in  $G$ . Thus, in particular, the basic subgroup  $B$  of  $G$  defined as the union of all subgroups  $S_n$  ( $n = 1, 2, 3, \dots$ ) of  $G$  (see (5), (6)) is always a serving subgroup of  $G$ . That a serving subgroup is not necessarily a direct summand, is shown by the important example of the basic subgroup.

Now, KULIKOV's definition of the basic subgroup [6], [9] is contained in the following

**THEOREM 3.** *A subgroup  $B$  of an abelian  $p$ -group  $G$  is a basic subgroup of  $G$  if and only if*

- 1)  $B$  is a direct sum of cyclic groups;
- 2)  $B$  is a serving subgroup in  $G$ ;
- 3)  $G/B$  is an algebraically closed group.

**PROOF.** By virtue of our above statements it is sufficient to show that a subgroup  $B$  having properties 1), 2), 3) is a basic subgroup of  $G$ . For this purpose we use the notations introduced in (1) and (6) where  $B_n$  denotes a direct sum  $\Sigma C(p^n)$ . In order to show that  $S_n$  is a direct summand of  $G$  we prove by induction on  $n$  that  $S_n$  is a direct summand of every containing



subgroup  $H$  of  $G$ :

$$(15) \quad H = S_n + K.$$

We put  $S_0 = 0$  and assume that (15) is true for some  $n \geq 0$ . Since a direct summand of a serving subgroup of  $G$  is itself a serving subgroup in  $G$ , we infer from condition 2) that  $S_n$  and

$$(16) \quad S_{n+1} = S_n + B_{n+1}$$

are serving subgroups in  $G$ . But then  $S_{n+1}/S_n$  is a serving subgroup in  $M/S_n$  for an arbitrary group  $M$  such that  $S_n \subseteq M \subseteq G$ , i. e.,

$$(S_{n+1}/S_n) \cap p^{n+1}(M/S_n) = p^{n+1}(S_{n+1}/S_n) = 0.$$

Consequently the  $p^{n+1}$ -regular subgroup  $S_{n+1}/S_n$  is a direct summand of  $M/S_n$  (see the Lemma in § 2):

$$M/S_n = (S_{n+1}/S_n) + (H/S_n)$$

with a suitable subgroup  $H \subseteq M$ . This means

$$M = \{S_{n+1}, H\}, \quad S_{n+1} \cap H = S_n,$$

i. e., by (16),

$$(17) \quad M = \{B_{n+1}, H\}, \quad B_{n+1} \cap H = 0.$$

Therefore we get from (17), (15) and (16):

$$M = B_{n+1} + H = S_{n+1} + K.$$

Thus we have shown that  $S_{n+1}$  is a direct summand of every containing subgroup  $M$  of  $G$ , and this completes the proof of the statement that  $S_n$  is a direct summand of  $G$  for every  $n$ .

Finally we must show that  $G$  has no direct summand of the form  $S_n + \{a\}$  with  $a \notin B$ . But this is clear for condition 3) implies

$$(18) \quad px = a + b$$

with some  $x \in G$  and  $b \in B$ , i. e.,  $b \in S_n$  for suitable  $n$ . Thus, by (18), a representation  $G = S_n + \{a\} + N$  is obviously impossible.

KALOUJNINE's definition of the basic subgroup differs from that of KULIKOV in replacing condition 3) by an equivalent condition of topological character provided that  $G$  contains no elements of infinite height. In this case, namely, the subgroups  $G, pG, p^2G, \dots$  have no element  $\neq 0$  in common and so this system may be taken as a complete system of neighborhoods of 0 in a well defined topology of  $G$  which is called *the natural topology* in  $G$ . Then it is easy to see that condition 3) is equivalent to the requirement that  $B$  is everywhere dense in  $G$ . In fact, condition 3) means that for an arbitrary (but fixed) element  $a \in G$  and natural number  $n$  there exists an element  $b \in B$  such that  $a - b \in p^n G$ . This is equivalent to the statement that every open set of the form  $a + p^n G$  intersects  $B$ , i. e.,  $B$  is everywhere dense in  $G$ . So we have obtained from Theorem 3 KALOUJNINE's characterization of the basic subgroup [3]:



THEOREM 4. A subgroup  $B$  of an abelian  $p$ -group  $G$  containing no elements of infinite height is a basic subgroup of  $G$  if and only if

- 1)  $B$  is a direct sum of cyclic groups;
- 2)  $B$  is a serving subgroup in  $G$ ;
- 3)  $B$  is everywhere dense in  $G$ .

### § 5. The basic subgroup as endomorphic image

Now we are going to prove the main result of this paper, namely, the

THEOREM 5. If  $B$  is a basic subgroup of an arbitrary abelian  $p$ -group  $G$ , then  $B$  is a homomorphic image of  $G$ .

The fundamental idea of the proof will be rather transparent if we first give the proof in the special case of the group  $G^*$  which served us in § 3 as an example. Suppose that the element  $a$  of  $G^*$  in (12) is of order  $p^r$ . Then obviously  $p^{n-r} | m_n$  ( $n \geq r$ ), i. e.,

$$(19) \quad m_n = m'_n p^{n-r} \quad (n = r, r+1, r+2, \dots).$$

Consequently,

$$(20) \quad a \rightarrow m_2 c_1 + m_4 c_2 + m_6 c_3 + \dots$$

is a well defined homomorphic mapping of  $G^*$  onto the basic subgroup  $B^* = \{c_1\} + \dots + \{c_n\} + \dots$  of  $G^*$  since the sum in (20) contains but a finite number of terms  $\neq 0$ ; more exactly, (19) implies

$$m_{2k} c_k = m'_{2k} p^{2k-r} c_k = 0 \quad \text{for } k \geq r,$$

taking into account that  $O(c_k) = p^k$ .

Passing now to the general case, let (1) be a basic subgroup of an arbitrary abelian  $p$ -group  $G$ , and let

$$(21) \quad m = \min_{n=0, 1, 2, \dots} \text{rank } (p^n B) = \text{rank } (p^m B)$$

(where we may suppose that  $m$  is fixed in the sequel).  $m$  is an invariant of  $G$ , it will be called the *critical number* of  $G$ .<sup>4</sup> We decompose

$$B = B_1 + B_2 + \dots + B_{m-1} + F_{m-1} = S_{m-1} + F_{m-1}$$

into the direct sum of its cyclic direct components and consider the direct components  $C(p^k)$  of  $F_{m-1}$ . (Clearly  $k \geq m$ .) We make correspond to each  $\{y\} = C(p^k)$  another direct component  $\{x\} = C(p^l)$  with  $l \geq 2k$  such that this correspondence between the whole set of direct components of  $F_{m-1}$  and some subset of these components be one-to-one. By the choice of  $m$  and  $m$  such a correspondence surely exists. Now we define a homomorphism  $x \rightarrow \varepsilon x$

<sup>4</sup> We emphasize that  $m = 0$  is equivalent to the statement that  $B$  is a bounded group. In this trivial case  $B = B_1 + \dots + B_{m-1} = S_{m-1}$  is a direct summand of  $G$ , so that trivially  $G \sim B$ . Therefore in proving Theorem 5 we may assume in the sequel  $m > 0$ .



( $x \in B$ ) of  $B$  onto itself induced by

$$(22) \quad \begin{cases} \varepsilon x = x & \text{if } x \text{ belongs to a cyclic direct component of } S_{m-1}; \\ \varepsilon x = y & \text{if } \{x\} = C(p^r) (\in F_{m-1}) \text{ corresponds to } \{y\} = C(p^r) (\in F_{m-1}); \\ \varepsilon x = 0 & \text{for all other direct components of } F_{m-1}. \end{cases}$$

Making use of this mapping  $\varepsilon$  we can define also a homomorphism of  $G$  onto  $B$  as follows. Let  $a$  be an arbitrary element of order  $p^r$  in  $G$ . Then, according to (5), we have a representation

$$(23) \quad a = b_1 + b_2 + \cdots + b_n + d_n \quad (b_i \in B_i, d_n \in D_n)$$

of "increasing length" of  $a$  successively for  $n = 1, 2, 3, \dots$ . Therefore we can associate with  $a$  the infinite "vector"

$$(23') \quad a \rightarrow \langle b_1, b_2, \dots, b_n, \dots \rangle \quad (b_i \in B_i)$$

and, by making use of (22), we define

$$(24) \quad \varepsilon a = \varepsilon b_1 + \varepsilon b_2 + \cdots + \varepsilon b_n + \dots$$

This mapping  $a \rightarrow \varepsilon a$  implies obviously a homomorphism of  $G$  onto  $B$  provided that the sum in (24) contains but a finite number of terms  $\neq 0$ . This is therefore the only thing we have yet to show. But this follows at once from (22), for  $O(a) = p^r$  implies  $O(b_n) \leq p^r$  and that the height of  $b_n$  is  $\geq n - r$  for  $n \geq r$ , consequently, by (22), we have

$$\varepsilon b_n = 0 \quad \text{if } n \geq 2r,$$

q. e. d.

## § 6. Applications

Let  $\mathfrak{p}$  be an arbitrary cardinal number. Following the terminology of L. FUCHS we shall say that an abelian  $p$ -group  $G$  has property  $A_{\mathfrak{p}}$  if it can be mapped homomorphically onto a direct sum of  $\mathfrak{p}$  groups, each isomorphic to the direct sum

$$(25) \quad C(p) + C(p^2) + C(p^3) + \cdots + C(p^r) + \dots$$

In connection with certain investigations, L. FUCHS raised the problem of characterizing all abelian  $p$ -groups  $G$  with property  $A_{\mathfrak{p}}$  [2]. Now we can show that this problem may be solved in terms of the basic subgroup of  $G$ , i. e., by a structural invariant of  $G$ . Namely, we have

**THEOREM 6.** *An abelian  $p$ -group  $G$  has property  $A_{\mathfrak{p}}$  if and only if a basic subgroup  $B$  of  $G$  possesses this property.<sup>5</sup>*

Since the basic subgroup  $B$  is a direct sum of cyclic groups, the problem of having property  $A_{\mathfrak{p}}$  may be decided trivially for  $B$ . We get also the following criterion:

<sup>5</sup> For an essentially simpler proof of this theorem see the following paper of L. FUCHS: On a property of basic subgroups, *Acta Math. Acad. Sci. Hung.*, 5 (1954), pp. 143–144.



THEOREM 7. Let  $m$  be the critical number<sup>6</sup> of  $G$ . Then  $G$  has property  $A_p$  if and only if  $m \cong p$ .

PROOF. If  $p \cong \aleph_0$ , then the validity of Theorems 6 and 7 follows immediately from  $G \sim B$  and, on the other hand, from the fact that (25) cannot be a homomorphic image of a group  $A + B$  where  $A$  is an algebraically closed group and  $B$  is a bounded group. Thus in the sequel we may assume  $p > \aleph_0$ . If the critical cardinal number  $m$  of  $G$  is greater than or equal to  $p$ , then the basic subgroup  $B$  of  $G$  has obviously property  $A_p$  and thus, by virtue of Theorem 5, the same is true also for  $G$ . Therefore it remains only to prove the following statement: *if  $\aleph_0 < p$  and  $m < p$ , then  $G$  cannot be mapped homomorphically onto a direct sum of  $p$  groups (25)*. Furthermore we may assume that  $G$  is a reduced group since  $G$  has property  $A_p$  if and only if the reduced part of  $G$  has property  $A_p$ . Now, let us suppose that  $G$  has property  $A_p$ . We choose  $n$  so that  $B$  contains no regular  $p^n$ -bounded subgroup of rank  $p$ . (This is always possible by the definition of the critical number  $m$ , and by our hypothesis  $p > m$ .) From this statement we deduce a contradiction to the hypothesis that  $G$  has property  $A_p$ , and this will complete the proof.

Since by our hypothesis  $G$  has property  $A_p$ ,  $G$  can be homomorphically mapped onto a direct sum

$$E = \sum_p C(p^n)$$

of  $p$  groups  $C(p^n)$ . Therefore  $G$  must have also a subgroup isomorphic to  $E$  which may be identified with  $E$ , i. e., the group  $E$  can be regarded as the endomorphic image of  $G$  under an endomorphism  $\eta$ :

$$(26) \quad \eta G = E = \sum_p C(p^n).$$

Now we observe that an endomorphism  $\eta$  of  $G$  is completely determined by its action on the basic subgroup  $B$ . As a matter of fact, in  $\eta$  and  $\eta'$  are endomorphisms of  $G$  such that  $\eta b = \eta' b$  holds for every element  $b \in B$ , then  $B$  is contained in the kernel of the endomorphism  $\eta' - \eta$ , and thus the image group  $(\eta' - \eta)G$  is a homomorphic image of the algebraically closed group  $G/B$ . This implies  $(\eta' - \eta)G = 0$ , since 0 is the only algebraically closed subgroup of the reduced group  $G$ . Therefore  $\eta' = \eta$ .

Now we associate with each element  $a \in G$  on basis of (23) and (23') the element

$$(27) \quad a \rightarrow \eta_1 b_1 + \eta_2 b_2 + \cdots + \eta_n b_n + \cdots.$$

This is a well-defined mapping of  $G$  into  $E$ , since the sum on the right hand side contains but a finite number of terms  $\neq 0$  which can be proved

<sup>6</sup> The critical number was defined in (21).



as the analogous statement in § 5 by considering the heights of the occurring elements. Moreover, the mapping (27) is an endomorphism of  $G$  which coincides with  $\eta$  on  $B$ . Consequently, as we have seen above, the mapping (27) is identical with  $\eta$ , and this means that the image group  $\eta G$  is a subgroup of  $E$ . But this contradicts (26) since by our hypothesis on  $n$  the group  $B$  contains no subgroup isomorphic to  $E$ . This contradiction completes the proof of Theorems 6 and 7.

As an immediate consequence of Theorem 7 we get

**THEOREM 8.** *Let  $G$  be an arbitrary abelian  $p$ -group with the critical number  $m$ . Then a necessary and sufficient condition that every abelian  $p$ -group of power  $\leq p$  be a homomorphic image of  $G$  is*

$$m \geq p.$$

**REMARK.** My hearty thanks are due to L. FUCHS whose kind remarks have made possible shorter proof of Theorem 5.

(Received 5 March 1954)

### Bibliography

- [1] L. FUCHS, On the structure of abelian  $p$ -groups, *Acta Math. Acad. Sci. Hung.*, **4** (1953), pp. 267—288.
- [2] L. FUCHS, On a special kind of duality in group theory. II, *Acta Math. Acad. Sci. Hung.*, **4** (1953), pp. 299—314.
- [3] L. KALOUJNINE, Sur les groupes abéliens primaires sans éléments de hauteur infinie, *Comptes Rendus Acad. Sci., Paris*, **225** (1947), pp. 713—715.
- [4] A. KERTÉSZ, On fully decomposable abelian torsion groups, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 225—232.
- [5] Л. Я. Куликов, К теории абелевых групп произвольной мощности, *Мат. Сборник*, **9** (51) (1941), pp. 165—181.
- [6] Л. Я. Куликов, К теории абелевых групп произвольной мощности, *Мат. Сборник*, **16** (58) (1945), pp. 129—162.
- [7] Л. Я. Куликов, Обобщенные примарные группы. I, *Труды Моск. Мат. Общ.*, **1** (1952), pp. 247—326.
- [8] Л. Я. Куликов, Обобщенные примарные группы. II, *Труды Моск. Мат. Общ.*, **2** (1953), pp. 85—167.
- [9] А. Г. Курош, *Теория групп* (Москва, 1953).
- [10] T. SZELE, On non-countable abelian  $p$ -groups, *Publicationes Math. (Debrecen)*, **2** (1952), pp. 300—301.
- [11] T. SZELE, On direct decomposition of abelian groups, *J. London Math. Soc.*, **28** (1953), pp. 247—250.



О БАЗИСНЫХ ПОДГРУППАХ АБЕЛЕВЫХ  $p$ -ГРУПП

Т. СЕЛЕ (Дебрецен)

(Резюме)

Автор рассматривает базисную подгруппу абелевой  $p$ -группы  $G$ , открытую Л. Я. Куликовым и Л. Калужниным. В работе доказывается, что базисная подгруппа группы  $G$  является всегда гомоморфным образом от  $G$ . В качестве применения автором решается проблема о гомоморфных образах абелевых  $p$ -групп, поставленная Л. Фуксом.







## ON A PROPERTY OF BASIC SUBGROUPS

By

L. FUCHS (Budapest)

(Presented by L. RÉDEI)

In a recent paper<sup>1</sup> T. SZELE has proved the following interesting results.

I. A basic subgroup  $B$  of an abelian  $p$ -group  $G$  is a homomorphic image of  $G$ .

II. If a direct sum of cyclic groups is a homomorphic image of an abelian  $p$ -group  $G$ , then the same direct sum is a homomorphic image of some (and hence each) basic subgroup  $B$  of  $G$ .

The starting point of the present simple note is the observation that the cited results imply at once

**THEOREM.** *If  $H$  is a homomorphic image of an abelian  $p$ -group  $G$ , then any basic subgroup  $C$  of  $H$  is a homomorphic image of any basic subgroup of  $G$ .*

As a matter of fact, by I one has  $H \sim C$  and hence  $G \sim H$  implies  $G \sim C$ . By the definition of basic subgroup,  $C$  is the direct sum of cyclic groups, so that II implies  $B \sim C$ , as we wished to show.

Now we intend to prove that the statement of our Theorem may be verified by a simple argument without appeal to II, by making use of I only, and then II will follow at once, thus providing a simpler proof of II.

As above, one may conclude to the existence of a homomorphism  $G \sim C$  which implies  $p^n G \sim p^n C$  for each non-negative integer  $n$ . Combining the latter homomorphism with the natural homomorphism  $p^n C \sim p^n C / p^{n+1} C$  we obtain a homomorphic mapping  $p^n G \sim p^n C / p^{n+1} C$  whose kernel clearly contains  $p^{n+1} G$ . Consequently, we have

$$(1) \quad p^n G / p^{n+1} G \sim p^n C / p^{n+1} C \quad (n = 0, 1, \dots).$$

Next take into account that for a basic subgroup  $B$  of  $G$  and for each non-negative integer  $n$  the isomorphism<sup>2</sup>

$$p^n G / p^{n+1} G \cong p^n B / p^{n+1} B$$

<sup>1</sup> On the basic subgroups of abelian  $p$ -groups, *Acta Math. Acad. Sci. Hung.* **5** (1954), pp. 129—141. For the definition of basic subgroups and their isomorphism we refer to this paper.

<sup>2</sup> This isomorphism is an easy consequence of the definition of basic subgroup.



holds, so that by (1) we are led to a homomorphism

$$(2) \quad p^n B/p^{n+1} B \cong p^n C/p^{n+1} C \quad (n = 0, 1, \dots).$$

The factorgroup  $p^n B/p^{n+1} B$  is a direct sum of cyclic groups of order  $p$  where the power of the set of the components is equal to that of those direct summands of  $B$  which are of order  $\cong p^{n+1}$ . In view of (2), this fact, together with the same fact in  $C$ , implies that in  $C$  the power of the set of the direct summands of order  $\cong p^{n+1}$  does not exceed the cardinal number defined analogously for  $B$  ( $n = 0, 1, \dots$ ). Since  $B$  and  $C$  are both direct sums of cyclic groups, it follows that there is a homomorphism  $B \sim C$ , as stated.

Finally, statement II follows immediately: if  $H$  is a direct sum of cyclic groups, then  $H$  is a basic subgroup of itself, and therefore, by our Theorem,  $G \sim H$  implies  $B \sim H$ .

It is worth while mentioning that the converse of Theorem is not true in general. A counterexample: let  $G$  be the discrete direct sum of the cyclic groups of order  $p, p^2, p^3, \dots$  respectively, and  $H$  the torsion subgroup of the complete direct sum of the same groups. Then  $B \cong C (\cong G)$ , but  $G \sim H$  is impossible by virtue of the powers of  $G$  and  $H$ .

(Received 5 May 1954)

## ОБ ОДНОМ СВОЙСТВЕ БАЗИСНОЙ ПОДГРУППЫ

Л. ФУКС (Будапешт)

(Резюме)

Автор доказывает следующую теорему:

Если  $H$  представляет собой гомоморфный образ примарной абелевой группы  $G$ , то базисная подгруппа  $C$  от  $H$  является гомоморфным образом базисной подгруппы  $B$  от  $G$ .



Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en un volume.

On est prié d'envoyer les manuscrits destinés à la rédaction et écrits à la machine à l'adresse suivante :

*Acta Mathematica, Budapest 62, Postafiók 440.*

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints (\$ 6.50) par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur des Livres et Journaux „Kultúra“ (Budapest, VI., Sztálin-út 21. Compte-courant No. 45-790-057-50-032) ou à l'étranger chez tous les représentants ou dépositaires.

---

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up one volume.

Manuscripts should be typed and addressed to :

*Acta Mathematica, Budapest 62, Postafiók 440.*

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica is 110 forints (\$ 6.50) a volume. Orders may be placed with „Kultúra“ Foreign Trade Company for Books and Newspapers (Budapest, VI., Sztálin-út 21. Account No. 45-790-057-50-032) or with representatives abroad.

---

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französischer, englischer und deutscher Sprache.

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, mit Maschine geschrieben, an folgende Adresse zu senden :

*Acta Mathematica, Budapest 62, Postafiók 440.*

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band 1.0 Forints (\$ 6.50). Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen „Kultúra“ (Budapest, VI., Sztálin-út 21. Bankkonto Nr. 45-790-057-50-032) oder bei seinen Auslandsvertretungen und Kommissionären.



## INDEX

<i>Hajós, G., and Rényi, A.,</i> Elementary proofs of some basic facts concerning order statistics. Г. Хайош и А. Реньи, Элементарное доказательство некоторых основных фактов теорий вариационных рядов . . . . .	1
<i>Varga, O.,</i> Bedingungen für die Metrisierbarkeit von affinzusammenhängenden Linien-elementmännigfaltigkeiten. О. Варга, Об условиях метризуемости аффинно-связных многообразий линейных элементов . . . . .	7
<i>Rédei, L.,</i> Über die Kantenbasen für endliche vollständige gerichtete Graphen. Л. Рэдэи, О фазисах ребер конечных ориентированных полных графов . . . . .	17
<i>Rédei, L.,</i> Über das Kreisteilungspolynom. Л. Рэдэи, О циклотомическом полиноме . . . . .	27
<i>Szász, P.,</i> Über die Trigonometrie des Poincaréschen Kreismodells der hyperbolischen ebenen Geometrie. П. Сас, О тригонометрии круговой модели Пуанкаре для гиперболической плоской геометрии . . . . .	29
<i>Szűsz, P.,</i> Über die Verteilung der Vielfachen einer komplexen Zahl nach dem Modul des Einheitsquadrats. П. Сюс, О распределении кратных комплексного числа по модулю квадрата единицы . . . . .	35
<i>Fejes Tóth, L.,</i> Über die dichteste Horozyklenlagerung. Л. Фееш Тот, О самом плотном расположении гороциклов . . . . .	41
<i>Erdős, J.,</i> The theory of groups with finite classes of conjugate elements. И. Эрдэш, Теория групп с конечными классами сопряженных элементов . . . . .	45
<i>Moór, A.,</i> Die oskulierenden Riemannschen Räume regulärer Cartanscher Räume. А. Моор, Соприкасающиеся римановы пространства регулярных картановых пространств . . . . .	59
<i>Soós, Gy.,</i> Über Gruppen von Affinitäten und Bewegungen in Finslerschen Räumen. Д. Шоош, Об аффинитетах и движениях в финслеровых пространствах . . . . .	73
<i>Fenyő, I.,</i> Über die Lösung der im Banachschen Raume definierten nichtlinearen Gleichungen. И. Феньё, О нелинейных уравнениях, определенных в пространстве Банаха . . . . .	85
<i>Fuchs, L.,</i> On the fundamental theorem of commutative ideal theory. Л. Фукс, Об основной теореме теории коммутативных идеалов . . . . .	95
<i>Szele, T.,</i> Simple proof of the Wedderburn—Artin structure theorem. Т. Селе, Простое доказательство структурной теоремы Веддерберна—Артина . . . . .	101
<i>Freud, G.,</i> Über die Konvergenz des Hermite—Fejérschen Interpolationsverfahrens. Г. Фрайд, О сходимости интерполяционного процесса Эрмита—Фейера . . . . .	109
<i>Szele, T.,</i> On the basic subgroups of abelian $p$ -groups. Т. Селе, О базисных подгруппах абелевых $p$ -групп . . . . .	129
<i>Fuchs, L.,</i> On a property of basic subgroups. Л. Фукс, Об одном свойстве базисной подгруппы . . . . .	143

Technikai szerkesztő: Fuchs László

A kiadásért felelős: Mestyán János.

Műszaki felelős: Tóth Ferenc

A kézirat beérkezett: 1954. IV. 16. — Példányszám: 800 — Terjedelem: 12<sup>1</sup>/<sub>2</sub> (A/5) iv, 43 ábra.

Cscngrádmegyei Nyomdaipari Vállalat, Szeged, 54-2667

Felelős vezető: Vincze György



# ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM  
HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,  
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,  
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS V

FASCICULI 3—4



MAGYAR TUDOMÁNYOS AKADÉMIA  
BUDAPEST, 1954

ACTA MATH. HUNG.



# ACTA MATHEMATICA HUNGARICA

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK  
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG ÉS KIADÓHIVATAL: BUDAPEST, V., ALKOTMÁNY-U. 21

Az Acta Mathematica orosz, francia, angol és német nyelven közöl értekezéseket a matematika köréből.

Az Acta Mathematica változó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet.

A közlésre szánt kéziratok géppel írva, a következő címre küldendőek:

*Acta Mathematica, Budapest 62, Postafiók 440.*

Ugyanerre a címre küldendő minden szerkesztőségi és kiadóhivatali levelezés.

Az Acta Mathematica előfizetési ára kötetenként belföldre 80 forint, külföldre 110 forint. Megrendelhető a belföld számára az „Akadémiai Kiadó”-nál (Budapest, V., Alkotmány-utca 21. Bankszámla 04-878-111-48), a külföld számára pedig a „Kultúra” Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin-út 21. Bankszámla 45-790-057-50-032 sz.) vagy külföldi képviselőinél és bizományosainál.

---

„Acta Mathematica” публикует трактаты из области математических наук на русском, французском, английском и немецком языках.

„Acta Mathematica” выходит отдельными выпусками разного объема. Несколько выпусков составляют один том.

Предназначенные для публикации рукописи (в напечатанном на машинке виде) следует направлять по адресу:

*Acta Mathematica* (Венгрия, Будапешт 62, п/я 440).

По этому же адресу направлять всякую корреспонденцию для редакции и администрации.

Подписная цена „Acta Mathematica” — 110 форинтов за том. Заказы в стране принимает *Akadémiai Kiadó* (Alkotmány-utca 21. Текущий счет № 04-878-111-48), а для заграницы, предприятие по внешней торговле книг и газет „Kultúra” (Budapest, VI., Sztálin-út 21. Текущий счет № 45-790-057-50-032), или его заграничные представительства и уполномоченные.



Az Acta Mathematica V. kötetét a jelen füzettel még nem zártuk le. E kötethez egy függelék fogunk megküldeni előfizetőinknek. E függelék tartalmazza a Magyar Tudományos Akadémia által 1952 decemberében, Bolyai János születésének 150. évfordulója alkalmából tartott ünnepi ülészak előadásait.

Пятый том Acta Mathematica будет дополнен приложением, содержащим доклады, прочитанные на торжественной сессии Академии Наук Венгрии в декабре 1952 года по случаю 150-летней годовщины со дня рождения Яноша Бояи.

Le tome V des Acta Mathematica sera complété par un recueil de conférences faites à la session de l'Académie des Sciences de Hongrie en décembre 1952 à l'occasion du 150<sup>ème</sup> anniversaire de la naissance de János Bolyai.

Volume V of the Acta Mathematica will be completed by a Supplement containing the papers read at the festival session of the Hungarian Academy of Sciences in December 1952 on the occasion of the 150th birthday of János Bolyai.

Der Band V der Acta Mathematica wird noch mit einem Supplement ergänzt werden, das die im Dezember 1952 auf der zum 150. Geburtstag von János Bolyai veranstalteten Festsitzung der Ungarischen Akademie der Wissenschaften gehaltenen Vorträge umfaßt.







## ON LINDELÖF'S CONJECTURE

By

P. TURÁN (Budapest), member of the Academy

1. In what follows, we denote the complex variable by  $s = \sigma + it$ , further if  $\frac{1}{2} \leq \alpha \leq 1$ ,  $T \geq 2$ , we denote by  $N(\alpha, T)$  the number of the zeros of the Riemann zeta-function in the parallelogram

$$(1.1) \quad \alpha \leq \sigma \leq 1, \quad 0 < t \leq T.$$

$c_1, c_2, \dots$  denote constants whose exact values are not essential and which might depend upon an  $\varepsilon$  or  $\eta$ -parameter; if so, the dependence will always be explicitly stated. Then the well-known theorem of INGHAM<sup>1</sup> states essentially the following:

if

$$(1.2) \quad \left| \zeta \left( \frac{1}{2} + it \right) \right| < c_1 (1 + |t|)^a \log^b (2 + |t|)$$

with numerical  $a$  and  $b \leq \frac{3}{2}$ , then we have the inequality

$$(1.3) \quad N(\alpha, T) < c_2 T^{2(1+2\alpha)(1-\alpha)} \log^5 T.$$

As well known,<sup>2</sup> the simplest admissible values for  $a$  and  $b$  are

$$a = \frac{1}{6}, \quad b = \frac{3}{2};$$

this gives for  $\frac{1}{2} \leq \alpha \leq 1$  the estimation

$$(1.4) \quad N(\alpha, T) \leq c_2 T^{\frac{8}{3}(1-\alpha)} \log^5 T.$$

It is well known the effect of this inequality; following an idea of HOHEISEL, INGHAM deduced from it the estimation

$$(1.5) \quad p_{n+1} - p_n < c_3(\varepsilon) p_n^{\frac{5}{8} + \varepsilon},$$

if  $n > n_0(\varepsilon)$  and  $p_n$  denotes the  $n^{\text{th}}$  prime. He showed further that if the

<sup>1</sup> A. E. INGHAM, On the difference between consecutive primes, *Quart. J. of Math. Oxf. Ser.*, **8** (1937), pp. 255–266.

<sup>2</sup> E. C. TITCHMARSH, *The theory of the Riemann zeta-function*, 2nd edition (Oxf., 1951), p. 81.



constant  $\frac{8}{3}$  in (1.4) could have been replaced by  $2 + \varepsilon$  with arbitrary small positive  $\varepsilon$ , then (1.5) could have been replaced by

$$(1.6) \quad p_{n+1} - p_n < c_4(\varepsilon) p_n^{\frac{1}{2} + c_5(\varepsilon)}$$

where

$$(1.7) \quad \lim_{\varepsilon \rightarrow 0} c_5(\varepsilon) = 0,$$

i. e. the fourth main-problem of the analytical number-theory would be essentially solved. Now the estimation (1.3) has the interesting consequence that the truth of the so-called LINDELÖF hypothesis implies already the truth of the inequality (1.6). This hypothesis, unproved sofar, asserts that the value of  $a$  in (1.2) can be chosen to be an arbitrarily small positive number. This hypothesis is definitely weaker than RIEMANN'S conjecture; for a long time it was thought that it is so weak that its proof has almost no consequence on the distribution of zeros. The significance of INGHAM'S theorem consists in showing that perhaps LINDELÖF'S conjecture is deeper than thought previously.

If the constant  $\frac{8}{3}$  in (1.4) can be replaced by  $d$  ( $\geq 2$ ) uniformly for  $\frac{1}{2} \leq \alpha \leq 1$ , then in (1.5) the exponent  $\frac{5}{8} + \varepsilon$  could have been replaced by  $(1 - \frac{1}{d} + \varepsilon)$ .

2. The present author introduced recently a method by which among others some results on the above-mentioned questions could have been achieved too.<sup>3</sup> So he proved<sup>4</sup> without any hypothesis the inequality

$$(2.1) \quad N(\alpha, T) < c_6 T^{2(1-\alpha) + 600(1-\alpha)^{\frac{101}{100}}} \log^6 T$$

valid uniformly for  $T \geq 2$  and<sup>5</sup> with a positive numerical  $c$  for

$$(2.2) \quad 1 - c \leq \alpha \leq 1$$

which constitutes the best inequality sofar for  $N(\alpha, T)$  "near to  $\alpha = 1$ " where in a sense it is the deepest. It is to a certain extent a similar situation in the investigation of the order of  $|\zeta(s)|$  for which the deepest results could have been achieved in the neighbourhood of the line  $\sigma = 1$ . Now concerning

<sup>3</sup> P. TURÁN, *Eine neue Methode in der Analysis und deren Anwendungen*, Akadémiai Kiadó (Budapest, 1953).

<sup>4</sup> L. c. <sup>3</sup> II. Teil. Anwendungen, § 14, p. 158.

<sup>5</sup> By performing the analysis more carefully the values 600 and  $\frac{101}{100}$  could have been replaced by much smaller resp. much greater values and also for  $c$  in (2.2) a numerical value could have been obtained. According to an unpublished improvement I replaced (2.1) by the estimation

$$N(\alpha, T) < c_6 T^{2(1-\alpha) + (1-\alpha)^{1,14}}.$$



this two principal questions arise. First, can the method in <sup>3</sup> work at all for  $\frac{1}{2} \leq \alpha \leq +1$ , i. e. also "far" from  $\sigma = 1$ ? At the second, can perhaps from the truth of LINDELÖF'S hypothesis much stronger estimation be derived for  $N(\alpha, T)$ ? For the first we shall see, the answer is yes.<sup>6</sup> As to the second it will turn out that everything depends on a value-distribution lemma in the half-plane  $\sigma > 1$  which will be treated here only by the simplest means. The possibilities of an improvement are discussed in **7**; this leads the author to the suspicion that LINDELÖF'S conjecture is deeper than thought before and perhaps it implies also the estimation

$$N(\alpha, T) < c_7(\varepsilon, \delta) T^\varepsilon$$

for  $\alpha \geq \frac{1}{2} + \delta$ ,  $T \geq 2$  with arbitrary small positive  $\varepsilon$  and  $\delta$ .

**3.** For the proof of the inequality (1.6), under the supposition of the validity of LINDELÖF'S conjecture, the following theorem is sufficient. If for  $\sigma \geq \frac{1}{2}$  and  $t \geq 1$  we have with an arbitrarily small  $\varepsilon > 0$

$$(3.1) \quad |\zeta(\sigma + it)| \leq c_8(\varepsilon) t^\varepsilon,$$

then, if  $\alpha \geq \frac{1}{2} + c_9(\varepsilon)$ ,  $T \geq 2$ , we have

$$(3.2) \quad N(\alpha, T) < c_{10}(\varepsilon) T^{2(1+c_{11}(\varepsilon))(1-\alpha)} \log^5 T$$

where  $c_9(\varepsilon)$  and  $c_{11}(\varepsilon)$  tend to 0 if  $\varepsilon \rightarrow 0$ .

We consider namely  $N(\alpha, T)$  for  $\frac{1}{2} \leq \alpha \leq \frac{1}{2} + c_9(\varepsilon)$ . Then, if  $c_9(\varepsilon) < \frac{1}{4}$ , we have

$$\begin{aligned} N(\alpha, T) &< c_{12} T \log T = c_{12} T^{1-\alpha(1-\alpha)} \log T \leq c_{12} T^{1-2c_9(\varepsilon)(1-\alpha)} \log T < \\ &< c_{12} T^{2(1+3c_9(\varepsilon))(1-\alpha)} \log T, \end{aligned}$$

i. e. choosing

$$c_{13}(\varepsilon) = \max(c_{11}(\varepsilon), 3c_9(\varepsilon))$$

we have

$$\lim_{\varepsilon \rightarrow 0} c_{13}(\varepsilon) = 0$$

and for  $\sigma \geq \frac{1}{2}$ ,  $T \geq 2$  uniformly

$$N(\alpha, T) < c_{14}(\varepsilon) T^{2(1+c_{13}(\varepsilon))(1-\alpha)} \log^5 T$$

with

$$c_{14}(\varepsilon) = \max(c_{12}, c_{10}(\varepsilon))$$

and INGHAM'S proof shows indeed that the inequality (1.6) follows with a suitable  $c_5(\varepsilon)$  tending to 0 with  $\varepsilon$ .

<sup>6</sup> Announced without proof in <sup>3</sup> p. 161.



4. In this paper we are going to show the following  
THEOREM. Suppose there is a  $\mathcal{F}$  with

$$(4.1) \quad \frac{1}{2} < \mathcal{F} < 1$$

such that for a sufficiently small positive  $\eta$  in the domain

$$(4.2) \quad \sigma \geq \mathcal{F}, \quad t \geq 1$$

the inequality

$$(4.3) \quad |\zeta(\sigma + it)| \leq c_{15}(\eta)t^{\eta^{1000}}$$

holds. Then for

$$(4.4) \quad (1 \geq) \sigma_1 \geq \mathcal{F} + 4\eta, \quad T > c_{16}(\eta)$$

we have

$$(4.5) \quad N(\sigma_1, T) < c_{17}(\eta) T^{2(1+3\eta)(1-\sigma_1)} \log^5 T.$$

It would be easy to modify after the pattern of 3 the theorem so that it remain valid also for  $\mathcal{F} = \frac{1}{2}$ ; i. e. include the form (3.1)—(3.2) of INGHAM's theorem. We can leave it to the interested reader. At a superficial investigation it seems to me, that supposing the inequality (4.3) only in a domain (4.2) with  $\mathcal{F} > \frac{1}{2}$ , INGHAM's method gives only an estimation for  $N(\sigma_1, T)$  worse than (4.5). The proof given here is essentially different from that of INGHAM; it operates with values assumed "far" in the half-plane  $\sigma > 1$  of a function related to  $\frac{\zeta'}{\zeta}(s)$  and uses LINDELÖF'S conjecture in the form that

$$N(\sigma_2, T+1) - N(\sigma_2, T) = o(\log T)$$

if only  $\sigma_2 \geq \mathcal{F} + \delta$ ,  $\delta > 0$  fixed and  $T \rightarrow \infty$ .

5. Before turning to the proof of the theorem, we need simple lemmas. We consider for  $\sigma > 1$ ,  $\xi \geq e^{10}$ , integer  $k$  satisfying the inequality

$$(5.1) \quad 10 \log \xi \geq k \geq 10,$$

the function  $g_k(s)$  defined by

$$(5.2) \quad g_k(s) = \sum_{n \leq \xi} \frac{A(n)}{n^s} \log^k \frac{n}{\xi}.$$

As easy to justify,  $g_k(s)$  is connected with  $\frac{\zeta'}{\zeta}(s)$  by the formula

$$(5.3) \quad g_k(s) = \frac{(-1)^k k!}{2\pi i} \int_{(\gamma)} \frac{\xi^w}{w^{k+1}} \frac{\zeta'}{\zeta}(s+w) dw$$

where

$$(5.4) \quad 1 - \sigma < \gamma < 0.$$



We define for  $A \geq 2$ ,  $B \geq 10$ ,  $C > 1$

$$(5.5) \quad S = S(A, B, C) = \sum_{n \geq A} n^{-C} \log^B \frac{n}{A}.$$

Then we have

LEMMA I.

$$S(A, B, C) < 3 \cdot B! \frac{C \cdot A^{1-C}}{(C-1)^{B+1}}.$$

The proof uses the integral-criterion and can be omitted; in <sup>3</sup> p. 191 we had a similar lemma but with the further restriction  $C \leq \frac{3}{2}$ .

We define further for

$$(5.6) \quad T \geq 2, \quad \sigma_0 > 1$$

the quantity  $J_k(T)$  by

$$(5.7) \quad J_k(T) = \int_T^{2T} |g_k(\sigma_0 + it)|^2 dt.$$

Then we have the

LEMMA II.

$$J_k(T) < c_{18} \left( 1 + \frac{1}{(\sigma_0 - 1)^4} \right) k!^2 \sigma_0^3 \log^4 \xi \left\{ \frac{T \xi^{1-2\sigma_0}}{(\sigma_0 - \frac{1}{2})^{2k+2}} + \frac{\xi^{2(1-\sigma_0)}}{(\sigma_0 - 1)^{2k+2}} \cdot 4^{\sigma_0} \right\}.$$

A similar lemma was proved in <sup>3</sup> p. 192 with the supplementary restriction  $\sigma_0 \leq \frac{3}{2}$ , but even therefore this is for our aims useless. However, the proof of this lemma II needs only slight changes, so a sketch of the proof will suffice. Replacing  $g_k(\sigma_0 + it)$  by its Dirichlet series, the usual argument leads to the inequality

$$(5.8) \quad J_k(T) \leq c_{19} \left\{ T \sum_{n \leq \xi} \frac{\log^2 n \log^{2k} \frac{n}{\xi}}{n^{2\sigma_0}} + \left( \sum_{n \leq \xi} \frac{\log n \log^k \frac{n}{\xi}}{n^{\sigma_0}} \right)^2 + \sum_{\xi \leq n < m \leq 2n} \frac{\log m \log n \log^k \frac{m}{\xi} \log^k \frac{n}{\xi}}{(mn)^{\sigma_0} \log \frac{m}{n}} \right\} \equiv c_{19} (J_1 + J_2 + J_3).$$

For  $J_1$  we have in terms of  $S$  in (5.5)

$$J_1 < 2T \{ \log^2 \xi \cdot S(\xi, 2k, 2\sigma_0) + S(\xi, 2k+2, 2\sigma_0) \},$$

i. e. using lemma I and also (5.1)

$$(5.9) \quad J_1 < c_{20} T \frac{\sigma_0 \xi^{1-2\sigma_0} \log^2 \xi}{(\sigma_0 - \frac{1}{2})^{2k+2}} k!^2.$$



Similarly

$$J_2 < \{ \log \xi \cdot S(\xi, k, \sigma_0) + S(\xi, k+1, \sigma_0) \}^2,$$

i. e., using again lemma I and (5. 1),

$$(5. 10) \quad J_2 < c_{21} \frac{\sigma_0^2 \xi^{2-2\sigma_0} \log^4 \xi \cdot k!^2}{(\sigma_0-1)^{2k+2}} \left( 1 + \frac{1}{(\sigma_0-1)^2} \right).$$

Finally,

$$J_3 < c_{22} \max_{j=0, 1, 2, 3} \log^{3-j} \xi \cdot S\left(\frac{\xi}{2}, 2k+j, 2\sigma_0-1\right),$$

i. e., using again lemma I and (5. 1),

$$(5. 11) \quad J_3 < c_{23} k!^2 \frac{\sigma_0^3 4^{\sigma_0} \xi^{2(1-\sigma_0)} \log^3 \xi}{(\sigma_0-1)^{2k+2}} \left( 1 + \frac{1}{(\sigma_0-1)^4} \right).$$

By (5. 8), (5. 9), (5. 10) and (5. 11), lemma II is proved.

**6.** What is essential in the sequel is a simple consequence of lemma II. For any real function  $h(t)$  defined in  $T \leq t \leq 2T$  and for any  $0 < \beta < 1$  we have

$$(6. 1) \quad |h(t)| \leq T^{-\frac{\beta}{2}} \left( \int_T^{2T} |h(x)|^2 dx \right)^{1/2},$$

except for a set  $L$  of measure  $m(L)$  satisfying the inequality

$$(6. 2) \quad m(L) \leq T^\beta.$$

For

$$\int_T^{2T} |h(x)|^2 dx \geq \int_{(L)} |h(x)|^2 dx \geq m(L) T^{-\beta} \int_T^{2T} |h(x)|^2 dx$$

which already proves (6. 1)—(6. 2). Applying this remark to  $h(t) = g_k(\sigma_0 + it)$ , we obtain that for  $T \leq t \leq 2T$  we have the inequality

$$(6. 3) \quad |g_k(\sigma_0 + it)| \leq c_{24} \left( 1 + \frac{1}{(\sigma_0-1)^2} \right) k! \sigma_0^2 \log^2 \xi \cdot T^{-\frac{\beta}{2}} \cdot \left\{ \frac{\sqrt{T} \xi^{\frac{1}{2}-\sigma_0}}{\left(\sigma_0 - \frac{1}{2}\right)^{k+1}} + 2^{\sigma_0} \frac{\xi^{1-\sigma_0}}{(\sigma_0-1)^{k+1}} \right\},$$

except for a set  $H_k$  of measure  $m(H_k)$  for which the inequality

$$(6. 4) \quad m(H_k) \leq T^\beta$$

holds. Now we restrict the value of the integer  $k$  by imposing

$$(6. 5) \quad \frac{1}{2} \log T \leq k+1 \leq 2 \log T;$$

then, calling  $H$  the union of the sets  $H_k$  for all values  $k$  of (6. 5), we obtained that on the segment

$$(6. 6) \quad \sigma = \sigma_0, \quad T \leq t \leq 2T$$



with the possible exception of a set  $H$  of measure  $m(H)$  satisfying

$$(6.7) \quad m(H) < 2T^\beta \log T$$

we have the inequality (6.3) for *all* integer values  $k$  satisfying (6.5).

For the sake of orientation we remark that for all  $T \leq t \leq 2T$  the inequality

$$\begin{aligned} |g_k(\sigma_0 + it)| &\leq \sum_{n \geq \xi} \frac{\log n \log^k \frac{n}{s\xi}}{n^{\sigma_0}} = \log \xi \cdot S(\xi, k, \sigma_0) + S(\xi, k+1, \sigma_0) < \\ &< c_{25} k! \log \xi \frac{\sigma_0 \xi^{1-\sigma_0}}{(\sigma_0 - 1)^{k+1}} \end{aligned}$$

trivially holds. Hence what makes (6.3) non-trivial, is the factor  $T^{-\frac{\beta}{2}}$ .

7. We call for the sake of convenience those values  $s$  of the segment (6.6), for which (6.3) is fulfilled for all values  $k$  of (6.5), "good" values. Let us divide the segment (6.6), starting from the point  $\sigma_0 + iT$ , into sub-intervals  $l_j$  open from the right, of length  $\frac{1}{\log^3 T}$  (the last one possibly  $< \frac{1}{\log^3 T}$ ). We call such a subinterval  $l_j$  "good" resp. "bad" if it contains at least one "good" value  $s$  resp. none. (6.3)—(6.7) give the result that at most

$$(7.1) \quad 2T^\beta \log^4 T$$

of them can be "bad" intervals.

This is a result on the distribution of values of  $g_k(s)$ . All improvements of it would imply a better estimation for  $N(\sigma_1, T)$ . One way of improvement can be the increasing of the exponent  $\frac{\beta}{2}$  in (6.3) with keeping (7.1) or what amounts to the same, the decreasing of the number of bad intervals in (7.1) with keeping the exponent  $\frac{\beta}{2}$  in (6.3). This, though difficult, seems not to be impossible.<sup>7</sup> But what we actually need, is a result of type (7.1), i. e. the

<sup>7</sup> To be more exact if in (6.3) the exponent  $\frac{\beta}{2}$  could be replaced by  $K\beta$ ,  $K > \frac{1}{2}$ , uniformly for  $0 \leq \beta \leq 1$ , then for any fixed  $\sigma > 0$  we could deduce uniformly for

$$\frac{1}{2} + \delta \leq \alpha \leq 1,$$

the estimation

$$N(\alpha, T) < C_{26}(\delta) T^{\frac{1}{K}(1+C_{27}(\delta))(1-\alpha)} \log^6 T$$

where  $C_{26}(\delta) > 0$  and tends to 0 with  $\delta$ , of course supposing the truth of LINDELÖF'S conjecture. This is particularly strong, when  $\alpha$  is near to  $\frac{1}{2}$ .



fact that all  $|g_k(\sigma_0 + it)|$  with (6.5) assume "not too big" values in "very many" preassigned little intervals and this may be the case even if  $m(H)$  cannot be diminished. To improve the value distribution theorem (7.1), i. e. to diminish the number of the bad intervals independently of the estimation of the measure  $m(H)$ , we are led to questions, in certain sense dual to those

I settled in.<sup>3</sup> In<sup>3</sup> I treated results according to which an expression<sup>8</sup>  $\left| \sum_{\nu=1}^n a_\nu z_\nu^x \right|$  assumes "often" values which are "not too small with respect to a norm". What we need are results according to which such an expression assumes "often" values which are "not too large with respect to a norm" (in particular with respect to the N. Wiener-norm). Such results could not be expected of course without any restrictions upon the numbers  $z_j$ ; a suitable restriction would be the lower limitation of  $\max_{1 \leq \mu, \nu \leq n} \left| \arccos \frac{z_\mu}{z_\nu} \right|$ .

8. After having established the value distribution theorem (7.1) we can turn to the proof of our theorem.  $\mathcal{G}$  being fixed with

$$(8.1) \quad \frac{1}{2} < \mathcal{G} < 1,$$

we choose  $\eta$  so small that

$$(8.2) \quad \mathcal{G} + 4\eta < 1 - \left( \frac{\eta}{100} \right)^{100},$$

$$(8.3) \quad \eta < \frac{1}{100},$$

$$(8.4) \quad \left( \frac{\eta}{100} \right)^{100} < c$$

where  $c$  denotes the constant in (2.2), further

$$(8.5) \quad \eta^{889} \log \frac{144C}{\eta^{990}} < \frac{1}{100^{101}}$$

and

$$(8.6) \quad \mathcal{G} > \frac{1}{2} + 10\eta \quad \left( > \frac{1}{2} + \eta^3 \right).$$

<sup>8</sup> Here  $a_\nu$  and  $z_\nu$  are fixed,  $x$  positive variable;  $z_\nu^x$  means  $e^{x \log z_\nu}$  with any fixed value of the logarithm. A norm means any positive expression of the quantities  $|z_\nu|^x$  resp.  $|a_\nu| |z_\nu|^x$ . So far we know only problems where the necessity of the H. Bohr-norm

$\sum_{\nu=1}^n |a_\nu| |z_\nu|^x$ , of the norms  $(\min_j |z_j|)^x$  or  $(\max_j |z_j|)^x$  and of the N. Wiener-norm

$$\left( \sum_{\nu=1}^n |a_\nu|^2 |z_\nu|^{2x} \right)^{1/2}$$

emerged.



Let  $\eta$  be a value satisfying all these restrictions and for which (4.2)—(4.3) holds; then we consider first the values  $\sigma_1$  with

$$(8.7) \quad 1 \cong \sigma_1 > 1 - \left(\frac{\eta}{100}\right)^{100}.$$

Owing to (8.4) we have in this case

$$\sigma_1 > 1 - c,$$

i. e. according to (2.1)

$$N(\sigma_1, T) < c_6 T^{2(1-\sigma_1)+600(1-\sigma_1)^{\frac{101}{100}}} \log^6 T.$$

But then using (8.7) we have

$$2(1-\sigma_1) + 600(1-\sigma_1)^{\frac{101}{100}} < 2(1+3\eta)(1-\sigma_1),$$

i. e., our theorem is proved for the values  $\sigma_1$  satisfying (8.7). Hence we may suppose

$$(8.8) \quad \mathcal{G} + 4\eta \cong \sigma_1 \cong 1 - \left(\frac{\eta}{100}\right)^{100}.$$

We determine  $\sigma_0$  of (6.3) by

$$(8.9) \quad \sigma_0 = 1 + \frac{1}{\eta}$$

and  $\beta$  by

$$(8.10) \quad \beta = \frac{2}{1-2\eta}(1-\sigma_1).$$

Then owing to (8.8) we have

$$(8.11) \quad \frac{2}{1-2\eta} \left(\frac{\eta}{100}\right)^{100} \cong \beta \cong \frac{2}{1-2\eta}(1-\mathcal{G}-4\eta),$$

which implies, using (8.6) and (8.3), the inequality

$$(8.12) \quad 0 < \beta < 1.$$

$T$  should be so large that

$$(8.13) \quad \frac{\log \log T}{\log T} < \frac{1}{2(1-2\eta)} \cdot \left(\frac{\eta}{100}\right)^{101},$$

$$(8.14) \quad \frac{1}{\log^3 T} < \frac{\eta}{1-2\eta} \cdot \frac{1}{2} \left(\frac{\eta}{100}\right)^{100} \quad (< \eta^3),$$

and

$$(8.15) \quad T > \max \left( e^{22}, \left(2 + \frac{1}{\eta}\right)^2 \right);$$

later on we shall have some more restriction upon  $T$ , all of the type  $T > T_0(\eta)$ .

On the integer  $k$  we impose at this moment only the restriction

$$(8.16) \quad \frac{1-(1-\eta)\beta}{1-(1-2\eta)\beta} \log T \cong k+1 \cong \frac{1-\eta}{1-2\eta+\eta\left(1-\frac{3}{2}\eta\right)} \log T.$$



Is that restriction not contradictory? Owing to (8.9) and (8.12) we have

$$\sigma_0 > 1 + \frac{\left(1 - \frac{3}{2}\eta\right)(1 - (1-\eta)\beta)}{\eta},$$

$$\frac{\left(1 - \frac{3}{2}\eta\right)(1 - (1-\eta)\beta)}{\sigma_0 - 1} < \eta,$$

$$\{1 - (1-\eta)\beta\} \left\{ (1-2\eta) + \frac{1 - \frac{3}{2}\eta}{\sigma_0 - 1} \right\} = (1-2\eta) - (1-\eta)(1-2\eta)\beta +$$

$$+ \frac{(1 - (1-\eta)\beta)\left(1 - \frac{3}{2}\eta\right)}{\sigma_0 - 1} < (1-\eta)(1 - (1-2\eta)\beta),$$

i. e.

$$\frac{1 - (1-\eta)\beta}{1 - (1-2\eta)\beta} < \frac{1-\eta}{1-2\eta + \frac{1 - \frac{3}{2}\eta}{\sigma_0 - 1}} = \frac{1-\eta}{1-2\eta + \eta\left(1 - \frac{3}{2}\eta\right)},$$

and hence (8.16) is compatible. Hence for all values  $\beta$  satisfying (8.11) we have

$$\frac{1-\eta}{1-2\eta + \eta\left(1 - \frac{3}{2}\eta\right)} - \frac{1 - (1-\eta)\beta}{1 - (1-2\eta)\beta} > c_{28}(\eta),$$

i. e. for  $T > c_{29}(\eta)$  the interval (8.16) contains integer values, indeed. From (8.3) we easily obtain

$$\frac{1-\eta}{1-2\eta + \eta\left(1 - \frac{3}{2}\eta\right)} < 2,$$

further from (8.12)

$$\frac{1 - (1-\eta)\beta}{1 - (1-2\eta)\beta} > \frac{1}{2};$$

hence for  $(k+1)$  we have also the rough inequalities

$$(8.17) \quad \frac{1}{2} \log T \leq k+1 \leq 2 \log T;$$

i. e., by the restriction (8.16), (6.5) is automatically fulfilled. Owing to (8.15) we have then certainly

$$(8.18) \quad k \geq 10.$$

Finally let the quantity  $\xi$  be given by

$$(8.19) \quad \xi = e^{k+1};$$



hence, by this choice, (5.1) is satisfied.<sup>9</sup> Hence there is no hindrance to apply the inequality (6.3) with all integer values  $k$  satisfying (8.16) for all "good" values  $s$  of 7.

9. We have defined in 7 "good" resp. "bad" subintervals of the interval (6.6). Correspondingly, we define "good" resp. "bad" horizontal strips as those in which the imaginary part  $t$  of the variable belongs to a good resp. bad subinterval  $l_j$ . Since the number of zeros of  $\zeta(s)$  in every bad strip (as a matter of fact, in every such strip in  $T \leq t \leq 2T$ ) is, as known,

$$< c_{29} \log T,$$

the "bad" strips for  $T \leq t \leq 2T$  contain at most

$$(9.1) \quad 2c_{29} T^\beta \log^5 T$$

zeros. We assert further that the "good" strips do not contain zeros for

$$(9.2) \quad \sigma \geq 1 - \left(\frac{1}{2} - \eta\right) \beta.$$

If we assume this as proved, this means that

$$(9.3) \quad N\left(1 - \left(\frac{1}{2} - \eta\right) \beta, 2T\right) - N\left(1 - \left(\frac{1}{2} - \eta\right) \beta, T\right) < 2c_{29} T^\beta \log^5 T.$$

The restrictions upon  $T$  were all of type  $T > c_{30}(\eta)$ ; if we choose  $T$  so large that

$$T^{1/s} > c_{31}(\eta),$$

then (9.3) can be applied replacing  $T$  by  $\frac{T}{2}, \frac{T}{2^2}, \dots, \frac{T}{2^\nu}$ , where  $\nu$  is uniquely determined by

$$\frac{1}{2} T^{1/s} \leq \frac{T}{2^\nu} < T^{1/s}.$$

Adding the resulting inequalities, we obtain, owing to (8.11),

$$N\left(1 - \left(\frac{1}{2} - \eta\right) \beta, T\right) < c_{32}(\eta) T^\beta \log^5 T + N\left(1 - \left(\frac{1}{2} - \eta\right) \beta, T^{1/s}\right).$$

<sup>9</sup> All applications of the method in question to the zeta-function referred to sums of the form  $\left| \sum_{\ell} \xi^{s_0 - a} \left(\frac{s_0 - a}{s_0 - \ell}\right)^{k+1} \right|$ , where the sum contains "not too many" zeta-roots,  $s_0$  and  $a$  are fixed,  $\xi$  and  $k+1$  to be chosen suitably, to make the sum "big". By the choice (8.19) we reduce the sum to a type  $\left| \sum_{j=1}^n z_j^{k+1} \right|$  where only *one* parameter,  $k+1$ , is at our disposal. So we certainly lose something by this step; this loss one could perhaps avoid by developing a similar theory of "double powersums"  $\sum_{\nu=1}^n b_\nu z_\nu^x w_\nu^y$ .



Applying e. g. CARLSON'S theorem to the estimation of the last term, we obtain

$$N\left(1 - \left(\frac{1}{2} - \eta\right)\beta, T\right) < c_{33}(\eta) T^\beta \log^5 T$$

which is, owing to (8. 10), only another form of our theorem, since

$$\frac{2}{1 - 2\eta} < 2(1 + 3\eta).$$

10. Hence we have only to prove the assertion (9. 2) for "good" strips. Suppose this is not true; then we have a "good" strip containing a zero  $\rho^* = \sigma^* + it^*$  with

$$(10. 1) \quad \sigma^* \geq 1 - \left(\frac{1}{2} - \eta\right)\beta.$$

Owing to the definition of a "good" strip, its intersection with the line  $\sigma = \sigma_0$  contains a "good" point  $s$ . Taking this value  $s$  and writing

$$(10. 2) \quad s = s_0 = \sigma_0 + it_0,$$

we start from the inequality (6. 3). We make use, on the other hand, of the representation

$$(10. 3) \quad \frac{1}{k!} \sum_{n \geq \xi} \frac{A(n)}{n^s} \log^k \frac{n}{\xi} = \frac{\xi^{1-s}}{(s-1)^{k+1}} - \sum_{\rho} \frac{\xi^{\rho-s}}{(s-\rho)^{k+1}} - \sum_{n=1}^{\infty} \frac{\xi^{-2n-s}}{(s+2n)^{k+1}}$$

valid for  $\sigma > 1$ ; this could be deduced from the representation (5. 3), following by and large RIEMANN'S contour-integration argument. Hence

$$\left| \frac{\xi^{1-s_0}}{(s_0-1)^{k+1}} - \sum_{\rho} \frac{\xi^{\rho-s_0}}{(s_0-\rho)^{k+1}} - \sum_{n=1}^{\infty} \frac{\xi^{-2n-s_0}}{(s_0+2n)^{k+1}} \right| < \\ < c_{34}(\eta) \frac{\log^2 \xi}{T^{\frac{\beta}{2}}} \left\{ \sqrt{T} \frac{\xi^{\frac{1}{2}-\sigma_0}}{\left(\sigma_0 - \frac{1}{2}\right)^{k+1}} + \frac{\xi^{1-\sigma_0}}{(\sigma_0-1)^{k+1}} \right\}.$$

Multiplying on both sides by

$$\left| \xi^{s_0-\rho^*} (s_0-\rho^*)^{k+1} \right| < \xi^{\sigma_0-\sigma^*} \left\{ (\sigma_0-\sigma^*) + \frac{1}{\log^3 T} \right\}^{k+1},$$

we obtain

$$\left| \xi^{1-\rho^*} \left(\frac{s_0-\rho^*}{s_0-1}\right)^{k+1} - \sum_{\rho} \xi^{\rho-\rho^*} \left(\frac{s_0-\rho^*}{s_0-\rho}\right)^{k+1} - \sum_{n=1}^{\infty} \xi^{-2n-\rho^*} \left(\frac{s_0-\rho^*}{s_0+2n}\right)^{k+1} \right| < \\ (10. 4) \quad < c_{34}(\eta) \frac{\log^2 \xi}{T^{\frac{\beta}{2}}} \left\{ \sqrt{T} \frac{\xi^{\frac{1}{2}-\sigma^*}}{\left(\sigma_0 - \frac{1}{2}\right)^{k+1}} + \xi^{1-\sigma^*} \left(\frac{\sigma_0-\sigma^*}{\sigma_0-1} + \frac{1}{\log^3 T}\right)^{k+1} \right\}.$$

This will be the inequality from which, by use of (10. 1) and suitable choice of  $k$  we shall get a contradiction.



11. First we shall simplify the right-side of this inequality. We remark first that owing to (10. 1), (8. 10) and (8. 8) we have

$$\sigma^* > \vartheta + \eta^3 > \frac{1}{2} + \eta^3,$$

i. e. using (8. 14)

$$\sigma^* > \frac{1}{2} + \frac{1}{\log^3 T}$$

and thus

$$(\sigma_0 - \sigma^*) + \frac{1}{\log^3 T} < \sigma_0 - \frac{1}{2}.$$

Hence

$$\begin{aligned} \sqrt{T} \xi^{\frac{1}{2} - \sigma^*} \left( \frac{\sigma_0 - \sigma^* + \frac{1}{\log^3 T}}{\sigma_0 - \frac{1}{2}} \right)^{k+1} &< \sqrt{T} \xi^{\frac{1}{2} - \sigma^*} = \\ &= \sqrt{T} e^{\left(\frac{1}{2} - \sigma^*\right)(k+1)} < \sqrt{T} \left( e^{-\frac{1}{2} + \left(\frac{1}{2} - \eta\right)\beta} \right)^{k+1}, \end{aligned}$$

using the definition of  $\xi$  and (10. 1). Using (8. 16), this is less than

$$(11. 1) \quad \sqrt{T} \exp \left\{ \frac{(1-2\eta)\beta - 1}{2} \cdot \frac{1 - (1-\eta)\beta}{1 - (1-2\eta)\beta} \log T \right\} = T^{\frac{1-\eta}{2}\beta}.$$

As to the second term on the right of (10. 4) we see using again (10. 1) and (8. 19) that this is less than

$$\left\{ e^{\left(\frac{1}{2} - \eta\right)\beta} \frac{(\sigma_0 - 1) + \left(\frac{1}{2} - \eta\right)\beta + \frac{1}{\log^3 T}}{\sigma_0 - 1} \right\}^{k+1}.$$

Since owing to (8. 14) and (8. 11)

$$\frac{1}{\log^3 T} < \frac{1}{2} \cdot \frac{\eta}{1-2\eta} \left( \frac{\eta}{100} \right)^{100} \cong \frac{\eta}{4} \beta,$$

it follows further that the term in question is less than

$$(11. 2) \quad \left\{ e^{\left(\frac{1}{2} - \eta\right)\beta} \left( 1 + \frac{1}{2} \frac{1 - \frac{3}{2}\eta}{\sigma_0 - 1} \beta \right) \right\}^{k+1} < \left\{ e^{\left(\frac{1}{2} - \eta\right)\beta + \frac{1}{2} \frac{1}{\sigma_0 - 1} \left(1 - \frac{3}{2}\eta\right)\beta} \right\}^{k+1}.$$

Using now the upper estimation of  $(k+1)$  in (8. 16), the exponent in (11. 2) is less than

$$\frac{1}{2} \left( (1-2\eta)\beta + \beta\eta \left( 1 - \frac{3}{2}\eta \right) \right) \cdot \frac{1-\eta}{1-2\eta + \eta \left( 1 - \frac{3}{2}\eta \right)} \log T = \beta \frac{1-\eta}{2} \log T.$$



From this and (11.1) we obtain the following consequence of (10.4):

$$(11.3) \quad \left| \xi^{1-\varrho^*} \left( \frac{s_0 - \varrho^*}{s_0 - 1} \right)^{k+1} - \sum_{\varrho} \xi^{\varrho - \varrho^*} \left( \frac{s_0 - \varrho^*}{s_0 - \varrho} \right)^{k+1} - \sum_{n=1}^{\infty} \xi^{-2n - \varrho^*} \left( \frac{s_0 - \varrho^*}{s_0 + 2n} \right)^{k+1} \right| < \\ < c_{35}(\eta) T^{-\frac{\eta\beta}{2}} \log^2 T.$$

12. Some terms on the left can be estimated trivially. Since from (8.19) and (8.17),  $\xi \leq T^2$ , the first term on the left is absolutely less than

$$(12.1) \quad T \left( \frac{2 + \frac{1}{\eta}}{T} \right)^{10} < \frac{1}{T}$$

using also (8.18) and (8.15). Also roughly the last sum on the left in (11.3) is absolutely less than

$$\frac{1}{\xi^2} \sum_{n=1}^{\infty} \left( \frac{\sigma_0 + 1}{\sigma_0 + 2n} \right)^{k+1} < \frac{1}{\xi^2} \left\{ 2\sigma_0 + \sum_{n > 2\sigma_0} \left( \frac{\sigma_0 + 1}{\sigma_0 + 2n} \right)^{k+1} \right\};$$

since from (8.17) and (8.19)  $\xi > \sqrt{T}$  and further

$$\frac{\sigma_0 + 1}{\sigma_0 + 2n} < \frac{\sigma_0}{n},$$

so we get

$$(12.2) \quad \left| \sum_{n=1}^{\infty} \xi^{-2n - \varrho^*} \left( \frac{s_0 - \varrho^*}{s_0 + 2n} \right)^{k+1} \right| < \frac{1}{T} \left( 2\sigma_0 + \frac{\sigma_0^{k+1}}{(2\sigma_0 - 1)^k} \right) < \frac{c_{36}(\eta)}{T}.$$

Now there remained only one sum on the left of (11.3); we estimate roughly the  $V$ -contribution of those zeros  $\varrho = \sigma_{\varrho} + it_{\varrho}$  for which

$$(12.3) \quad |t_{\varrho} - t_0| \geq 3\sigma_0.$$

We shall make use of the known fact that for all real values  $\tau$  the number of  $\varrho$ 's in

$$0 < \sigma < 1, \quad \tau \leq t \leq \tau + 1$$

is less than

$$(12.4) \quad c_{37} \log(2 + |\tau|).$$

It is sufficient to investigate the  $\varrho$ 's with  $t_{\varrho} \geq t_0 + 3\sigma_0$ ; the part corresponding to  $t_{\varrho} \leq t_0 - 3\sigma_0$  can be dealt with analogously. Using (10.1) we obtain

$$(12.5) \quad |V| < c_{38} \xi^{\left(\frac{1}{2} - \eta\right)\beta} \sum_{n \geq 3\sigma_0} \left( \frac{\sigma_0}{n} \right)^{k+1} \log(2T + n + 2) < c_{39} e^{\frac{k+1}{2}\beta \log T} \frac{\log T}{3^{k+1}} < \\ < c_{40} e^{-\frac{k+1}{2}} \log T < c_{40} T^{-\frac{1}{4}} \log T$$

owing to (8.17).

Next we estimate, again roughly, the contribution  $W$  of the zeros belonging to the domain

$$\sigma \leq 1 - \frac{\beta}{2}, \quad |t - t_0| \leq 3\sigma_0.$$



We remark that from (10. 1), (8. 14) and (8. 11) it follows

$$\begin{aligned} |s_0 - \varrho^*| &\leq \sigma_0 - 1 + \left(\frac{1}{2} - \eta\right)\beta + \frac{1}{\log^3 T} < \\ &< \sigma_0 - 1 + \left(\frac{1}{2} - \eta\right)\beta + \frac{\eta\beta}{2} < \sigma_0 - 1 + \frac{\beta}{2} \leq |s_0 - \varrho|, \end{aligned}$$

i. e. using again (12. 4)

$$|W| < c_{41}(\eta) \xi^{1 - \frac{\beta}{2} - (1 - (\frac{1}{2} - \eta)\beta)} \log T.$$

Taking into account the definition of  $\xi$  and also (8. 17), it follows

$$(12. 6) \quad |W| < c_{41}(\eta) e^{-(k+1)\eta\beta} \log T \leq c_{41}(\eta) T^{-\frac{\eta\beta}{2}} \log T.$$

Collecting (11. 3), (12. 1), (12. 2), (12. 5) and (12. 6) we obtained, inserting the explicit expression of  $\xi$

$$(12. 7) \quad Z \equiv \left| \sum_{\substack{|t_\varrho - t_0| \leq 3\sigma_0 \\ \sigma_\varrho \geq 1 - \frac{\beta}{2}}} \left( e^{\varrho - \varrho^*} \frac{s_0 - \varrho^*}{s_0 - \varrho} \right)^{k+1} \right| < c_{42}(\eta) T^{-\frac{\eta\beta}{2}} \log^2 T.$$

13. The integer  $k$  was subjected so far only to the restriction (8. 16). Now we determine its value exactly so that we should get a *lower* estimation for  $Z$  which will lead to a contradiction with (12. 7). This will be performed by the following theorem:<sup>10</sup>

If

$$(13. 1) \quad \max_{j=1, \dots, K} |z_j| \geq 1$$

with  $K \leq N$  and  $m > 0$  is given arbitrarily, then there is an integer  $\nu_0$  such that

$$(13. 2) \quad m \leq \nu_0 \leq m + N$$

and

$$(13. 3) \quad |z_1^{\nu_0} + z_2^{\nu_0} + \dots + z_K^{\nu_0}| \geq \left( \frac{1}{48e^2} \cdot \frac{N}{m + 2N} \right)^N.$$

The sum in  $Z$  has evidently the character  $z_1^{k+1} + \dots + z_K^{k+1}$  where the role of the numbers  $z_j$  is played by the numbers

$$e^{\varrho - \varrho^*} \frac{s_0 - \varrho^*}{s_0 - \varrho}$$

and  $K$  is the number of  $\varrho$ 's in the domain

$$(13. 4) \quad \sigma \geq 1 - \frac{\beta}{2}, \quad |t - t_0| \leq 3\sigma_0.$$

<sup>10</sup> See l. c. <sup>3</sup>, p. 52. with  $b_1 = b_2 = \dots = b_k = 1$ .



Since  $\varrho^*$  is among the  $\varrho$ 's in question, (13. 1) is fulfilled indeed. We choose

$$(13. 5) \quad m = \frac{1 - (1 - \eta)\beta}{1 - (1 - 2\eta)\beta} \log T \quad (< 2 \log T).$$

14. We need in applying the theorem (13. 1)—(13. 3) an upper estimation for  $K$ . Here and only here will the modified LINDELÖF'S conjecture be used. To get this upper estimation we shall use JENSEN'S inequality in the form that, if  $f(s)$  is regular for  $|s - s'| \leq R$  and  $0 < r < R$ , then the number of zeros of  $f(s)$  in the circle  $|s - s'| \leq r$  does not exceed

$$(14. 1) \quad \frac{1}{\log \frac{R}{r}} \max_{|s - s'| \leq R} \log \left| \frac{f(s)}{f(s')} \right|.$$

Let

$$U > \frac{2}{\eta^3} + 1$$

and we consider for  $T \leq U \leq 2T$  the quantity

$$N(\mathcal{G} + \eta^3, U + 1) - N(\mathcal{G} + \eta^3, U).$$

To apply (14. 1) we choose

$$s' = \frac{1}{\eta^3} + i \left( U + \frac{1}{2} \right), \quad R = \frac{1}{\eta^3} - \mathcal{G}, \quad f(s) = \zeta(s).$$

Since then  $|f(s')| > \frac{1}{2}$ , we obtain using (4. 3) that for  $|s - s'| \leq R$

$$(14. 2) \quad \log \left| \frac{f(s)}{f(s')} \right| < \log \left\{ 2c_{15}(\eta) \left( U + \frac{1}{2} + \frac{1}{\eta^3} \right)^{\eta^{1000}} \right\} < c_{43}(\eta) + \eta^{1000} \log T.$$

All points of the parallelogram  $\mathcal{A}$  defined by

$$(14. 3) \quad 1 \leq \sigma \leq \mathcal{G} + \eta^3, \quad U \leq t \leq U + 1$$

have a distance from  $s'$  not greater than

$$(14. 4) \quad \sqrt{\frac{1}{4} + \left( \frac{1}{\eta^3} - \mathcal{G} - \eta^3 \right)^2}.$$

Choosing this as  $r$  we have

$$\begin{aligned} \frac{R}{r} &= \frac{\frac{1}{\eta^3} - \mathcal{G}}{\sqrt{\frac{1}{4} + \left( \frac{1}{\eta^3} - \mathcal{G} - \eta^3 \right)^2}} = \frac{1}{\sqrt{\left( 1 - \frac{\eta^6}{1 - \mathcal{G}\eta^3} \right)^2 + \frac{1}{4} \frac{\eta^6}{(1 - \mathcal{G}\eta^3)^2}}} \cong \\ &\cong \frac{1}{\sqrt{1 - \frac{\eta^6}{1 - \mathcal{G}\eta^3} \left( \frac{3}{2} - \frac{\eta^6}{1 - \mathcal{G}\eta^3} \right)}} > \left( 1 - \frac{5}{4} \frac{\eta^6}{1 - \mathcal{G}\eta^3} \right)^{-\frac{1}{2}} > \frac{1}{\sqrt{1 - \eta^6}} > e^{\frac{1}{2}\eta^6}. \end{aligned}$$



Owing to (14.4) and (14.2) we have

$$\sum_{\rho \in \Delta} 1 < \sum_{|e^{-s'}| \leq r} 1 < \frac{2}{\eta^6} \{c_{43}(\eta) + \eta^{1000} \log T\} = c_{44}(\eta) + 2\eta^{994} \log T < < c_{44}(\eta) + \eta^{993} \log T.$$

But we need the estimation of the number of zeros in the domain (13.4). We assert that

$$1 - \frac{\beta}{2} \geq \vartheta + \eta^3,$$

i. e. the parallelogram

$$1 \geq \sigma \geq 1 - \frac{\beta}{2}, \quad U \leq t \leq U + 1$$

is contained in that of (14.3). For we have owing to (8.3)

$$\begin{aligned} 4\eta &\geq \eta^3 + 2\eta(1 - \vartheta - \eta^3), \\ 1 - \vartheta - 4\eta &\leq (1 - 2\eta)(1 - \vartheta - \eta^3), \\ \frac{2}{1 - 2\eta}(1 - \vartheta - 4\eta) &\leq 2(1 - \vartheta - \eta^3), \end{aligned}$$

i. e. owing to (8.11)

$$\beta \leq 2(1 - \vartheta - \eta^3), \quad 1 - \frac{\beta}{2} \geq \vartheta + \eta^3,$$

indeed. Hence

$$K < (2 + 6\sigma_0)(c_{44}(\eta) + \eta^{993} \log T) < c_{45}(\eta) + \eta^{991} \log T < \eta^{990} \log T$$

for  $T > c_{46}(\eta)$ , using (8.9). Hence we may choose

$$(14.5) \quad N = \eta^{990} \log T.$$

Now we choose for the value  $(k + 1)$  in (12.7) that given by (13.1)—(13.3). We have to verify that (8.16) is by this choice satisfied. By (13.5) we have only to show that

$$N < \left( \frac{1 - \eta}{1 - 2\eta + \eta \left(1 - \frac{3}{2} \eta\right)} - \frac{1 - (1 - \eta)\beta}{1 - (1 - 2\eta)\beta} \right) \log T,$$

or, using (14.5), that

$$(14.6) \quad \eta^{990} < \frac{1 - \eta}{1 - 2\eta + \eta \left(1 - \frac{3}{2} \eta\right)} - \frac{1 - (1 - \eta)\beta}{1 - (1 - 2\eta)\beta}.$$

But this is true, since from (8.11) we conclude

$$\beta < \frac{2}{1 - 2\eta}(1 - \vartheta) < \frac{1}{1 - 2\eta},$$



i. e.

$$\begin{aligned}
 r_i^{990} &< \frac{3}{2} r_i^2 < \frac{\frac{3}{2} r_i^2 + \beta(1-r_i)r_i \left(1 - \frac{3}{2} r_i\right)}{\left(1-r_i - \frac{3}{2} r_i^2\right) (1-(1-2r_i)\beta)} = \\
 &= \frac{\frac{3}{2} r_i^2 + \beta \left\{ r_i(1-r_i) - \frac{3}{2} r_i^2(1-r_i) \right\}}{\left(1-r_i - \frac{3}{2} r_i^2\right) (1-(1-2r_i)\beta)} = \\
 &= \frac{\frac{3}{2} r_i^2 + \beta \left\{ (1-r_i)^2 - (1-r_i)(1-2r_i) - \frac{3}{2} r_i^2(1-r_i) \right\}}{\left(1-r_i - \frac{3}{2} r_i^2\right) (1-(1-2r_i)\beta)} = \frac{1-r_i}{1-r_i - \frac{3}{2} r_i^2} \\
 &= \frac{1-(1-r_i)\beta}{1-(1-2r_i)\beta} = \frac{1-r_i}{1-2r_i+r_i\left(1-\frac{3}{2} r_i\right)} = \frac{1-(1-r_i)\beta}{1-(1-2r_i)\beta},
 \end{aligned}$$

indeed. Hence (13.3) gives at once, using (13.5) and choosing in (12.7)  $k+1 = r_0$ ,

$$Z > \left( \frac{1}{48e^2} \cdot \frac{r_i^{990}}{2+2r_i^{990}} \right)^{r_i^{990} \log T},$$

or a fortiori

$$Z > \left( \frac{r_i^{990}}{1440} \right)^{r_i^{990} \log T} = T^{-r_i^{990} \log \frac{1440}{r_i^{990}}}.$$

This and (12.7) give

$$(14.7) \quad T^{-r_i^{990} \log \frac{1440}{r_i^{990}}} < c_{42}(r_i) T^{-\frac{r_i \beta}{2}} \log^2 T.$$

15. But it is easy to show that for  $T > c_{47}(r_i)$  this is not true. From (14.7) we conclude, by using (8.11), to the inequality

$$(14.8) \quad T^{r_i^{990} \log \frac{1440}{r_i^{990}}} > \frac{1}{c_{42}(r_i)} T^{\frac{100}{1-2r_i}} \left( \frac{r_i}{100} \right)^{101} \log^{-2} T.$$

But (8.13) gives

$$\frac{1}{\log^2 T} > T^{-\frac{1}{1-2r_i}} \left( \frac{r_i}{100} \right)^{101},$$

and for  $T > c_{48}(r_i)$

$$\frac{1}{c_{42}(r_i)} > T^{-\frac{1}{1-2r_i}} \left( \frac{r_i}{100} \right)^{101},$$

i. e., from (14.8) for  $T > c_{49}(r_i)$

$$T^{r_i^{990} \log \frac{1440}{r_i^{990}}} > T^{\frac{98}{1-2r_i}} \left( \frac{r_i}{100} \right)^{101} > T \left( \frac{r_i}{100} \right)^{101}.$$



or

$$t_1^{889} \log \frac{1440}{t_1^{990}} > \frac{1}{100^{101}}$$

which contradicts (8.5).

*(Received 14 June 1954)*

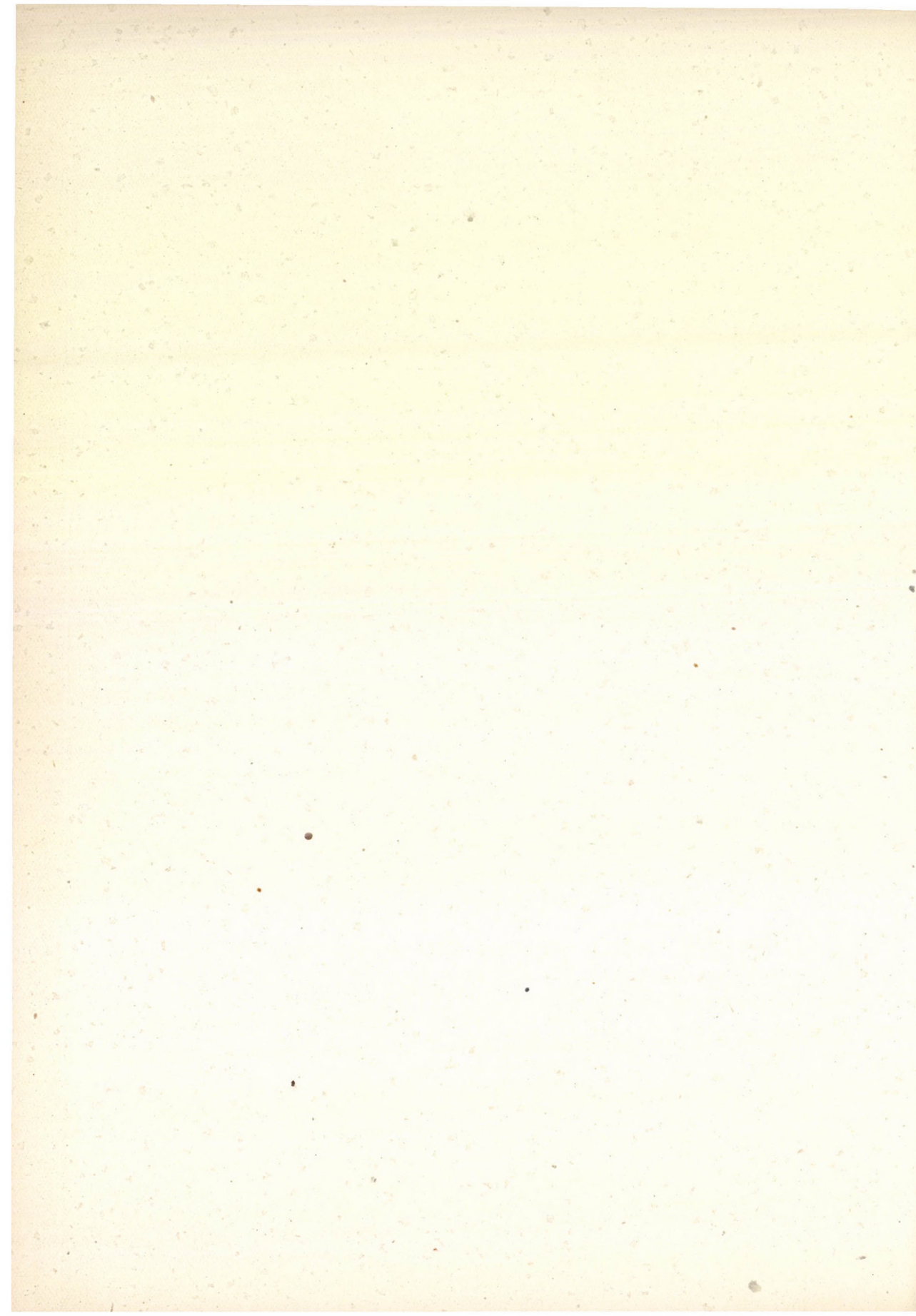
## О ГИПОТЕЗЕ ЛИНДЕЛЕФА

П. ТУРАН (Будапешт)

## (Резюме)

В своей книге<sup>3</sup> автор ввел новый метод и дал там же различные применения этого метода. В настоящей работе этот метод применяется к дальнейшему изучению  $\zeta$ -функции Римана, а именно к изучению т. н. гипотезы Линделефа. Автор доказывает известную теорему Ингэма в немного усиленной форме. Согласно этой теореме, если при выполнении (4.1) оценка (4.3) при произвольно малом  $\eta > 0$  имеет силу для  $\zeta$ -функции Римана в области (4.2), то при выполнении (4.4) имеется оценка (4.5) для числа корней. Дается краткое обсуждение возможностей усовершенствования доказательства.







# KONVEXES POLYEDER ALS GEOMETRISCHER ORT

Von

GYULA SZ.-NAGY (Szeged), Mitglied der Akademie

(Aus dem Nachlaß des Verfassers bearbeitet durch L. FEJES TÓTH)

Wir betrachten den geometrischen Ort derjenigen Punkte des euklidischen Raumes, deren Abstandssumme von  $n$  festen Fundamentebenen einen konstanten Wert aufweist. Werden die Abstände an verschiedenen Seiten der Fundamentebenen mit entgegengesetztem Vorzeichen in Betracht gezogen, so ist natürlich der betrachtete Ort eine Ebene oder der ganze Raum. Wir wollen aber hier den absoluten Betrag des obigen algebraischen Abstandes, d. h. den effektiven Abstand, ins Auge fassen. Dann erweist sich unser Ort im allgemeinen als der Rand eines konvexen Vielflachs.

Wir fassen unser Resultat im folgenden Satz zusammen:

*Es seien im euklidischen Raum  $n$  feste Ebenen  $E_1, \dots, E_n$  vorgegeben, die nicht alle zu einer bestimmten Geraden parallel sind.<sup>1</sup> Dann ist der geometrische Ort  $G(k)$  der Punkte  $P$ , deren Abstandssumme  $A(P)$  von  $E_1, \dots, E_n$  den konstanten Wert  $k > k_0 = \text{Min } A(P)$  aufnimmt, der Rand eines konvexen Polyeders, dessen Kanten auf Fundamentebenen liegen. Der „Kern“  $G(k_0)$  der Örter  $G(k)$  ist ein konvexes Polyeder, das auch in ein Polygon, in eine Strecke oder in einen Punkt entarten kann.*

Das verschiedene Verhalten der Örter  $G(k)$  im Falle  $k > k_0$  bzw.  $k = k_0$  wird plausibel, wenn wir anstatt der Menge  $G(k)$  der Punkte mit  $A(P) = k$  die Menge  $M(k)$  derjenigen Punkte  $P$  ins Auge fassen, für die  $A(P) \leq k$  ausfällt. Wir werden nämlich zeigen, daß  $M(k)$  immer ein konvexes Polyeder ist, das aber für  $k = k_0$  auch entarten kann. Im Falle  $k > k_0$  ist  $M(k)$  eben das von  $G(k)$  begrenzte Polyeder. Im Falle  $k = k_0$  sind aber  $G(k)$  und  $M(k)$  offensichtlich identisch, wodurch verständlich wird, daß  $G(k_0)$  auch eine dreidimensionale Mannigfaltigkeit sein kann.

<sup>1</sup> Sind die Fundamentebenen zu einer Geraden parallel, so läßt sich die Untersuchung unmittelbar auf das entsprechende ebene Problem zurückführen, wenn nur nicht sämtliche Fundamentebenen parallel sind und die ganze Fragestellung sich sogar auf den linearen Fall reduziert.



Die Tatsache, daß  $M(k)$  eine konvexe Menge ist, folgt aus der Bemerkung, daß  $A(P)$  als Summe von konvexen Funktionen selbst konvex ist. Gilt daher für zwei Punkte  $P_1$  und  $P_2$ :  $A(P_1) \leq k$  und  $A(P_2) \leq k$ , so gilt  $A(P) \leq k$  auch für jeden Punkt  $P$  der Strecke  $P_1P_2$ . Folglich gehört mit  $P_1$  und  $P_2$  auch die Strecke  $P_1P_2$  zu  $M(k)$ .

$M(k)$  liegt offenbar im Durchschnitt der Parallelschichten der Breite  $2k$ , welche die Fundamentebenen symmetrisch umgeben. Mit Rücksicht auf die Voraussetzung, daß nicht alle Fundamentebenen zu einer Geraden parallel sind, liegt dieser Durchschnitt, und damit auch  $M(k)$ , im Endlichen. Wir beachten ferner, daß die Fundamentebenen den Raum in eine endliche Anzahl<sup>2</sup> von konvexen Raumteilen zerlegen und daß  $A(P)$  in diesen Raumteilen linear verläuft.<sup>3</sup> Folglich wird  $M(k)$  durch eine endliche Anzahl von Ebenen begrenzt.  $M(k)$  erweist sich also als ein konvexes Polyeder, dessen jede Fläche Durchschnitt je einer Ebene und eines der betrachteten Raumteile ist. Daraus folgt, daß die Kanten von  $M(k)$  auf Fundamentebenen liegen.

Wir haben noch zu zeigen, daß  $G(k)$  für  $k > k_0$  mit dem Rande von  $M(k)$  zusammenfällt. Dazu bemerken wir zunächst, daß die Funktion  $A(P)$  auf dem Rande von  $M(k)$  den konstanten Wert  $k$  annimmt. Wäre nämlich in einem Randpunkt  $A(P) < k$ , so würde wegen der Stetigkeit der Funktion  $A(P)$  diese Ungleichung auch in einer Umgebung von  $P$  gelten, also würde eine ganze Umgebung von  $P$  zu  $M(k)$  gehören, was unmöglich ist. Andererseits folgt aus der Konvexität der Funktion  $A(P)$ , daß  $A(P)$  in  $M(k)$  entweder konstant ist, oder im Inneren von  $M(k)$  nur kleinere Werte als  $k$  annimmt. Nun ist die erste Alternative unmöglich, da es wegen  $k > k_0$  in  $M(k)$  Punkte gibt, wo  $A(P) = k_0 < k$  ist. Damit ist unsere Behauptung bewiesen.

Die Tatsache, daß der Kern  $G(k_0) \equiv M(k_0)$  eine 0-, 1-, 2- oder 3-dimensionale Mannigfaltigkeit sein kann, sieht man an Beispielen leicht ein. Gehen die Fundamentebenen z. B. durch einen Punkt  $S$ , so besteht der Kern aus dem einzigen Punkte  $S$ . Sind dagegen die Fundamentebenen Flächenebenen eines regulären Körpers  $K$ , so ist der Kern mit  $K$  identisch. Es seien nun  $E_1, E_2, E_3$  drei Ebenen, die einen regulären unendlichen prismatischen Raumteil begrenzen. In diesem Raumteil ist die Abstandssumme von  $E_1, E_2, E_3$  konstant, etwa  $= a$ , während sie im Äußeren des Prismas  $> a$  ausfällt. Führt man noch eine Ebene  $E_4$  ein, die das Prisma in einem Dreieck  $\triangle$  durchschneidet, so wird die Abstandssumme eines Punktes von  $E_1, E_2, E_3, E_4$  offenbar dann minimal, wenn der Punkt auf  $\triangle$  liegt. In diesem

<sup>2</sup> Diese Anzahl ist nach einem Satz von STEINER höchstens  $\binom{n+1}{3} + n + 1$ . Vgl. E. STEINITZ und H. RADEMACHER, *Vorlesungen über die Theorie der Polyeder* (Berlin, 1934), S. 285–286.

<sup>3</sup>  $A(P)$  kann in einem (im Endlichen liegenden) Raumteil auch den konstanten Wert  $k$  aufnehmen. Dann ist  $M(k)$  mit diesem Raumteil identisch.



Fall ist also der Kern eine zweidimensionale Mannigfaltigkeit. Führt man noch eine Ebene  $E_5$  so ein, daß sie das Dreieck  $\triangle$  in einer Strecke  $s$  trifft, so wird die Abstandssumme eines Punktes von  $E_1, \dots, E_5$  dann minimal, wenn der Punkt auf  $s$  liegt. Also ist der Kern eindimensional.

Damit haben wir unseren Satz bewiesen.

Unser Satz gilt natürlich auch im Falle, wenn die von den Fundamentebenen gemessenen Abstände mit vorgegebenen positiven Gewichten in Betracht genommen werden. Sind unter den Gewichten auch negative vorhanden, so begrenzt  $G(k)$  im allgemeinen kein konvexes Polyeder.

*(Eingegangen am 12. Oktober 1954.)*

## ВЫПУКЛЫЙ МНОГОГРАННИК КАК ГЕОМЕТРИЧЕСКОЕ МЕСТО

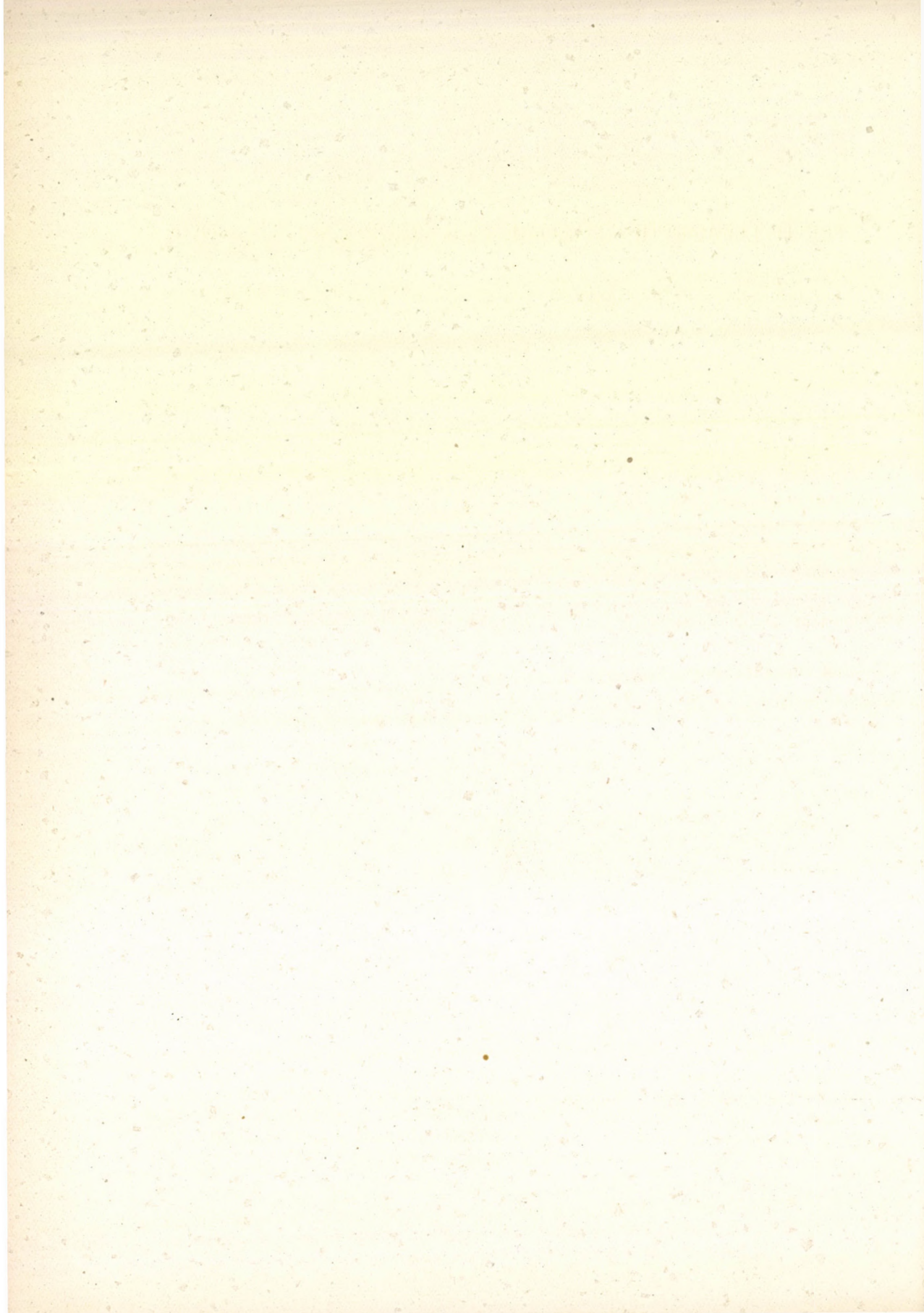
Д. С.-НАДЬ (Сегед)

(Эту посмертную работу автора подготовил к печати Л. Фееш Тот)

(Резюме)

Исследуется геометрическое место тех точек евклидова пространства, для которых сумма расстояний от  $n$  данных плоскостей постоянна; при этом плоскости предполагаются непараллельными одной и той же прямой. Доказывается, что если сумма расстояний минимальна, то геометрическое место может быть только выпуклым многогранником, выпуклым полигоном, отрезком прямой или точкой. Во всех других случаях геометрическое место является границей выпуклого многогранника, грани которого лежат в данных плоскостях.







# DIE HOLOMORPHENTHEORIE FÜR GRUPPEN UND RINGE

Von

LADISLAUS RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

## § 1. Einleitung

Die verschiedenen Zweige der Algebra haben sich mehr oder weniger voneinander unabhängig entwickelt und eine zielbewußte Tätigkeit zur Vereinheitlichung der begrifflich zusammengehörenden Teile ist erst in der neueren Zeit bemerkbar, aber man hat in dieser Beziehung gewiß noch vieles zu tun. Ein wichtiges Mittel dieser Vereinheitlichung besteht im Ausbau von analogen Theorien auf den verschiedenen Forschungsgebieten.

In unserer Arbeit wird es sich um eine neue gruppen-ringtheoretische Analogie handeln, die zwischen der schon fertigen *Holomorphentheorie für Gruppen* und einer erst hier aufzustellenden *Holomorphentheorie für Ringe* besteht. Dabei verstehen wir unter „Holomorphentheorie für Gruppen“ denjenigen Teil der Gruppentheorie, der sich vor allem auf die zwei, bekanntlich eng zusammenhängenden Begriffe „charakteristische Untergruppen“ und „Holomorph“ einer Gruppe, ferner auf den mit diesen ebenfalls eng zusammenhängenden Begriff „vollständige Gruppe“ bezieht. Entsprechend werden auch in der „Holomorphentheorie“ für Ringe die drei analogen Begriffe auftreten. Die ersten zwei dieser Begriffe werden wir ähnlich *charakteristische Unterlinge* bzw. *Holomorphe*<sup>1</sup> eines Ringes nennen (ihre Definition erfolgt später). Es wird eine Überraschung, daß als zum dritten der erwähnten gruppentheoretischen Begriffe, nämlich zur „vollständigen Gruppe“ analoger ringtheoretischer Begriff der „Ring mit Einselement“ auftreten wird.<sup>2</sup>

<sup>1</sup> Die Mehrzahl „Holomorphe“ ist kein Druckfehler! Es wird sich nämlich herausstellen, daß ein Ring im Gegensatz zur Einzigkeit des Holomorphen einer Gruppe im allgemeinen mehrere Holomorphe hat. Diese Mehrheit liegt — wie wir sehen werden — völlig in der Natur und ist etwa dem Umstand zuzuschreiben, daß der Begriff des Ringes mehr zusammengesetzt ist als der der Gruppe.

<sup>2</sup> Die Ringe mit Einselement haben von jeher eine bevorzugte Rolle unter allen Ringen gespielt (z. B. in der Teilbarkeitslehre, Idealtheorie, linearen Algebra usw.). Freilich findet diese Bevorzugung eine formale Erklärung im Umstand, daß die Anwesenheit des Einselementes beträchtliche rechnerische Vereinfachungen nach sich zieht. Demgegenüber erblicken wir in der obigen Analogie sozusagen den *inneren Grund* für die bevorzugte Stellung der Ringe mit Einselement. Diese sollten wir fortan eigentlich auch „vollständige Ringe“ nennen, was wir trotzdem nicht tun werden, raten aber dem Leser, überall statt „Ring mit Einselement“ auch „vollständigen Ring“ zu verstehen.



Gleich wollen wir noch folgendes erwähnen. Bekanntlich spielen in allen Teilen der Algebra gewisse Abbildungen, nämlich der Homo-, Iso-, Auto-, Endo-, Meromorphismus, eine Führerrolle. Als ein wichtiges Ergebnis unserer Untersuchungen wird sich zu diesen fünf Begriffen (ausschließlich für Ringe) ein sechster, nicht minder wichtiger gesellen. Dieser wird gewisse Paare von Abbildungen eines Ringes in sich bedeuten, ein solches Paar von Abbildungen werden wir einen *Doppelhomothetismus* nennen. Die Doppelhomothetismen spielen, wie wir sehen werden, nicht nur in der Holomorphentheorie, sondern auch in anderen Fragen bezüglich Ringe eine Rolle, und bilden, wie unglaublich das auch lautet, ein Analogon zu den Automorphismen einer Gruppe.

Nachdem wir hiermit kurz über den Inhalt unserer Arbeit berichtet haben, wollen wir zur vorherigen Beleuchtung der erst im § 2 beginnenden Untersuchungen einiges über die Analogien zwischen gruppen- und ringtheoretischen Begriffen bemerken.<sup>3</sup> Dementsprechend kann der übrige Teil dieser etwas langen Einleitung flüchtig gelesen werden.

$\Gamma$  und  $P$  bezeichnen durchgehend eine Gruppe bzw. einen Ring.

$P^+$  bezeichnet den Modul von  $P$ , d. h. den durch die Elemente von  $P$  gebildeten Modul.

$A \cdot\cdot B$  bezeichnet, daß  $A, B$  zwei analoge Begriffe sind, von denen  $A$  gruppentheoretisch,  $B$  ringtheoretisch ist.

Eine der wichtigsten Analogien solcher Art ist

- (1) Homomorphismus von  $\Gamma \cdot\cdot$  Homomorphismus von  $P$ .

Mit dieser hängt

- (2) Faktorgruppe von  $\Gamma \cdot\cdot$  Faktoring<sup>4</sup> von  $P$

eng zusammen, da links im wesentlichen (d. h. von Isomorphie abgesehen) die sämtlichen homomorphen Bilder von  $\Gamma$ , rechts die von  $P$  stehen.

Kraft (2) gilt die ebenfalls sehr starke Analogie

- (3) Normalteiler von  $\Gamma \cdot\cdot$  Ideal von  $P$ ,

da die Faktorgruppen und Faktoringe auf bekannte Weise durch die Normalteiler bzw. Ideale angebbar sind.

Die Analogie

- (4) Untergruppe von  $\Gamma \cdot\cdot$  Unterring von  $P$ ,

ist weniger stark, da statt ihrer oft die eine oder andere der (allerdings

<sup>3</sup> Es ist freilich stets eine Geschmackssache, ob man und in welchem Grade gewisse Begriffe als analog betrachtet. Wenn wir deshalb über Analogien sprechen und eventuell auch ihren Grad mit „genau, stark, schwach“ usw. bezeichnen, so soll das stets nur eine *Auffassung* bedeuten.

<sup>4</sup> Wir sagen „Faktoring“ statt „Restklassenring“.



schwächeren) Analogien

(5) Untergruppe von  $\Gamma$   $\cdot \cdot$  Untermodul von  $P$ ,

(6) Untergruppe von  $\Gamma$   $\cdot \cdot$  einseitiges Ideal von  $P$

zu Kraft tritt, ferner kommt in gewissen Untersuchungen mehr die ebenfalls schwache Analogie

(7) Normalteiler von  $\Gamma$   $\cdot \cdot$  einseitiges Ideal von  $P$

(statt (6)) zur Geltung.

Dem Anschein nach ist

(8) Automorphismus von  $\Gamma$   $\cdot \cdot$  Automorphismus von  $P$

eine ähnlich starke Analogie wie (1). Wenn man aber (3) beachtet, und bedenkt, daß sich die Normalteiler von  $\Gamma$  als die gegenüber gewissen (nämlich den inneren) Automorphismen von  $\Gamma$  zulässigen Untergruppen charakterisieren lassen, dagegen eine entsprechende Charakterisierung der Ideale von  $P$  durch Automorphismen unmöglich ist,<sup>5</sup> so fühlt man sich gezwungen, die Analogie (8) abzulehnen. Eine andere „Schwäche“ der Analogie (8) ist, daß bekanntlich eine Gruppe mit mehr als zwei Elementen stets auch nichtidentische Automorphismen hat, wogegen es viele Ringe mit nur einem (dem identischen) Automorphismus gibt.<sup>6</sup> Schon aus diesem Grunde kann die rechte Seite von (8) nicht die ähnlich wichtige Rolle spielen wie die linke Seite, und man weiß auch aus Erfahrungen, daß — im Gegensatz zur fortwährenden Anwendung der Automorphismen in der Gruppentheorie — die Automorphismen eines Ringes sehr selten zu Wort kommen.<sup>7</sup> Nicht weniger entscheidendes gegen die Analogie (8) ist endlich, daß die Automorphismen von  $\Gamma$  eine Gruppe bilden, wogegen die Automorphismen von  $P$  keinen Ring bilden: wegen des ersteren existiert die „Schreiersche Erweiterung von  $\Gamma$  mit seiner vollen Automorphismengruppe“, wegen des letzteren fehlt ein analoger Begriff für  $P$ , der ebenfalls durch  $P$  und seine volle Automorphismengruppe bestimmt wäre.<sup>8</sup>

<sup>5</sup> Die inneren Automorphismen einer Gruppe haben innerhalb der Automorphismen eines Ringes ohne Einselement überhaupt keine Analoga. (Hierauf kommen wir bald zurück.)

<sup>6</sup> Z. B. hat der Ring  $I$  der ganzen Zahlen nur den identischen Automorphismus. Betrachten wir ferner den Polynomring  $I[x]$ . Selbst dieser hat bekanntlich unendlich viele Automorphismen, und zwar lassen sich seine sämtlichen Automorphismen durch  $x \rightarrow x + c$  und  $x \rightarrow -x + c$  angeben ( $c \in I$ ). Bezeichnet aber  $f(x)$  ( $\in I[x]$ ) ein Polynom, wofür alle  $\pm f(\pm x + c)$  verschieden sind, so hat der Ring  $f(x)I[x]$  offenbar nur den identischen Automorphismus. Ähnlich beschaffen ist auch der Ring  $\{f^{-1}(x), I[x]\}$ . — Wie hier so auch später bezeichnet  $\{\dots\}$  die durch die eingeklammerten Elemente erzeugte (algebraische) Struktur.

<sup>7</sup> In (allerdings wenigen) Spezialfällen, so vor allem für gewisse Körper und Schiefkörper sind jedoch die Automorphismen überaus wichtig, wie z. B. in der Theorie von GALOIS. Durch solche Sonderheiten wird jedoch das oben Gesagte nicht beeinflußt.

<sup>8</sup> Diesen jetzt noch fehlenden analogen Begriff werden wir später dennoch finden.



Dagegen spielen die Endomorphismen des Moduls eines Ringes eine ähnlich oftmalige Rolle in der Ringtheorie wie die Automorphismen einer Gruppe in der Gruppentheorie. Deshalb kann

(9) Automorphismus von  $\Gamma$   $\cdot \cdot$  Endomorphismus von  $P^+$

einigermaßen als ein Ersatz für die nach vorigem sehr schwache Analogie (8) angesehen werden. Allerdings ist auch (9) für eine schwache Analogie zu halten schon aus dem Grunde, daß im allgemeinen in die rechte Seite von (9) die in  $P$  definierte Multiplikation gar nicht eingeht.

In Spezialfällen geschieht das doch, denn die Abbildungen

(10) 
$$\varrho \rightarrow \alpha\varrho, \varrho \rightarrow \varrho\alpha \quad (\varrho \in P)$$

definieren für jedes  $\alpha (\in P)$  zwei Endomorphismen von  $P^+$ . Diese nennen wir nach BOURBAKI [1]<sup>9</sup> den durch  $\alpha$  induzierten *inneren Links-* bzw. *Rechtshomothetismus* von  $P$ .<sup>10</sup> Beide nennen wir kurz auch *innere Homothetismen*. Dann gilt der „Teil“<sup>11</sup>

(11) innerer Automorphismus von  $\Gamma$   $\cdot \cdot$  innerer Homothetismus von  $P$

der schwachen Analogie (9) für eine starke Analogie. In der Tat ist (11) mit (3) gut verträglich, denn einerseits sind die Normalteiler von  $\Gamma$  (wie schon erwähnt) die gegenüber allen inneren Automorphismen von  $\Gamma$  zulässigen Untergruppen, andererseits sind die Ideale von  $P$  ähnlich die gegenüber allen inneren Homothetismen von  $P$  zulässigen Unterringe. (Vgl.<sup>5</sup>.)

Damit beenden wir auch schon die Aufzählung der im wesentlichen bekannten Analogien zwischen gruppen- und ringtheoretischen Grundbegriffen, deren Zahl man noch vermehren könnte, aber für unseren Zweck, das spätere zu beleuchten, genügt das bisherige.

Jetzt wollen wir kurz einige neue Untersuchungen erwähnen, in denen Analogien zwischen Gruppen- und Ringtheorie ausgebaut wurden.

EVERETT [3] hat die Schreiersche Erweiterungstheorie für Ringe aufgestellt.<sup>12</sup> Diese zwei analogen Theorien (die Schreiersche und Everettsche) beruhen völlig auf der Analogie (3) und bilden einen äußerst wichtigen Teil der Gruppen- bzw. Ringtheorie.

<sup>9</sup> Mit [ ] verweisen wir an das Literaturverzeichnis am Schluß unserer Arbeit.

<sup>10</sup> Selbst BOURBAKI verwendet die Benennungen „homothétie à gauche“ bzw. „à droite“. Hiernach hätten wir „Homothétie“ statt „Homothetismus“ sagen sollen, uns gefällt aber die Verwendung von „-ismus“ wegen der Ähnlichkeit mit Homo-, Iso-, Automorphismus usw.

<sup>11</sup> Wir nennen (11) in dem Sinne einen Teil von (9), daß beide Seiten von (11) einen Spezialfall der zwei Seiten von (9) bilden.

<sup>12</sup> Bezüglich der Grundlagen der Schreierschen Erweiterungstheorie für Gruppen und Ringe halten wir uns stets an RÉDEI [6]. — Gleichförmigkeitshalber werden wir über Schreiersche statt Everettsche Erweiterungsringe sprechen, ferner sagen wir statt „Schreiersche Erweiterungsgruppe“ und „Schreierscher Erweiterungsring“ in beiden Fällen kurz „Schreiersche Erweiterung“, wenn daraus kein Mißverständnis entstehen kann.



FUCHS [4] hat den wichtigen Begriff der Frattinischen Untergruppe auf den Fall von Gruppen mit Operatoren übertragen und gezeigt, daß im Fall eines  $P$  mit Einselement das Jacobsonsche Radikal von  $P$  mit der Frattinischen Untergruppe von  $P^+$  übereinstimmt, wenn man die inneren Rechtshomothetismen von  $P$  als Operatoren von  $P^+$  deutet.

RÉDEI [9], [10] hat zwei verschiedene ringtheoretische Analoga der Hamiltonschen Gruppen untersucht (die zweite dieser Untersuchungen noch nicht beendet).<sup>13</sup>

Mit den letzten zwei Beispielen sind wir von unserem eigentlichen Gegenstand abgewichen, aber wir kommen jetzt darauf zurück. Und zwar wollen wir statt der schwachen Analogie (8) unter Beibehaltung der linken Seite und passender Veränderung der rechten Seite eine starke Analogie aufstellen. Dabei wollen wir an die vorige Bemerkung anknüpfen, daß (11) eine starke Analogie ist. Deshalb beabsichtigen wir (11) so zu „erweitern“, daß hierdurch die gewünschte Analogie entsteht, was auf eine „richtige“ Verallgemeinerung der rechten Seite von (11) ankommt. Nun liegt diese Verallgemeinerung auf der Hand, so daß man in (10) das Element  $\alpha(\in P)$  durch ein Element  $a$  eines Oberringes  $\bar{P}$  von  $P$  ersetzt, das man aber der Bedingung unterwirft, daß alle Produkte  $a\varrho, \varrho a$  in  $P$  liegen, damit auf diese Weise wirklich Abbildungen von  $P$  in sich entstehen. Da die gesagte Bedingung bedeutet, daß der Ring  $P$  als Ideal in  $\bar{P}$  enthalten, d. h.  $\bar{P}$  eine Schreiersche Erweiterung von  $P$  ist, so greifen wir zur folgenden Definition:

Ist  $a$  ein Element einer Schreierschen Erweiterung von  $P$ , so nennen wir die zwei Abbildungen

$$(12) \quad \varrho \rightarrow a\varrho, \varrho \rightarrow \varrho a \quad (\varrho \in P)$$

den durch  $a$  induzierten *Links-* bzw. *Rechtshomothetismus* von  $P$ .<sup>14</sup> Beide nennen wir kurz auch *Homothetismen*. Es ist klar, daß die inneren Homothetismen wirklich Spezialfälle der Homothetismen sind. Einen Homothetismus von  $P$ , der kein innerer ist, nennen wir einen *äußeren Links-* bzw. *Rechtshomothetismus* von  $P$ .<sup>15</sup>

<sup>13</sup> Bekanntlich nennt man eine Gruppe Hamiltonisch, wenn alle Untergruppen normal sind. Auf Grund der Analogie (3) und der Analogien (4), (5), (6) lassen sich drei verschiedene ringtheoretische Analoga definieren, von denen die ersten zwei von RÉDEI [10], [9] „Vollidealringe im weiteren Sinn“ bzw. „Vollidealringe“ genannt wurden.

<sup>14</sup> Die Möglichkeit nach weiterer Verallgemeinerung liegt vor, so daß man in den zwei Abbildungen (12) nur voraussetzt, daß  $a$  ein Element eines Oberringes von  $P$  ist, der  $P$  als Links-, bzw. Rechtsideal enthält.

<sup>15</sup> BOURBAKI [1] spricht über „homothétie externe“ in wesentlich anderem Sinne. Nur ungern geraten wir durch unsere obige Definition der Homothetismen mit der Definition der „homothétie externe“ von BOURBAKI in einen Konflikt, trotzdem können wir auf die Benennung „Homothetismus“ der obigen Abbildungen nicht verzichten. Wir haben hierzu zwei Gründe. Der eine ist, daß wir in unseren Homothetismen die natürlichste Verallge-



Ferner nennen wir das (geordnete) System der zwei Homothetismen (12) den durch  $a$  induzierten *Doppelhomothetismus* von  $P$ .<sup>16</sup> Gehört dabei  $a$  in  $P$ , so sprechen wir von einem *inneren Doppelhomothetismus* von  $P$ , dagegen verstehen wir unter einem *äußeren Doppelhomothetismus* von  $P$  einen solchen, der kein innerer ist.

Wie oben schon gesagt wurde, werden für uns die Doppelhomothetismen von großer Wichtigkeit sein. Diesbezüglich bemerken wir vor allem, daß an Stelle von (11) sich die noch stärkere Analogie

(13) innerer Automorphismus von  $\Gamma \dots$  innerer Doppelhomothetismus von  $P$  für viel nützlicher erweisen wird. Einerseits gilt nämlich das über (11) bemerkte offenbar auch für (13). Andererseits bilden die inneren Doppelhomothetismen von  $P$  — wie wir sehen werden — nach einfachen Rechnungsregeln einen Ring, der sich als genaues Analogon der Gruppe der inneren Automorphismen von  $\Gamma$  betrachten lassen wird. (Eine der letzteren ähnliche Bemerkung gilt für (11) nicht.)

Ferner besteht als eine „Fortsetzung“ von (13) die unseres Erachtens sehr starke Analogie

(14) Automorphismus von  $\Gamma \dots$  Doppelhomothetismus von  $P$ ,

die wir auch schon oben angemeldet haben; diese mag an Stelle der sehr schwachen Analogie (8), und zwar mit viel mehr Recht als (9), für richtig hingenommen werden. Die eigentliche Rechtfertigung dieser Auffassung wird durch unsere späteren Betrachtungen geliefert, hier beschränken wir uns diesbezüglich auf einige kurze Bemerkungen.

Bekanntlich gewinnt man alle Automorphismen von  $\Gamma$  und nur diese als durch die Elemente  $a$  der Schreierschen Erweiterungen von  $\Gamma$  gelieferten Abbildungen

$$\varrho \rightarrow a^{-1}\varrho a \quad (\varrho \in \Gamma).$$

Die Analogie zwischen dieser Definition und der der Doppelhomothetismen ist auffallend.

Im Gegensatz zu den über die Automorphismen von  $P$  gesagten (vgl. auch<sup>16</sup>) gibt es im Fall  $P \neq 0$  stets mindestens zwei Doppelhomothetismen von  $P$ .

meinerung der inneren Homothetismen erblicken. Der zweite ist, daß unseres Wissens die Benennung „homothétie externe“ (im Sinne von BOURBARI) sich bisher wenig in der Literatur verbreitet hat. [Wir können auch noch einen dritten Grund aufführen. Die Algebra entlehnt nämlich den Begriff „Homothetie“ von der Geometrie in der man unter Homothetie (bis auf Translation) eine Punkttransformation  $x \rightarrow cx$  ( $c$  reell) des  $n$ -dimensionalen Euklidischen Raumes versteht ( $x$  bezeichnet den Ortsvektor des Punktes), die also im wesentlichen in der Verwendung eines „Ähnlichkeitsfaktors“ bedeutet, wie auch unser Homothetismus, wogegen ähnliches für die „homothétie externe“ von BOURBARI nicht gilt.]

<sup>16</sup> Ein Doppelhomothetismus besteht also aus einem Links- und einem Rechtshomothetismus, die aber nicht beliebig sind, sondern beide durch *ein* Element induziert werden können, wie das oben unzweideutig gesagt wurde.



Unter diesen kommen nämlich der *triviale* und der *identische Doppelhomothetismus* stets vor; so nennen wir die beiden Spezialfälle, wo alle Bildelemente  $a\varrho, \varrho a$  gleich 0 (Fall  $a=0$  von (12)) bzw. gleich  $\varrho$  sind.<sup>17</sup>

Die Doppelhomothetismen (ähnlich wie die inneren) werden sich nach einfachen Regeln addieren und multiplizieren lassen. Wohl bildet dann die Menge aller Doppelhomothetismen von  $P$  im allgemeinen keinen Ring, dagegen wird diese Menge gewisse Ringe enthalten, die die ähnliche Rolle im Zusammenhang mit  $P$  spielen, wie die volle Automorphismengruppe von  $I$ , weshalb sie als (genaue) ringtheoretische Analoga dieser Gruppe angesehen werden können. Die gesagten Ringe werden wir verwenden um mit ihnen gewisse Schreiersche Erweiterungen von  $P$  zu bilden, worauf wir schon in<sup>18</sup> angespielt haben; diese Erweiterungsringe werden die Holomorphe von  $P$  sein.

Endlich werde bemerkt, daß die charakteristischen Untergruppen und -ringe ähnlich mit der Analogie (14) zusammenhängen werden, wie auch die Normalteiler und Ideale mit der Analogie (11) (oder (13)) im schon bemerkten Zusammenhang stehen.

Übrigens werden wir im § 2 bei (30) bis (33) die Definition (12) der Doppelhomothetismen durch eine mehr explizite ersetzen.

Unsere Arbeit, was noch übrig ist, gliedert sich so:

Im § 2 machen wir einige Vorbereitungen.

Im § 3 stellen wir die im wesentlichen bekannten Grundlagen der Holomorphentheorie für Gruppen in einer zu unseren Zwecken geeigneten Form zusammen.

Im § 4 erfolgt die Begründung der analogen Theorie für Ringe.

Im § 5 untersuchen wir als Spezialfall die Ringe mit Einselement (als Analogon zu den vollständigen Gruppen).

Im § 6 beschäftigen wir uns mit einigen weiteren Spezialfällen.

Obwohl es sich hauptsächlich um die Theorie der Ringe handeln wird, so werden wir doch auch in der Gruppentheorie wenig neues erzielen.<sup>18</sup>

## § 2. Vorbereitungen

Neben den im § 1 eingeführten Bezeichnungen  $I, P, P^+$  führen wir die folgenden, ebenfalls durchgängigen Bezeichnungen ein:

<sup>17</sup> Das letztere ist auch ein möglicher Fall, der nämlich so entsteht, daß man für  $a$  in (12) das Einselement einer passenden Schreierschen Erweiterung von  $P$  nimmt, die bekanntlich stets existiert. Gleich bemerken wir, daß der triviale Doppelhomothetismus von  $P$  stets ein innerer, dagegen der identische dann und nur dann ein innerer ist, wenn  $P$  ein Einselement enthält.

<sup>18</sup> Das Hauptproblem dieser Arbeit, eine Holomorphentheorie für Ringe zu schaffen, tauchte mir während einer Besprechung mit meinem Aspiranten, Herrn O. STEINFELD, auf. Er und die Herren L. FUCHS, G. POLLÁK, J. SZENDREI, Prof. L. KALMÁR haben mit ihren nützlichen Bemerkungen zur endgültigen Abfassung meiner Arbeit beigetragen, weshalb ich ihnen herzlichst danke.



$\alpha, \beta, \dots$  bezeichnen beliebige Elemente von  $I$  bzw.  $P$ , je nachdem von Gruppen oder Ringen die Rede ist. Insbesondere bezeichnet  $\varepsilon$  das Einselement von  $I$  oder  $P$ , letzteres natürlich nur dann, wenn  $P$  ein Ring mit Einselement ist. Die Null in  $P$  wird mit  $0$  bezeichnet.

$I_*, P_*$  bezeichnen das Zentrum von  $I$  bzw. den Annulator von  $P$ ; unter dem letzteren verstehen wir das Ideal derjenigen  $\nu$ , für die  $\nu P = P \nu = 0$  gilt. Die ähnlichen Bezeichnungen  $I_*, P_*$  verwenden wir deshalb, weil sich

$$(15) \quad (I_* =) \text{ Zentrum von } I \dots (P_* =) \text{ Annulator von } P$$

(im Zusammenhang mit (14)) als eine starke Analogie erweisen wird.<sup>19</sup>

$\alpha_*$  bezeichnet das durch  $\alpha$  repräsentierte Element der Faktorgruppe  $I/I_*$  bzw. des Faktorringes  $P/P_*$ , d. h. die Klasse  $\alpha I_*$  bzw.  $\alpha + P_*$ .

$X < Y$  bezeichnet, daß entweder  $X, Y$  Gruppen sind und  $X$  Untergruppe von  $Y$  ist, oder sie Ringe sind und  $X$  Unterring von  $Y$  ist.

$X \triangleleft Y$  bezeichnet, daß  $X < Y$  gilt und dabei  $X$  normal<sup>20</sup> in  $Y$  ist.

$X \blacktriangleleft Y$  bezeichnet, daß  $X < Y$  gilt und dabei  $X$  charakteristisch<sup>21</sup> in  $Y$  ist.

$a, b, \dots$  bezeichnen entweder Abbildungen von  $I$  in sich oder (geordnete) Paare von Abbildungen, kurz *Doppelabbildungen* von  $P$  in sich; genau gesprochen verstehen wir unter einer Doppelabbildung von  $P$  in sich das System von zwei Abbildungen von  $P$  in sich, die wir als den *ersten* bzw. *zweiten Bestandteil* der Doppelabbildung unterscheiden.

$e$  bezeichnet insbesondere die identische Abbildung von  $I$ , bzw. die (aus zwei identischen Abbildungen bestehende) *identische Doppelabbildung* von  $P$ , je nachdem von  $I$  oder von  $P$  die Rede ist.

Für eine beliebige Abbildung  $a$  von  $I$  in sich bezeichnen wir mit  $\alpha^a$  das entsprechende Bild von  $\alpha$ . Dann ist  $a$  gleich der Abbildung  $\alpha \rightarrow \alpha^a$ . In der Menge dieser Abbildungen definieren wir das Produkt  $ab$  als dasjenige Element der Menge, wofür

$$(16) \quad \alpha^{ab} = (\alpha^a)^b$$

gilt.

Für eine beliebige Doppelabbildung  $a$  von  $P$  in sich bezeichnen wir mit  $aa, \alpha a$  die entsprechenden zwei Bilder von  $\alpha$ , so daß dann  $a$  aus den zwei Abbildungen  $\alpha \rightarrow aa, \alpha \rightarrow \alpha a$  besteht (in *dieser* Reihenfolge). In der Menge dieser Doppelabbildungen definieren wir die Summe  $a + b$  und das Produkt

<sup>19</sup> Dagegen ist

Zentrum von  $I \dots$  Zentrum von  $P$

eine sehr schwache Analogie zu nennen.

<sup>20</sup> Einen Unterring nennen wir normal, wenn er ein Ideal ist. Die Bezeichnung „ $X \triangleleft Y$ “, läßt sich auch so erklären, daß  $Y$  eine Schreiersche Erweiterung von  $X$  ist (mit beliebiger Faktorstruktur  $Y/X$ ).

<sup>21</sup> Wir definieren „charakteristisch“ erst in den §§ 3, 4, weshalb das Zeichen „ $\blacktriangleleft$ “ vorläufig noch nicht verwendet wird.



$ab$  als dasjenige Element der Menge, wofür

$$(17) \quad (a+b)\alpha = a\alpha + b\alpha, \quad \alpha(a+b) = \alpha a + \alpha b,$$

$$(18) \quad ab\alpha = a(b\alpha), \quad \alpha ab = (\alpha a)b$$

gelten.<sup>22</sup>

Für Teilmengen  $H$  von  $\Gamma$  bzw.  $\mathcal{P}$  definieren wir die Bilder  $H^\alpha$  und  $aH, Ha$  wie üblich als die Menge der entsprechenden Bilder der Elemente.

Von was für eine Menge  $M$  von Abbildungen von  $\Gamma$  oder Doppelabbildungen von  $\mathcal{P}$  in sich die Rede sein wird, so werden wir eine Teilmenge  $H$  von  $\Gamma$  bzw.  $\mathcal{P}$  *zulässig* nennen, wenn  $H'' \subseteq H$  bzw.  $aH, Ha \subseteq H$  ( $a \in M$ ) gilt.

Übrigens werden wir unter allen Abbildungen von  $\Gamma$  in sich nur die Automorphismen von  $\Gamma$  in Betracht ziehen. Diese bilden auf Grund der durch (16) definierten Multiplikation eine Gruppe, die *volle Automorphismengruppe* von  $\Gamma$ . Ihre Untergruppen nennen wir schlechthin die Automorphismengruppen von  $\Gamma$ .

Ferner werden wir unter allen Doppelabbildungen von  $\mathcal{P}$  in sich nur diejenigen betrachten, die aus zwei Endomorphismen von  $\mathcal{P}^+$  bestehen; diese nennen wir die *Doppelendomorphismen* von  $\mathcal{P}^+$ . Hiervon bilden die Doppelhomothetismen von  $\mathcal{P}$  nach der Definition (12) offenbar einen Spezialfall. Nur dieser Spezialfall wird uns später interessieren, aber wir haben auch auf den allgemeinen Fall einen Blick zu werfen.

Zu diesem Zweck bezeichnen wir einen Augenblick mit  $R$  den vollen Endomorphismenring von  $\mathcal{P}^+$ . Das ist, wie üblich, so gemeint, daß man insbesondere das Produkt  $AB$  von zwei Elementen von  $R$  durch  $AB\alpha = A(B\alpha)$  definiert. Mit  $R'$  bezeichnen wir den zu  $R$  inversen Ring, der nämlich aus  $R$  so entsteht, daß man von der Multiplikation  $AB$  auf die „inverse Multiplikation“  $BA$  übergeht. Dann ersieht man aus (17), (18) sofort, daß die sämtlichen Doppelendomorphismen von  $\mathcal{P}^+$  einen Ring bilden, den wir den *vollen Doppelendomorphismenring* von  $\mathcal{P}^+$  nennen, und zwar daß dieser die direkte Summe des vollen Endomorphismenringes von  $\mathcal{P}^+$  und des hierzu inversen Ringes ist.<sup>23</sup> Alle später aufzutretenden Doppelhomothetismenringe von  $\mathcal{P}$ , was wir hier ein für allemal schon im voraus bemerken wollen, werden lauter Unterringe des vollen Doppelendomorphismenringes von  $\mathcal{P}^+$  sein. Nur mit solchen Unterringen dieses Ringes werden wir es zu tun haben.

Wir setzen noch unsere Vorbereitungen fort, wobei wir nunmehr zwei Fälle unterscheiden.

<sup>22</sup> Wenn wir etwa  $aba\beta$  oder  $a\beta ab$  schreiben, so meinen wir damit  $(ab)(a\beta)$  bzw.  $(a\beta)(ab)$ . Ähnlich ist die linke Seite der Gleichungen (18) zu verstehen.

<sup>23</sup> Man bemerke, daß der Begriff des vollen Doppelendomorphismenringes von  $\mathcal{P}^+$  nicht von der in  $\mathcal{P}$  gültigen Multiplikation abhängt (also für jeden Modul statt  $\mathcal{P}^+$  einen Sinn hat).



**Fall  $\Gamma$ :**

Ist  $A$  eine (nicht notwendig die volle) Automorphismengruppe von  $\Gamma$ , so definieren wir in der Menge aller Paare

$$(19) \quad (a, \alpha) \quad (a \in A, \alpha \in \Gamma)$$

das Produkt

$$(20) \quad (a, \alpha)(b, \beta) = (ab, \alpha^b \beta).$$

So entsteht wegen (16) eine Gruppe mit dem Einselement  $(e, e)$ , die wir mit

$$(21) \quad A \bullet \Gamma$$

bezeichnen. Diese Gruppe ist nämlich (vgl. RÉDEI [6] Korollar 1 von Satz 1, S. 260) eine faktorenfreie (also zerfallende) Schreiersche Erweiterung von  $\Gamma$  mit  $A$ , die wir deshalb kurz die *A zugehörige zerfallende Erweiterung* von  $\Gamma$  nennen.<sup>24</sup> In dieser bilden die Elemente  $(e, \alpha)$  einen mit  $\Gamma$  isomorphen Normalteiler, demgemäß wir die übliche Einbettung  $(e, \alpha) \rightarrow \alpha$  stets ausgeführt denken, wodurch

$$(22) \quad \Gamma \triangleleft (A \bullet \Gamma)$$

in Erfüllung geht:

Ist  $A$  insbesondere eine *innere Automorphismengruppe* (d. h. eine Gruppe von inneren Automorphismen) von  $\Gamma$ , so können wir der Gruppe (21) eine mehr explizite Form geben. Zu diesem Zweck betrachten wir den durch  $\alpha$  induzierten inneren Automorphismus

$$(23) \quad \rho \rightarrow \alpha^{-1} \rho \alpha$$

von  $\Gamma$ . Dieser hängt nur von der Klasse  $\alpha_* = \alpha \Gamma_*$  ab, weshalb wir ihn dieser Klasse  $\alpha_*$  zuordnen können. Diese Zuordnung ist auch rückwärts eindeutig. Hiernach dürfen wir (23) den „Automorphismus  $\alpha_*$ “ nennen und diesen ebenfalls mit  $\alpha_*$  bezeichnen, woraus kein Mißverständnis entstehen wird. Dann gilt

$$(24) \quad \rho^{\alpha_*} = \alpha^{-1} \rho \alpha,$$

ferner ist auch die (16) entsprechende Bedingung  $\rho^{\alpha_* \beta_*} = (\rho^{\alpha_*})^{\beta_*}$  erfüllt. Folglich ist die gesagte Zuordnung ein Isomorphismus zwischen  $\Gamma/\Gamma_*$  und der *vollen inneren Automorphismengruppe* (d. h. der Gruppe aller inneren Automorphismen) von  $\Gamma$ , weshalb wir die letztere nach Identifizierung der entsprechenden Elemente einfach mit  $\Gamma/\Gamma_*$  bezeichnen dürfen. Also läßt sich obiges  $A$  als eine beliebige Untergruppe von  $\Gamma/\Gamma_*$  annehmen. Dann besteht  $A \bullet \Gamma$  aus den Elementen

$$(25) \quad (\alpha_*, \beta) \quad (\alpha_* \in A \subseteq (\Gamma/\Gamma_*), \beta \in \Gamma)$$

<sup>24</sup> Man beachte, daß es im allgemeinen mehrere, wesentlich verschiedene zerfallende Schreiersche Erweiterungen von  $\Gamma$  mit  $A$  gibt, denn es gehört unter anderem auch das direkte Produkt von  $A$  und  $\Gamma$  darunter, daß aber die Gruppe (21) durch  $A$  und  $\Gamma$  eindeutig bestimmt ist.



und als Produktregel gilt nach (20), (24)

$$(26) \quad (\alpha_*, \beta)(\gamma_*, \delta) = (\alpha_*\gamma_*, \gamma^{-1}\alpha\gamma\delta),$$

wobei man rechts für  $\gamma$  einen beliebigen Repräsentanten der Klasse  $\gamma_*$  einzusetzen hat.

Wenn dabei  $\Gamma$  *zentrumlos* ist (d. h.  $\Gamma_* = \varepsilon$  gilt), so dürfen wir  $\Gamma/\Gamma_* = \Gamma$ ,  $\alpha_* = \alpha$  schreiben. Entsprechend lassen sich jetzt nach (25), (26) die Elemente der Gruppe  $A \bullet \Gamma$  und die Produktregel in ihr durch

$$(27) \quad (\alpha, \beta) \quad (\alpha \in A \subseteq \Gamma, \beta \in \Gamma),$$

$$(28) \quad (\alpha, \beta)(\gamma, \delta) = (\alpha\gamma, \gamma^{-1}\beta\gamma\delta)$$

angeben.

Gleich zeigen wir, daß  $A \bullet \Gamma$  in diesem Fall ein direktes Produkt ist:

$$(29) \quad A \bullet \Gamma \approx A \otimes \Gamma,$$

wobei „ $\approx$ “, „ $\otimes$ “ die Isomorphie bzw. das direkte Produkt bezeichnet. Zu diesem Zweck nehme man die Permutation

$$(\alpha, \beta) \rightarrow \Pi(\alpha, \beta) = (\alpha, \alpha\beta)$$

der Elemente von  $A \bullet \Gamma$  zu Hilfe und gehe von (28) auf die neue Multiplikation

$$(\alpha, \beta) \times (\gamma, \delta) = \Pi(\Pi^{-1}(\alpha, \beta)\Pi^{-1}(\gamma, \delta))$$

über. Die rechte Seite ist wegen  $\Pi^{-1}(\alpha, \beta) = (\alpha, \alpha^{-1}\beta)$  und (28) gleich

$$\Pi((\alpha, \alpha^{-1}\beta)(\gamma, \gamma^{-1}\delta)) = \Pi(\alpha\gamma, \gamma^{-1}\alpha^{-1}\beta\delta) = (\alpha\gamma, \beta\delta).$$

Hieraus folgt (vgl. RÉDEI [8] § 2) die Richtigkeit von (29).

### Fall P:

Wir wollen hier die den vorigen analogen Vorbereitungen schaffen. Zu diesem Zwecke verfahren wir am bequemsten so, daß wir vorläufig die Definition (12) der Doppelhomothetismen außer acht lassen und statt deren von der folgenden Definition ausgehen, von der wir aber später unten zeigen werden, daß diese mit der vorigen äquivalent ist.<sup>25</sup>

Unter einem Doppelhomothetismus von P verstehen wir eine Doppelabbildung  $a$  von P in sich mit den Eigenschaften:

$$(30) \quad a(\alpha + \beta) = a\alpha + a\beta, \quad (\alpha + \beta)a = \alpha a + \beta a,$$

$$(31) \quad a\alpha\beta = (a\alpha)\beta, \quad \alpha\beta a = \alpha(\beta a),$$

$$(32) \quad (a\alpha)\beta = \alpha(a\beta),$$

$$(33) \quad (a\alpha)a = a(a\alpha).$$

<sup>25</sup> Diese „zweite“ Definition entspricht mehr der üblichen Definition der Automorphismen von  $\Gamma$  und ist mehr explizit, dagegen begrifflich weniger einfach als die „erste“ Definition (12).



Wegen (30) ist jeder Doppelhomothetismus von  $P$  ein Doppelendomorphismus von  $P^+$ . Entsprechend werden wir unter einem Ring von Doppelhomothetismen von  $P$  stets solche Unterringe des vollen Doppelendomorphismenringes von  $P^+$  verstehen, die aus Doppelhomothetismen von  $P$  bestehen. Nun bilden die sämtlichen Doppelhomothetismen von  $P$ , wie im § 1 schon erwähnt, im allgemeinen keinen Ring. Das werden wir erst im § 6 mit einem Beispiel zeigen (eilig ist es nicht). Dagegen zeigen wir hier zu einer ersten Orientierung folgendes:

Für zwei Doppelhomothetismen  $a, b$  von  $P$  ist  $a + b$  dann und nur dann ebenfalls ein Doppelhomothetismus von  $P$ , wenn die Bedingung

$$(34) \quad (aa)b + (ba)a = a(ab) + b(aa)$$

erfüllt ist.

Wir wissen nämlich, daß  $a + b$  jedenfalls ein Doppelendomorphismus von  $P^+$ , also (30) für  $a + b$  statt  $a$  erfüllt ist. Auch ist wegen (17) klar, daß (31), (32) mit  $a$  und  $b$  zusammen auch für  $a + b$  erfüllt sind. Damit also  $a + b$  ein Doppelhomothetismus von  $P$  ist, ist notwendig und hinreichend, daß (33) für  $a + b$  statt  $a$  erfüllt ist. Diese Bedingung lautet so:

$$((a+b)a)(a+b) = (a+b)(a(a+b)).$$

Dies läßt sich wegen (33) und der ähnlichen Gleichung für  $b$  statt  $a$ , ferner wegen (17) in (34) umformen, womit die Behauptung bewiesen ist.

Nun wird aber das eben gewonnene Kriterium (34) außer acht bleiben, denn wir werden es (glücklicherweise) auf einmal nur mit solchen Doppelhomothetismen  $a, b$  von  $P$  zu tun haben, für die die Bedingung

$$(35) \quad (aa)b = a(ab), \quad (ba)a = b(aa)$$

erfüllt ist; zwei Doppelhomothetismen  $a, b$  von  $P$  mit dieser Eigenschaft (35) nennen wir *befreundet*.<sup>26</sup> Ferner verstehen wir unter einer *Menge* oder einem *Ring von befreundeten Doppelhomothetismen* von  $P$  eine Menge bzw. einen Ring von paarweise befreundeten Doppelhomothetismen von  $P$ . Uns werden in der Hauptsache (aus der Menge aller Doppelhomothetismen von  $P$ ) nur die Ringe von befreundeten Doppelhomothetismen von  $P$  interessieren. Über diese einen Überblick zu verschaffen ist unser nächster Zweck.

Da (35) für  $a = b$  in (33) übergeht, so sehen wir vor allem, daß jeder Doppelhomothetismus von  $P$  mit sich selbst befreundet ist, d. h. für sich eine Menge von befreundeten Doppelhomothetismen von  $P$  bildet. Das bedeutet insbesondere die Existenz der Mengen von befreundeten Doppelhomothetismen von  $P$ .

<sup>26</sup> Die Bedeutung obiger Definition liegt darin, daß — wie wir später zeigen werden — zwei Doppelhomothetismen von  $P$  dann und nur dann befreundet sind, wenn sie durch zwei Elemente induziert werden, die in *eine* Schreiersche Erweiterung von  $P$  gehören. — Da (34) eine Folgerung aus (35) ist, so gilt nach obigem jedenfalls, daß die Summe  $a + b$  zweier befreundeter Doppelhomothetismen  $a, b$  von  $P$  auch ein Doppelhomothetismus ist. Dieses Resultat wird durch die späteren weit überholt.



Es ist wichtig, daß jede Menge von befreundeten Doppelhomothetismen von  $P$  in einer ebensolchen maximalen<sup>27</sup> Menge enthalten ist.

Ist nämlich  $M_1 \subset M_2 \subset \dots$  eine (unendliche) Kette von Mengen von befreundeten Doppelhomothetismen von  $P$ , so ist offenbar auch die Vereinigungsmenge von  $M_1, M_2, \dots$  eine Menge von befreundeten Doppelhomothetismen von  $P$ . Hieraus folgt nach dem Lemma von KURATOWSKI-ZORN die Richtigkeit der Behauptung.

Wir zeigen ferner, daß der durch eine Menge  $M$  von befreundeten Doppelhomothetismen von  $P$  erzeugte Ring  $\{M\}$  stets ein Ring von befreundeten Doppelhomothetismen von  $P$  ist.

Zunächst ist nämlich  $\{M\}$ , wie wir wissen, ein Unterring des vollen Doppelendomorphismenringes von  $P^+$ . Zum Beweis der übrigen Behauptung wollen wir eine Eigenschaft einer beliebigen Menge  $M'$  von Doppelendomorphismen von  $P^+$  *permanent* nennen, wenn diese Eigenschaft auch der Menge der Elemente des Ringes  $\{M'\}$  zukommt. Es folgt aus (17), (18) mit Induktion, daß die über alle Elemente bzw. alle Elementpaare  $a, b$  unserer Menge  $M$  vorausgesetzten Bedingungen (31), (32), (35)<sup>28</sup> lauter permanente Eigenschaften von  $M$  ausdrücken. Das beendet den Beweis unserer Behauptung.

Jede maximale Menge von befreundeten Doppelhomothetismen von  $P$  bildet einen Ring; diese Ringe nennen wir die *maximalen Ringe von befreundeten Doppelhomothetismen von  $P$* .

Ist nämlich  $M$  eine maximale Menge von befreundeten Doppelhomothetismen von  $P$ , so ist  $\{M\}$  nach vorigem ein Ring von befreundeten Doppelhomothetismen von  $P$ . Da dabei  $M \subseteq \{M\}$  gilt, so folgt aus der Maximalitätseigenschaft von  $M$  die Gleichung  $M = \{M\}$ , also die Richtigkeit der Behauptung.

Aus obigem folgt nunmehr auch, daß jeder Ring (und sogar überhaupt jede Menge) von befreundeten Doppelhomothetismen von  $P$  in mindestens einem ebensolchen maximalen Ring enthalten ist.

Betrachten wir nunmehr einen beliebigen (nicht notwendig maximalen) Ring  $D$  von befreundeten Doppelhomothetismen von  $P$ . Wir wiederholen, daß das nach (17), (18), (30), (31), (32), (35)<sup>28</sup> folgendes bedeutet:  $D$  ist ein Ring von Doppelabbildungen von  $P$  in sich mit den Eigenschaften

$$(36) \quad a(\alpha + \beta) = a\alpha + a\beta, \quad (\alpha + \beta)a = \alpha a + \beta a,$$

$$(37) \quad (\alpha + \beta)a = \alpha a + \beta a, \quad a(\alpha + \beta) = a\alpha + a\beta,$$

$$(38) \quad a\alpha\beta = (a\alpha)\beta, \quad \alpha\beta a = \alpha(\beta a),$$

$$(39) \quad a\beta\alpha = a(\beta\alpha), \quad \alpha a\beta = (\alpha a)\beta,$$

$$(40) \quad (\alpha a)\beta = \alpha(a\beta),$$

$$(41) \quad (a\alpha)b = a(\alpha b)$$

für alle  $a, b \in D$  und  $\alpha, \beta \in P$ .

<sup>27</sup> „Maximal“ wird stets im üblichen mengentheoretischen Sinne gebraucht.

<sup>28</sup> Man braucht (33) nicht zu nennen, da dies in (35) enthalten ist.



Für jeden dieser Ringe  $D$  definieren wir in der Menge aller Paare

$$(42) \quad (a, \alpha) \quad (a \in D, \alpha \in P)$$

die zwei Verknüpfungen (Summe und Produkt)

$$(43) \quad (a, \alpha) + (b, \beta) = (a + b, \alpha + \beta), \quad (a, \alpha)(b, \beta) = (ab, \alpha b + a\beta + \alpha\beta).$$

So entsteht nach RÉDEI [6] Satz 4 nebst Korollar<sup>29</sup> eine faktorenfreie (also zerfallende) Schreiersche Erweiterung von  $P$  mit  $D$ , die wir mit

$$(44) \quad D \bullet P$$

bezeichnen und die  $D$  zugehörige zerfallende Erweiterung von  $P$  nennen.<sup>30</sup> In diesem Ring bilden die Elemente  $(0, \alpha)$  ein mit  $P$  isomorphes Ideal, demgemäß wir die Einbettung  $(0, \alpha) \rightarrow \alpha$  stets ausgeführt denken, wodurch

$$(45) \quad P \triangleleft (D \bullet P)$$

in Erfüllung geht.<sup>31</sup>

Hier schalten wir den Beweis ein, daß die für die Doppelhomothetismen bei (12) bzw. (30) bis (33) angegebenen zwei Definitionen äquivalent sind. Diese Definitionen unterscheiden wir kurz als erste bzw. zweite Definition.

Zum Beweis bezeichne  $a$  zuerst einen Doppelhomothetismus von  $P$  im Sinne der ersten Definition. Das bedeutet nach (12), daß es einen Ring  $\bar{P}$  mit  $P \triangleleft \bar{P}$  und ein Element  $\bar{a}$  in ihm gibt, so daß

$$(46) \quad a\alpha = \bar{a}\alpha, \quad \alpha a = \alpha\bar{a}$$

gelten. (Die rechten Seiten bedeuten gewöhnliche Produkte im Ring  $\bar{P}$ .) Nun sind (30) bis (33) für  $\bar{a}$  statt  $a$  trivial erfüllt. Folglich ist  $a$  eine Doppelabbildung von  $P$  in sich, für die wegen (46) die Gleichungen (30) bis (33) ebenfalls erfüllt sind, somit ist  $a$  ein Doppelhomothetismus von  $P$  auch im Sinne der zweiten Definition.

Umgekehrt bezeichne jetzt  $a$  einen Doppelhomothetismus wie eben gesagt. Dann ist  $a$  in einem maximalen Ring  $D$  von befreundeten Doppelhomothetismen von  $P$  enthalten. (Übrigens würden wir auch mit dem Ring  $\{a\}$  statt  $D$  auskommen.) Wir nehmen den Ring (44) zu Hilfe, in dem wir jetzt aber die Einbettung  $(0, \alpha) \rightarrow \alpha$  vorläufig nicht ausführen. In diesem Ring bilden die Elemente  $(0, \alpha)$  ein Ideal  $P_1$ . Hiernach induziert das Element  $(a, 0)$  von (44) einen Doppelhomothetismus von  $P_1$  im Sinne der ersten Definition,

<sup>29</sup> Wir berichtigen einen augenscheinlichen Fehler im oben zitierten Korollar, und zwar sind dort die Bedingungen (56) bis (60) noch mit

$$ab\gamma = a(b\gamma), \quad abc = (ab)c$$

zu ergänzen.

<sup>30</sup> Mutatis mutandis gilt<sup>24</sup> auch hierfür.

<sup>31</sup> Schon aus obigem ersieht man die Bedeutung der Doppelhomothetismen insbesondere im Zusammenhang mit der Schreierschen Erweiterungstheorie. Im wesentlichen sind die Doppelhomothetismen (ohne diese Benennung) auch schon in RÉDEI [7] Satz 2 aufgetreten.



dessen zwei Bestandteile nach (12) die Abbildungen

$$(0, \varrho) \rightarrow (a, 0) \quad (0, \varrho) \rightarrow (0, \varrho) \quad (a, 0)$$

sind. Hierfür läßt sich nach (43<sub>2</sub>)

$$(0, \varrho) \rightarrow (0, a\varrho), \quad (0, \varrho) \rightarrow (0, \varrho a)$$

schreiben. Dies geht nach der Einbettung eben in (12) über, womit wir gezeigt haben, daß  $a$  ein Doppelhomothetismus von  $P$  auch im Sinne der ersten Definition ist. Damit haben wir die Äquivalenz beider Definitionen der Doppelhomothetismen ausgewiesen.

Es ist nach (36) bis (41) klar, daß die sämtlichen inneren Doppelhomothetismen von  $P$  einen Ring von befreundeten Doppelhomothetismen von  $P$  bilden, den wir kurz den *vollen inneren Doppelhomothetismenring* von  $P$  nennen.

Ist  $D$  insbesondere ein *innerer Doppelhomothetismenring* von  $P$ , worunter wir einen beliebigen Unterring des vorher genannten Ringes verstehen, so können wir dem Ring (44) eine mehr explizite Form geben. Zu diesem Zweck betrachten wir den durch  $\alpha$  induzierten (inneren) Doppelhomothetismus von  $P$ , der also aus den zwei Abbildungen

$$(47) \quad \varrho \rightarrow \alpha\varrho, \quad \varrho \rightarrow \varrho\alpha$$

besteht. Dieser hängt offenbar nur von der Klasse  $\alpha_* = \alpha + P_*$  ab, weshalb wir ihn dieser Klasse  $\alpha_*$  zuordnen können. Diese Zuordnung ist auch rückwärts eindeutig. Hiernach dürfen wir (47) den „Doppelhomothetismus  $\alpha_*$ “ nennen und diesen ebenfalls mit  $\alpha_*$  bezeichnen. Dann gilt

$$(48) \quad \alpha_*\varrho = \alpha\varrho, \quad \varrho\alpha_* = \varrho\alpha.$$

Dabei ist die gesagte Zuordnung ein Isomorphismus zwischen  $P/P_*$  und dem vollen inneren Doppelhomothetismenring von  $P$ , weshalb wir den letzteren nach Identifizierung der entsprechenden Elemente einfach mit  $P/P_*$  bezeichnen dürfen. Also läßt sich obiges  $D$  als ein beliebiger Unterring von  $P/P_*$  annehmen. Dann besteht  $D \bullet P$  aus den Elementen

$$(49) \quad (\alpha_*, \beta) \quad (\alpha_* \in D \subseteq (P/P_*), \beta \in P)$$

und als Verknüpfungsregeln gelten nach (43), (48)

$$(50) \quad (\alpha_*, \beta) + (\gamma_*, \delta) = (\alpha_* + \gamma_*, \beta + \delta), \quad (\alpha_*, \beta) (\gamma_*, \delta) = (\alpha_*\gamma_*, \alpha\delta + \beta\gamma + \beta\delta),$$

wobei auf der rechten Seite der zweiten Gleichung beliebige Repräsentanten  $\alpha, \gamma$  der Klasse  $\alpha_*$  bzw.  $\gamma_*$  zu nehmen sind.

Wenn dabei  $P$  *annullatorfrei* ist (d. h.  $P_* = 0$  gilt), so dürfen wir  $P/P_* = P$ ,  $\alpha_* = \alpha$  schreiben. Entsprechend lassen sich jetzt nach (49), (50) die Elemente des Ringes  $D \bullet P$  und die Verknüpfungsregeln in ihm durch

$$(51) \quad (\alpha, \beta) \quad (\alpha \in D \subseteq P, \beta \in P),$$

$$(52) \quad (\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta), \quad (\alpha, \beta) (\gamma, \delta) = (\alpha\gamma, \alpha\delta + \beta\gamma + \beta\delta)$$

angeben.



Gleich zeigen wir, daß  $D \bullet P$  in diesem Fall eine direkte Summe ist:<sup>32</sup>

$$(53) \quad D \bullet P \approx D \oplus P,$$

wobei „ $\oplus$ “ die direkte Summe bezeichnet. Zu diesem Zweck nehme man die Permutation

$$(\alpha, \beta) \rightarrow \Pi(\alpha, \beta) = (\alpha, \alpha + \beta)$$

der Elemente von  $D \bullet P$  zu Hilfe und gehe von (52<sub>2</sub>) auf die neue Multiplikation

$$(\alpha, \beta) \times (\gamma, \delta) = \Pi(\Pi^{-1}(\alpha, \beta)\Pi^{-1}(\gamma, \delta))$$

über. Die rechte Seite ist wegen  $\Pi^{-1}(\alpha, \beta) = (\alpha, -\alpha + \beta)$  und (52<sub>2</sub>) gleich

$$\Pi((\alpha, -\alpha + \beta)(\gamma, -\gamma + \delta)) = \Pi(\alpha\gamma, -\alpha\gamma + \beta\delta) = (\alpha\gamma, \beta\delta).$$

Da ferner (52<sub>1</sub>) gegenüber der ähnlichen „Transformation“ invariant bleibt, so folgt (vgl. RÉDEI [8] § 2) die Richtigkeit von (53).

### § 3. Die Grundlagen der Holomorphentheorie für Gruppen

Wir stellen hier die bekannten grundlegenden Definitionen und Sätze bezüglich der charakteristischen Untergruppen und des Holomorphes einer Gruppe zusammen.

DEFINITION 1<sub>0</sub>. Eine Untergruppe  $\Gamma'$  der Gruppe  $\Gamma$  nennen wir charakteristisch, wenn  $\Gamma'$  in allen Schreierschen Erweiterungen von  $\Gamma$  normal ist. (In Zeichen:  $\Gamma' \triangleleft \Gamma$  bedeutet, daß  $\Gamma' < \Gamma$  gilt und aus  $\Gamma \triangleleft \bar{\Gamma}$  stets  $\Gamma' \triangleleft \bar{\Gamma}$  folgt.)

SATZ 1<sub>0</sub>. (Erstes Kriterium für die charakteristischen Untergruppen.) Eine Untergruppe  $\Gamma'$  der Gruppe  $\Gamma$  ist dann und nur dann charakteristisch, wenn  $\Gamma'$  gegenüber allen Automorphismen von  $\Gamma$  zulässig<sup>33</sup> ist.

DEFINITION 2<sub>0</sub>. Unter dem Holomorph der Gruppe  $\Gamma$  verstehen wir die der vollen Automorphismengruppe  $A$  von  $\Gamma$  zugehörige zerfallende Erweiterung  $A \bullet \Gamma$  von  $\Gamma$ .<sup>34</sup>

<sup>32</sup> Darauf hat mich Herr J. SZENDREI aufmerksam gemacht und erst so wurde ich auch auf (29) aufmerksam.

<sup>33</sup> D. h.  $\Gamma^{a'} \subseteq \Gamma'$  für alle Automorphismen  $a$  von  $\Gamma$ . Da auch  $a^{-1}$  ein Automorphismus von  $\Gamma$  ist, so bedeutet diese Bedingung, daß  $\Gamma^{a^{-1}} = \Gamma'$  für alle  $a$  gilt, d. h.  $\Gamma'$  gegenüber allen Automorphismen von  $\Gamma$  (sogar) invariant ist. Man darf also wegen Satz 1<sub>0</sub> auch diese Eigenschaft als Definition zum Ausgang nehmen, wie das in der Literatur üblich ist, und dann gilt Definition 1<sub>0</sub> als Satz. Die von uns vorgenommene Vertauschung von Definition und Satz ist also durchaus gestattet und wird sich im § 4 lohnen.

<sup>34</sup> Man pflegt das Holomorph von  $\Gamma$  als eine Permutationsgruppe der Elemente von  $\Gamma$  zu definieren (vgl. ZASSENHAUS [16] S. 46), aber der Leser sieht leicht, daß beide Definitionen im wesentlichen übereinstimmen. Obige Definition ist begrifflich einfacher und scheint uns zu allen Zwecken geeigneter zu sein als die althergebrachte. Mit ihr sind wir in der Literatur nicht begegnet, aber nach einer brieflichen Mitteilung von Herrn L. FUCHS hat er sie auch entdeckt. — Erst nach der Abfassung unserer Arbeit wurden wir einer der obigen ähnlichen Definition des Holomorphes der Gruppe in der folgenden Arbeit gewahr: W. H. MILLS, On the nonisomorphism of certain holomorphs, *Transactions of the Amer. Math. Soc.*, 74 (1953), S. 428—443.



SATZ 2<sub>0</sub>. (Zweites Kriterium für die charakteristischen Untergruppen.)  
*Eine Untergruppe  $\Gamma'$  der Gruppe  $\Gamma$  ist dann und nur dann charakteristisch, wenn  $\Gamma'$  im Holomorph von  $\Gamma$  normal ist.*

Für den Beweis der Sätze 1<sub>0</sub>, 2<sub>0</sub> verweisen wir auf die Lehrbücher.

#### § 4. Die Grundlagen der Holomorphentheorie für Ringe

Ähnlich zur Definition 1<sub>0</sub> definieren wir, wie folgt:<sup>35</sup>

DEFINITION 1. *Einen Unterring  $P'$  des Ringes  $P$  nennen wir charakteristisch, wenn  $P'$  in allen Schreierschen Erweiterungen von  $P$  normal (d. h. ein Ideal) ist. (In Zeichen:  $P' \blacktriangleleft P$  bedeutet, daß  $P' < P$  gilt und aus  $P \triangleleft \bar{P}$  stets  $P' \triangleleft \bar{P}$  folgt.)*

BEMERKUNG. Hiernach sind die charakteristischen Unterringe lauter Ideale, wie auch die charakteristischen Untergruppen lauter Normalteiler sind. Diese Analogie war von einer trefflichen Begriffsbildung auch mit Recht zu erwarten, da — wie bemerkt — Normalteiler und Ideale durchaus analoge Begriffe sind. Hätten wir aber an die im Satz 1<sub>0</sub> enthaltene Eigenschaft der charakteristischen Untergruppen angeknüpft (mit der man diese — wie in<sup>33</sup> erwähnt — zu definieren pflegt), so wären wir zu denjenigen Unterringen von  $P$  gekommen, die gegenüber allen Automorphismen von  $P$  zulässig, also (vgl. den Anfang von<sup>33</sup>) invariant sind. Nun bilden zwar diese „invarianten Unterringe“ einen bekanntlich sehr wichtigen Begriff, doch ist dieser keineswegs als das Analogon zu den charakteristischen Untergruppen anzusehen, schon aus dem Grunde, daß die invarianten Unterringe keine Ideale zu sein brauchen. Das erklärt, warum wir uns bei obiger Definition 1 entschlossen haben. Aus diesen Auseinandersetzungen sieht man auch, daß die gruppen-ringtheoretischen Analogien nicht immer etwas selbstverständliches, sondern oft ziemlich verborgen sind.

SATZ 1. (Erstes Kriterium für die charakteristischen Unterringe.) *Ein Unterring  $P'$  des Ringes  $P$  ist dann und nur dann charakteristisch, wenn  $P'$  gegenüber allen Doppelhomothetismen von  $P$  zulässig<sup>36</sup> ist.*

BEMERKUNG. Durch die augenscheinliche Analogie der Sätze 1<sub>0</sub>, 1 wurde auch die Analogie (14) wiederholt bekräftigt.

Zum Beweis von Satz 1 betrachten wir einen charakteristischen Unterring  $P'$  und einen Doppelhomothetismus  $a$  von  $P$ . Letzterer wird wegen der Äquivalenz beider Definitionen der Doppelhomothetismen durch ein Element

<sup>35</sup> Analoge gruppen- und ringtheoretische Definitionen oder Sätze werden in der ganzen Arbeit mit 1<sub>0</sub>, 2<sub>0</sub>, ... bzw. 1, 2, ... numeriert.

<sup>36</sup> D. h.  $aP', P'a \subseteq P$  für alle Doppelhomothetismen  $a$  von  $P$ .



$\bar{a}$  einer Schreierschen Erweiterung  $\bar{P}$  von  $P$  induziert. Wegen der Definition 1 gilt  $P' \triangleleft \bar{P}$ , also ist

$$(54) \quad \bar{a}P', P'\bar{a} \subseteq P'.$$

Hierfür läßt sich

$$(55) \quad aP', P'a \subseteq P'$$

schreiben, womit wir „nur dann“ bewiesen haben.

Um auch „dann“ zu beweisen, nehmen wir (55) für ein  $P' (< P)$  und für alle Doppelhomothetismen  $a$  von  $P$  an. Wir betrachten ein  $\bar{P}$  mit  $P \triangleleft \bar{P}$ . Da jedes Element  $\bar{a}$  von  $\bar{P}$  einen Doppelhomothetismus von  $P$  induziert, so folgt aus (55) das Bestehen von (54). Dies bedeutet das Bestehen von  $P' \triangleleft \bar{P}$ . Somit haben wir Satz 1 bewiesen.

Ähnlich zur Definition 2<sub>0</sub> definieren wir, wie folgt:

DEFINITION 2. *Unter den Holomorphen des Ringes  $P$  verstehen wir die den maximalen Ringen  $D$  von befreundeten Doppelhomothetismen von  $P$  zugehörigen zerfallenden Erweiterungen  $D \bullet P$  von  $P$ .*

BEMERKUNG. Die Benennung „Holomorphe des Ringes“ ist auch durch den folgenden Satz 2 rechtfertigt, der das Analogon zum Satz 2<sub>0</sub> bildet. Die Bildung eines Holomorphes  $D \bullet P$  aus  $P$  und  $D$  geschieht nach (43) sehr einfach. Will man also alle Holomorphe von  $P$  kennen, so kommt das auf die Bestimmung der sämtlichen maximalen Ringen  $D$  von befreundeten Doppelhomothetismen von  $P$  an. Diese Frage für spezielle Ringe  $P$  zu untersuchen ist eine würdige und schwierige Aufgabe, mit der wir uns noch kaum beschäftigt haben. (Für den sehr einfachen Fall von „ $P$  mit Einselement“ s. Satz 3.) Für den allgemeinen Fall bemerken wir folgendes. Aus (35) folgt sofort, daß zwei Doppelhomothetismen von  $P$  befreundet sind, wenn mindestens der eine ein innerer oder der identische ist. Bezeichne  $D_0 (\approx P/P_*)$  den vollen inneren Doppelhomothetismenring von  $P$  und  $e$  (wie immer) den identischen Doppelhomothetismus von  $P$ . (Wir wissen, daß  $e$  nicht in  $D_0$  gehört, wenn  $P$  kein Einselement hat.) Nach dem Gesagten muß  $\{e, D_0\}$  in allen  $D$  enthalten sein, und sogar echt enthalten, wenn es außerhalb von  $\{e, D_0\}$  noch mindestens einen Doppelhomothetismus von  $P$  gibt. Ferner gilt offenbar: Dann und nur dann gibt es mehr als ein  $D$  (d. h. mehr als ein Holomorph von  $P$ ), wenn  $P$  zwei (äußere) Doppelhomothetismen hat, die nichtbefreundet sind. Wir bemerken auch: Aus  $e \in D$  und (43) folgt, daß jedes Holomorph von  $P$  das Einselement  $(e, 0)$  hat.

SATZ 2. (Zweites Kriterium für die charakteristischen Unterringe.) *Ein Unterring  $P'$  des Ringes  $P$  ist dann und nur dann charakteristisch, wenn  $P'$  in allen Holomorphen von  $P$  normal (d. h. ein Ideal) ist.*

„Nur dann“ folgt sofort aus den Definitionen 1, 2.



Um auch „dann“ zu beweisen, nehmen wir das Erfülltsein der im Satz 2 genannten Bedingung an und betrachten einen Doppelhomothetismus  $a$  von  $P$ . Nach § 2 ist  $a$  in einem maximalen Ring  $D$  von befreundeten Doppelhomothetismen von  $P$  enthalten. Nach der Annahme enthält das Holomorph  $D \bullet P$  den Unterring  $P'$  als Ideal. Andererseits induziert das Element  $(a, 0)$  von  $D \bullet P$ , wie wir gesehen haben, eben den Doppelhomothetismus  $a$ , woraus (55) folgt. Nach Satz 1 ist also  $P' \triangleleft P$ , womit Satz 2 bewiesen ist.

## § 5. Vollständige Gruppen. Ringe mit Einselement

Die im vorigen entwickelte Analogie zwischen Gruppen- und Ringtheorie läßt sich noch weiter ausbauen. So werden wir hier sehen, daß die Ringe mit Einselement sich zu den vollständigen Gruppen<sup>37</sup> analog verhalten.

**SATZ 3<sub>0</sub>.** *Eine Gruppe ist dann und nur dann vollständig, wenn sie ein direkter Faktor in allen ihren Schreierschen Erweiterungen ist. Wenn eine Gruppe ein direkter Faktor in ihrem Holomorph ist, so ist sie vollständig oder das direkte Produkt einer vollständigen Gruppe und einer Gruppe zweiter Ordnung.*

**BEMERKUNG.** Die Behauptung „nur dann“ ist bekannt (s. SPEISER [11] Satz 110, wo nur die endlichen Gruppen betrachtet werden). Die Behauptung „dann“ ist im wesentlichen ein Satz von BAER [1]. Die zweite Hälfte von Satz 3<sub>0</sub> ist unseres Wissens neu. Wir werden Satz 3<sub>0</sub> sehr leicht im ganzen Umfange beweisen.

**SATZ 3.** *Ein Ring hat dann und nur dann ein Einselement, wenn er ein direkter Summand in allen seinen Schreierschen Erweiterungen ist. Die Ringe mit Einselement lassen sich auch dadurch charakterisieren, daß sie nur innere Doppelhomothetismen haben.<sup>38</sup> Folglich sind ihre sämtlichen Ideale charakteristisch, ferner haben sie nur ein Holomorph.<sup>39</sup> Umgekehrt wenn ein Ring ein direkter Summand in seinem irgendwelchen Holomorph ist, so ist er ein Ring mit Einselement (folglich hat er nur ein Holomorph).<sup>40</sup>*

<sup>37</sup> Bekanntlich nennt man eine Gruppe vollständig, wenn sie nur innere Automorphismen hat und ohne Zentrum ist (vgl. SPEISER [11], S. 125). Eine vollständige Gruppe ist also isomorph mit ihrer vollen Automorphismengruppe.

<sup>38</sup> Da ein Ring mit Einselement auch annullatorfrei ist, so induzieren seine Elemente lauter verschiedene Doppelhomothetismen, folglich ist er isomorph mit seinem „vollen Doppelhomothetismenring“ (vgl. den Schluß von <sup>37</sup>). — Im allgemeinen, wenn die sämtlichen Doppelhomothetismen von  $P$  einen Ring bilden, so können wir diesen den vollen Doppelhomothetismenring von  $P$  nennen. Dieser kann natürlich auch dann existieren, wenn  $P$  kein Einselement hat.

<sup>39</sup> Das Holomorph von einem Ring  $P$  mit Einselement läßt sich einfach als  $P \bullet P$  angeben (s. Fall  $D = P$  von (51), (52)).

<sup>40</sup> Ausführlicher gesprochen: Hat ein Ring kein Einselement, so ist er in keinem Holomorph von ihm ein direkter Summand. Hat er das Einselement, so hat er nur ein Holomorph und ist ein direkter Summand in diesem.



Die Behauptung „dann“ ist eine Folgerung aus der letzten Behauptung des Satzes. Die Behauptung „nur dann“ war in einem Spezialfall schon TSCHEBOTAREFF [15] bekannt.<sup>40a</sup> Nach dem Anfang der Sätze 3<sub>0</sub>, 3 sind die vollständigen Gruppen und die Ringe mit Einselement als genaue Analoga voneinander anzusehen. Es ist merkwürdig, daß Satz 3 im ganzen viel reichhaltiger ist als Satz 3<sub>0</sub>.<sup>41</sup> Selbst das im ersten Teil von Satz 3 gelöste Problem der Bestimmung der ringtheoretischen Analoga der vollständigen Gruppen hat schon früher SZENDREI [14] aufgeworfen und beantwortet.

BEWEIS VON SATZ 3<sub>0</sub>. Wir fangen es mit „nur dann“ an. Zu diesem Zweck betrachten wir eine vollständige Gruppe  $\Gamma$  und eine Schreiersche Erweiterung  $\bar{\Gamma}$  von ihr. Wir haben zu zeigen, daß  $\Gamma$  als direkter Faktor in  $\bar{\Gamma}$  enthalten ist. Wegen der Annahme ist für jedes Element  $\bar{\alpha}$  von  $\bar{\Gamma}$  die Abbildung

$$\varrho \rightarrow \bar{\alpha}^{-1} \varrho \bar{\alpha} \quad (\varrho \in \Gamma)$$

ein innerer Automorphismus von  $\Gamma$ . Da ferner  $\Gamma$  ohne Zentrum ist, so gibt es zu  $\bar{\alpha}$  ein einziges Element  $\bar{\alpha}'$  von  $\bar{\Gamma}$ , das diesen Automorphismus induziert. Dann ist  $\bar{\alpha}^{-1} \bar{\alpha}'$  mit allen Elementen von  $\Gamma$  vertauschbar, folglich gibt es in jeder Klasse von  $\bar{\Gamma}$  nach  $\Gamma$  ein einziges, mit allen Elementen von  $\Gamma$  vertauschbares Element. Diese Elemente bilden eine Untergruppe  $\Gamma_0$  von  $\bar{\Gamma}$ , ferner ist  $\Gamma_0$  ein Repräsentantensystem der Klassen nach  $\Gamma$ . Hiernach ist  $\bar{\Gamma}$  das direkte Produkt von  $\bar{\Gamma}$  und  $\Gamma_0$ , womit wir „nur dann“ bewiesen haben.

Jetzt beweisen wir die letzte Behauptung von Satz 3<sub>0</sub>. Hierzu nehmen wir an, daß eine Gruppe  $\Gamma$  ein direkter Faktor in ihrem Holomorph  $A \bullet \Gamma$  ist, wobei  $A$  die volle Automorphismengruppe von  $\Gamma$  bezeichnet. Nach (20) und RÉDEI [6] Satz 3 gibt es wegen der Annahme eine Abbildung

$$a \rightarrow a'$$

von  $A$  in  $\Gamma$ , wofür

$$(56) \quad (ab)^{-1} a' b' = \varepsilon, \quad a'^{-1} \varrho a' = \varrho^a$$

gilt. Aus (56<sub>2</sub>) folgt, daß  $A$  aus lauter inneren Automorphismen von  $\Gamma$  besteht. Dann läßt sich  $A = \Gamma / \Gamma_*$  setzen, ferner folgt aus (56<sub>1</sub>)

$$(57) \quad (\alpha_* \beta_*)' = \alpha'_* \beta'_*, \quad \alpha_*^{-1} \varrho \alpha_* = \varrho^{\alpha_*}$$

<sup>40a</sup> Erst nach der Abfassung dieser Arbeit bemerkten wir, daß die Behauptung „nur dann“ von Satz 3 in voller Allgemeinheit auch schon in der folgenden Arbeit vorkommt: B. BROWN and N. H. MCCOY, The maximal regular ideal of a ring, *Proceedings of the Amer. Math. Soc.*, 1 (1950), S. 165–171 (Lemma 3).

<sup>41</sup> Auch sonst benehmen sich die Ringe mit Einselement in mancher Hinsicht abweichend von den vollständigen Gruppen. Es ist z. B. trivial, daß jeder Ring eine Schreiersche Erweiterung mit Einselement hat, nach dem Schluß der Bemerkung nach Definition 2 gilt sogar, daß jedes Holomorph eines Ringes ein Ring mit Einselement ist, demgegenüber ist unseres Wissens unbekannt, ob jede Gruppe sich nach SCHREIER zu einer vollständigen Gruppe erweitern läßt.



wobei

$$\alpha_* \rightarrow \alpha'_*$$

eine Abbildung von  $\Gamma/\Gamma_*$  in  $\Gamma$  ist. Aus (57<sub>2</sub>) und  $\varrho^{\alpha_*} = \alpha^{-1}\varrho\alpha$  folgt, daß  $\alpha'_*$  in die Klasse  $\alpha_* = \alpha\Gamma_*$  gehört. Da ferner die  $\alpha'_*$  wegen (57<sub>1</sub>) eine Untergruppe  $\Gamma'$  von  $\Gamma$  bilden, so zerfällt  $\Gamma$  in das Produkt  $\Gamma'\Gamma_*$ , worunter wir verstehen, daß sich jedes Element von  $\Gamma$  eindeutig als ein Produkt  $\varrho\sigma$  ( $\varrho \in \Gamma', \sigma \in \Gamma_*$ ) schreiben läßt. Da aber  $\Gamma_*$  das Zentrum von  $\Gamma$  ist, so folgt hieraus

$$\Gamma = \Gamma' \otimes \Gamma_*$$

ferner muß  $\Gamma'$  zentrumlos sein. Hat ein direktes Produkt nur innere Automorphismen, wie das nach obigem für  $\Gamma$  der Fall ist, so gilt ähnliches auch für die Faktoren. Hieraus folgt, daß  $\Gamma'$  vollständig und  $\Gamma_*$  (als Abelsche Gruppe) aus höchstens zwei Elementen besteht. Damit haben wir die letzte Behauptung von Satz 3<sub>0</sub> bewiesen.

Aus Satz 3<sub>0</sub> haben wir nur noch „dann“ zu beweisen. Da das Holomorph eine Schreiersche Erweiterung ist, so genügt es folgendes zu beweisen: Ist eine Gruppe

$$(58) \quad \Gamma = A \otimes B$$

das direkte Produkt von zwei Untergruppen  $A, B$ , von denen  $A$  ( $\neq \varepsilon$ ) Abelsch und endlich ist, so hat  $\Gamma$  eine Schreiersche Erweiterung  $\bar{\Gamma}$ , in der  $\Gamma$  kein direkter Faktor ist. Wir nehmen eine echte Obergruppe  $\bar{A}$  von  $A$ , die keinen mit  $A$  isomorphen direkten Faktor enthält, und setzen

$$(59) \quad \bar{\Gamma} = \bar{A} \otimes B.$$

Dabei läßt sich  $\Gamma < \bar{\Gamma}$  annehmen, und dann gilt sogar  $\Gamma \triangleleft \bar{\Gamma}$ . Es genügt zu zeigen, daß  $\Gamma$  kein direkter Faktor in  $\bar{\Gamma}$  ist. Wir nehmen an, daß  $\Gamma$  ein direkter Faktor in  $\bar{\Gamma}$  ist. Dann gilt wegen (58)

$$(60) \quad \bar{\Gamma} = A_1 \otimes A \otimes B$$

mit einer Untergruppe  $A_1$  von  $\bar{A}$ . Aus (59), (60) folgt die Isomorphie<sup>42</sup>

$$\bar{A} \approx A_1 \otimes A.$$

Dieser Widerspruch beendet den Beweis von Satz 3<sub>0</sub>.

BEWEIS VON SATZ 3. Wir fangen es mit der zweiten Behauptung an. Hierzu betrachten wir einen Ring  $P$  und nehmen zuerst an, daß  $P$  das Element  $\varepsilon$  hat. Dann folgt aus (31) für jeden Doppelhomothetismus  $a$  von  $P$ :

$$a\varrho = a\varepsilon\varrho = (a\varepsilon)\varrho, \quad \varrho a = \varrho\varepsilon a = \varrho(\varepsilon a).$$

Andererseits gilt nach (32)

$$(\varepsilon a)\varepsilon = \varepsilon(a\varepsilon),$$

<sup>42</sup> Gilt nämlich  $G = A \otimes B = A' \otimes B$ , wobei  $A, A', B$  Untergruppen der Gruppe  $G$  sind, so gilt  $A \approx A'$ .



d. h.  $a\varepsilon = \varepsilon a$ . Hiernach ist  $a$  gleich dem durch  $a\varepsilon (= \varepsilon a)$  induzierten inneren Doppelhomothetismus von  $P$ .

Umgekehrt wenn  $P$  nur innere Doppelhomothetismen hat, so ist auch der identische ein solcher, woraus die Existenz eines Elementes  $\alpha$  mit  $\alpha\varrho = \varrho, \varrho\alpha = \varrho$  folgt. Dann ist  $\alpha$  eben das Einselement von  $P$ . Somit haben wir die zweite Behauptung von Satz 3 bewiesen.

Die in der Mitte von Satz 3 erwähnten Folgerungen sind auch richtig, teils wegen Satz 1, teils weil die inneren Doppelhomothetismen paarweise befreundet sind.

Um die Behauptung „nur dann“ von Satz 3 zu beweisen, nehmen wir im Ring  $P$  die Existenz des Einselementes  $\varepsilon$  an und betrachten eine Schreierische Erweiterung  $\bar{P}$  von  $P$ . Wir haben zu zeigen, daß  $P$  ein direkter Summand in  $\bar{P}$  ist. Für jedes Element  $\bar{\alpha}$  von  $\bar{P}$  bilden die zwei Abbildungen

$$\varrho \rightarrow \bar{\alpha}\varrho, \quad \varrho \rightarrow \varrho\bar{\alpha}$$

einen Doppelhomothetismus von  $P$ . Folglich gibt es wegen der Annahme und des schon Bewiesenen ein  $\alpha (\in P)$  mit

$$(61) \quad \bar{\alpha}\varrho = \alpha\varrho, \quad \varrho\bar{\alpha} = \varrho\alpha.$$

Wegen der Annahme gilt auch  $P_* = 0$ , folglich ist  $\alpha$  eindeutig durch  $\bar{\alpha}$  bestimmt. Schreibt man (61) in der form

$$(\bar{\alpha} - \alpha)\varrho = \varrho(\bar{\alpha} - \alpha) = 0,$$

so sieht man, daß

$$\bar{\alpha} = \alpha + \bar{\nu}$$

gilt, wobei  $\bar{\nu} (\in \bar{P})$  die Eigenschaft

$$(62) \quad \bar{\nu}P = P\bar{\nu} = 0$$

hat. Die sämtlichen  $\bar{\nu}$  mit der Eigenschaft (62) bilden ein Ideal  $\mathfrak{n}$  von  $\bar{P}$  (den „Annulator von  $P$  in  $\bar{P}$ “), das wegen der Annahme kein von 0 verschiedenes gemeinsames Element mit  $P$  hat. Wir haben gewonnen, daß  $\bar{P}$  die direkte Summe seiner Ideale  $P, \mathfrak{n}$  ist, womit wir die Behauptung „nur dann“ von Satz 3 bewiesen haben.

Um die letzte Behauptung von Satz 3 zu beweisen, betrachten wir einen Ring  $P$  und ein Holomorph  $D \bullet P$  von ihm, wobei  $D$  ein maximaler Ring von befreundeten Doppelhomothetismen von  $P$  ist. Wir nehmen an, daß  $P$  als direkter Summand in  $D \bullet P$  enthalten ist, und dann haben wir zu beweisen, daß  $P$  das Einselement hat. Wegen der Annahme gibt es nach (43) und RÉDEI [6] Satz 6 eine Abbildung

$$a \rightarrow a'$$

von  $D$  in  $P$ , wofür<sup>43</sup>

$$(63) \quad \left. \begin{aligned} a' + b' - (a + b)' &= 0, & a'b + ab' + a'b' - (ab)' &= 0, \\ \alpha b + \alpha b' &= 0, & \alpha\beta + \alpha'\beta &= 0, \end{aligned} \right\} (a, b \in D; \alpha, \beta \in P)$$

<sup>43</sup> Wir berichtigen den Druckfehler im zitierten Satz, wo nämlich in (74) für das erste  $(ab)'$  richtig  $(a + b)'$  stehen soll.



gelten. Aus (63<sub>3, 4</sub>) folgt, daß  $D$  aus lauter inneren Doppelhomothetismen von  $P$  besteht. Da diese, wie erwähnt, mit jedem Doppelhomothetismus von  $P$  befreundet sind, so folgt aus der Maximaleigenschaft von  $D$ , daß  $P$  überhaupt nur innere Doppelhomothetismen haben kann. Nach dem oben bewiesenen Teil von Satz 3 bedeutet dies die Existenz des Einselementes in  $P$ , womit wir die letzte Behauptung von Satz 3 bewiesen haben.

Da jedes Holomorph eine Schreiersche Erweiterung ist, so folgt hieraus auch schon die Behauptung „dann“ von Satz 3, womit der Beweis dieses Satzes beendet ist.

### § 6. Weitere Anwendungen. Beispiele

Nach Satz 3 sind alle Ideale eines Ringes mit Einselement charakteristisch. Ein Gegenstück hierzu bildet ein neulich gefundener Satz von NAGATA [5],<sup>44</sup> der so lautet: Ist  $\mathfrak{p}$  ein Primideal<sup>45</sup> in  $P$  und  $P \triangleleft \bar{P}$ , so gilt  $\mathfrak{p} \triangleleft \bar{P}$ . Dieser Satz läßt sich nunmehr auch so aussprechen:

SATZ 4. *Die Primideale sind charakteristisch.*

Wir führen auf Grund von Satz 1 einen kurzen Beweis aus. Bezeichne  $\mathfrak{p}$  ein Primideal von  $P$  und  $a$  einen Doppelhomothetismus von  $P$ . Im Fall  $\mathfrak{p} = P$  ist der Satz richtig, weshalb wir nur noch den Fall  $\mathfrak{p} \neq P$  zu betrachten brauchen. Wegen (32), (31<sub>1</sub>) gilt  $P(a\mathfrak{p})$ ,  $(a\mathfrak{p})P \subseteq \mathfrak{p}$ , folglich gelten

$$P(\mathfrak{p} + a\mathfrak{p}) \subseteq \mathfrak{p}, \quad \mathfrak{p} + a\mathfrak{p} \triangleleft P.$$

Aus diesen und der Annahme ergibt sich

$$\mathfrak{p} + a\mathfrak{p} \subseteq \mathfrak{p},$$

also  $a\mathfrak{p} \subseteq \mathfrak{p}$ . Ebenso folgt  $\mathfrak{p}a \subseteq \mathfrak{p}$ . Beide besagen nach Satz 1 die Richtigkeit von Satz 4.

Merkwürdig an Satz 4 ist, daß er (nach unserer Ansicht) überhaupt kein gruppentheoretisches Analogon hat. Nach beiden Sätzen 3, 4 scheinen die charakteristischen Unterringe eine häufigere Erscheinung zu sein als die charakteristischen Untergruppen.

Gibt es überhaupt Ideale, die nicht charakteristisch sind? Diese Frage bejahen wir mit einem Beispiel. Bezeichne  $p$  eine Primzahl,  $I$  den Ring der ganzen Zahlen. Die Polynome

$$p^2 f(x) + p x g(x) \qquad (f(x), g(x) \in I[x])$$

<sup>44</sup> S. auch STEINFELD [12]. Auch diese Arbeit von STEINFELD, die ich schon im Manuskript kannte, hat dem Entstehen meiner „Holomorphentheorie für Ringe“ beigetragen.

<sup>45</sup> Wie das neuerdings üblich ist, verstehen wir unter einem Primideal eines beliebigen Ringes  $P$  ein solches Ideal  $\mathfrak{p}$  von  $P$ , wofür aus  $ab \in \mathfrak{p}$  ( $a, b \in P$ ) stets  $a \in \mathfrak{p}$  oder  $b \in \mathfrak{p}$  folgt. Die „Primideale im alten Sinn“ nennt man heute (vollständig oder) komplett; diese werden bekanntlich durch die Eigenschaft definiert, daß aus  $a\beta \in \mathfrak{p}$  stets  $a \in \mathfrak{p}$  oder  $\beta \in \mathfrak{p}$  folgt. Die kompletten Primideale sind ein Spezialfall der Primideale und für kommutative Ringe fallen beide Begriffe zusammen.



und

$$p^2 h(x) + p x k(x^2) \quad (h(x), k(x) \in I[x])$$

bilden je einen Unterring  $P$  bzw.  $P'$  von  $I[x]$ . Für diese gilt  $P' \triangleleft P$ ,  $P \triangleleft I[x]$  offenbar, dagegen gilt  $P' \triangleleft I[x]$  nicht mehr. Hiernach ist  $P'$  in der Tat ein Ideal in  $P$ , das kein charakteristischer Unterring ist.

Wir beweisen den folgenden merkwürdigen Satz, der ebenfalls kein gruppentheoretisches Analogon zu haben scheint:

**SATZ 5.** *Alle Holomorphe eines kommutativen nullteilerfreien Ringes sind kommutativ.* (Vgl. Satz 6.)

Bezeichne nämlich  $P$  einen solchen Ring. Für jeden Doppelhomothetismus  $\alpha$  von  $P$  gilt nach (32)  $(\alpha\alpha)\alpha = \alpha(\alpha\alpha)$ , woraus nach der Annahme zuerst  $\alpha(\alpha\alpha) = \alpha(\alpha\alpha)$ , dann

$$(64) \quad \alpha\alpha = \alpha\alpha$$

folgt. (Hiernach besteht jetzt jeder Doppelhomothetismus von  $P$  aus zwei gleichen Endomorphismen von  $P^+$ .) Sind nun  $a, b$  zwei befreundete Doppelhomothetismen von  $P$ , so folgt aus (18), (64), (35)

$$ab\alpha = a(b\alpha) = a(\alpha b) = (\alpha a)b = b(\alpha a) = ba\alpha,$$

d. h.

$$(65) \quad ab = ba.$$

Da nach (43<sub>2</sub>) die Elemente eines Holomorphes von  $P$  nach der Regel

$$(a, \alpha)(b, \beta) = (ab, \alpha b + a\beta + \alpha\beta)$$

multipliziert werden, woraus auch

$$(b, \beta)(a, \alpha) = (ba, \beta a + b\alpha + \beta\alpha)$$

folgt, so zeigen (64), (65) die Richtigkeit von Satz 5.

Beispiele für Ringe mit mehr als einem Holomorph bilden die Zeroringe.<sup>46</sup> Es gilt nämlich der folgende:

**SATZ 6.** *Ein Zeroring  $P$  hat dann und nur dann nur ein Holomorph, wenn der volle Endomorphismenring von  $P^+$  kommutativ ist.<sup>47</sup> Hat er mehrere Holomorphe, so gibt es darunter auch nichtkommutative.*

Wir schicken folgendes voran. Wegen der Annahme gilt  $\alpha\beta = 0$  unbeschränkt, also sind (31), (32) identisch erfüllt. Folglich stimmen jetzt nach (30), (33) die Doppelhomothetismen von  $P$  mit denjenigen Doppelendomorphismen von  $P^+$  überein, die aus zwei vertauschbaren Endomorphismen von  $P^+$  bestehen.

Ist nunmehr der volle Endomorphismenring von  $P^+$  kommutativ, so ist die Bedingung (35) auch erfüllt. Das bedeutet, daß jetzt alle Doppelhomothetismen

<sup>46</sup> Zeroring bedeutet einen Ring  $P$  mit  $P^2 = 0$  (d. h.  $P_* = P$ ).

<sup>47</sup> Das interessante Problem, wann der volle Endomorphismenring eines Moduls kommutativ ist, haben SZELE u. SZENDREI [13] unlängst in Angriff genommen und teilweise gelöst. Nach ihren Resultaten gibt es verhältnismäßig sehr wenig Moduln dieser Art.



tismen von  $P$  paarweise befreundet sind, d. h. einen Ring bilden, der dann der einzige maximale Doppelhomothetismenring von  $P$  ist, folglich hat  $P$  nur ein Holomorph.

Ist dagegen der volle Endomorphismenring von  $P^+$  nichtkommutativ, so betrachten wir zwei nichtvertauschbare Endomorphismen  $A, B$  von  $P^+$ . Bezeichne  $a, b$  (nach obiger Bemerkung) die zwei Doppelhomothetismen von  $P$ , so daß beide Bestandteile von  $a$  gleich  $A$  und die von  $b$  gleich  $B$  sind. Diese  $a, b$  sind nach (35) nichtbefreundet, gehören somit in zwei verschiedene maximale Ringe von befreundeten Doppelhomothetismen von  $P$ . Da durch sie zwei verschiedene Holomorphe von  $P$  bestimmt sind, so haben wir die erste Hälfte von Satz 6 bewiesen.

Wird zu den vorigen  $A, B$  je ein trivialer Endomorphismus  $0$  von  $P^+$  hinzugenommen, so entstehen offenbar zwei befreundete Doppelhomothetismen  $a = (A, 0)$ ,  $b = (B, 0)$  von  $P$ , für die  $ab \neq ba$  gilt. Diese sind also in einem nichtkommutativen maximalen Ring von befreundeten Doppelhomothetismen von  $P$  enthalten. Durch diesen ist ein nichtkommutatives Holomorph von  $P$  bestimmt, womit wir den Beweis von Satz 6 beendet haben.

Wir bemerken, daß selbstverständlich sich der Begriff der charakteristischen Reihen auf die Ringe übertragen läßt. Und zwar verstehe man unter einer *charakteristischen Reihe* von  $P$  jede nicht verfeinbare endliche Folge

$$P = P_0, P_1, \dots, P_r = 0,$$

in der  $P_{i+1}$  ein echter Unterring von  $P_i$  ( $i = 0, \dots, r-1$ ) ist und  $P_i \triangleleft P$  ( $i = 0, \dots, r$ ) gilt. Im wesentlichen handelt es sich also um die Kompositionsreihen der (additiven) Gruppe  $P^+$ , wenn man dieser die sämtlichen Doppelhomothetismen von  $P$  als Operatoren zuordnet. Entsprechend lassen sich die bekannten Sätze von SCHREIER und JORDAN—HÖLDER anwenden.

Als Ergänzung zu den im § 1 erwähnten Analogien schreiben wir noch die folgenden (sehr starken) Analogien hin, auf die wir während unserer Betrachtungen gestoßen sind:

charakteristische Untergruppe von  $\Gamma$  . . . charakteristischer Unterring von  $P$ ,  
 Holomorph von  $\Gamma$  . . . Holomorphe von  $P$ ,  
 vollständige Gruppe . . . Ring mit Einselement,  
 charakteristische Reihe von  $\Gamma$  . . . charakteristische Reihe von  $P$ .

(Eingegangen am 3. Februar 1954.)



## Literaturverzeichnis

- [1] R. BAER, Absolute retracts in group theory, *Bull. of the Amer. Math. Soc.*, **52** (1946), S. 501—506.
- [2] N. BOURBAKI, *Éléments de Mathématique*, première partie, Livre II: *Algèbre*, Chap. I (Paris, 1942), S. 1—165.
- [3] C. I. EVERETT, An extension theory for rings, *American Journal of Math.*, **64** (1942), S. 363—370.
- [4] L. FUCHS, A remark on the Jacobson radical, *Acta Sci. Math. Szeged*, **14** (1952), S. 167—168.
- [5] M. NAGATA, On the theory of radicals in a ring, *Journal of the Math. Soc. of Japan*, **3** (1951), S. 330—344.
- [6] L. RÉDEI, Die Verallgemeinerung der Schreierschen Erweiterungstheorie, *Acta Sci. Math. Szeged*, **14** (1952), S. 252—273.
- [7] L. RÉDEI, Über gewisse Ringkonstruktionen durch schiefes Produkt, *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 185—189.
- [8] L. RÉDEI, Die Anwendung des schiefen Produktes in der Gruppentheorie, *Journal f. d. reine u. angew. Math.*, **188** (1950), S. 201—227.
- [9] L. RÉDEI, Die Vollidealringe, *Monatshefte f. Math.*, **56** (1952), S. 89—95.
- [10] L. RÉDEI, Vollidealringe im weiteren Sinn, I, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 243—268.
- [11] A. SPEISER, *Die Theorie der Gruppen*, dritte Aufl. (Berlin, 1937).
- [12] O. STEINFELD, Über Idealquotienten und Primideale, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 289—298.
- [13] T. SZELE and J. SZENDREI, On abelian groups with commutative endomorphism ring, *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 309—323.
- [14] J. SZENDREI, On rings admitting only direct extensions, *Publicationes Math. (Debrecen)*, **3** (1954), S. 180—182.
- [15] Н. Г. Чебо тар ё в, Введение в теорию алгебр (Москва—Ленинград, 1949).
- [16] H. ZASSENHAUS, *Lehrbuch der Gruppentheorie* (Leipzig und Berlin, 1937).



## ТЕОРИЯ ГОЛОМОРФОВ ГРУПП И КОЛЕЦ

Л. РЭДЭИ (Сегед)

(Резюме)

Целью настоящей работы является перенесение теории голоморфов групп (под которой автор разумеет часть теории групп, изучающую характеристические подгруппы и голоморф группы) в теорию колец. Теория голоморфов колец может быть построено в тесной аналогии с теорией голоморфов групп. Автор определяет характеристические подкольца как подкольца, являющиеся идеалами во всяком шрейеровом (эвереттовом) расширении кольца. В противоположность теории групп кольцо вообще может иметь несколько голоморфов. Голоморфы являются некоторыми специальными шрейеровыми расширениями данного кольца. Эти шрейеровы расширения безфакторны, и сами факторкольца, при помощи которых расширение происходит, суть кольца, составленные из некоторых пар отображений данного кольца. Эти пары отображений называются автором двойными гомотетизмами. Определяются они так: Возьмем любой элемент из любого шрейерового расширения данного кольца, и перемножим этим (фиксированным) элементом справа и слева элементы данного кольца. Таким образом получаем два отображения кольца в себя (являющиеся, конечно, специальными эндоморфизмами аддитивной группы кольца). Такие пары отображений и суть двойные гомотетизмы. Двойные гомотетизмы можно складывать и умножать по простым правилам, однако множество всех двойных гомотетизмов данного кольца вообще говоря не образует кольцо. Тем не менее из множества всех двойных гомотетизмов можно выбирать некоторые „максимальные“ подмножества, образующие кольцо. Именно эти „максимальные кольца двойных гомотетизмов“ и играют при вышеупомянутых шрейеровых расширениях роль факторколец. Интересная особенность теории голоморфов колец заключается в том, что простые идеалы (в частности вполне простые идеалы) так же как и все идеалы колец с единицей характеристичны. Вообще оказывается, что кольца с единицей играют в теории колец роль точных аналогов совершенных групп. Так, например, справедлива следующая теорема: Кольца с единицей и только они суть кольца, допускающие только прямые шрейеровы расширения. Эта теорема аналогична известной теоретико групповой теореме о том, что совершенные группы и только они содержатся в качестве прямого множителя во всяком их шрейеровом расширении. На основе этой аналогии можно сказать, что кольца с единицей играют по существу роль „совершенных колец“.







# EINE NEUE DEFINITION DES HOLOMORPHES DER GRUPPE UND DER HOLOMORPHE DES RINGES

Von

J. SZENDREI (Szeged)  
(Vorgelegt von L. RÉDEI)

1. Die Untersuchungen der analogen Probleme in der Gruppen- und Ringtheorie haben vielfach dazu beigetragen, daß die gemeinsamen Züge der beiden Theorien aufgeklärt wurden, anderenteils haben diese Untersuchungen auf die wesentliche Verschiedenheit des Gruppen- und Ringbegriffes hingewiesen. Neulich hat Prof. L. RÉDEI [4]<sup>1</sup> die ringtheoretischen Analoga der charakteristischen Untergruppen und des mit diesen eng zusammenhängenden Holomorphes einer Gruppe eingeführt und damit die Holomorphentheorie der Ringe begründet. Der wesentliche Unterschied zwischen beiden Holomorphentheorien liegt darin, daß ein Ring — im Gegensatz zur Unizität des Holomorphes einer Gruppe — im allgemeinen mehrere Holomorphe besitzt.<sup>2</sup>

In dieser Arbeit wird bewiesen, daß man je eine begrifflich sehr einfache Definition des Holomorphes einer Gruppe und der Holomorphen eines Ringes geben kann, die die Analogie dieser Begriffe zugleich von einer anderen Seite beleuchtet. Diese neuen Definitionen sind weniger explizit, aber wir werden ihre Äquivalenz mit den bekannten expliziten Definitionen leicht ausweisen.

2. Es ist bekannt, daß sich das Holomorph einer Gruppe unter anderem auch so definieren läßt (W. H. MILLS [1], L. RÉDEI [4]):

DEFINITION  $A_0$ . *Unter dem Holomorph der Gruppe  $\Gamma$  verstehen wir die der vollen Automorphismengruppe  $A^*$  von  $\Gamma$  zugehörige zerfallende Schreiersche Erweiterung  $A^* \circ \Gamma$  von  $\Gamma$ .*<sup>3</sup>

Die durch L. RÉDEI [4] eingeführte Definition der Holomorphen eines Ringes lautet so:

<sup>1</sup> Mit [ ] verweisen wir auf die Arbeiten am Ende dieses Artikels.

<sup>2</sup> Die hier verwendeten Begriffe und weitere Analoga sind in der Arbeit von L. RÉDEI [4] zu finden.

<sup>3</sup> Siehe RÉDEI [4].



DEFINITION  $B_0$ . Unter den Holomorphen des Ringes  $P$  verstehen wir die den maximalen Ringen  $D^*$  von befreundeten Doppelhomothetismen von  $P$  zugehörigen zerfallenden Schreierschen Erweiterungen  $D^* \circ P$  von  $P$ .<sup>3</sup>

3.  $A$  sei eine multiplikative Struktur (das heißt eine Menge, in der eine nicht notwendig assoziative Multiplikation definiert ist) mit Einselement.  $\Gamma$  sei eine beliebige Gruppe. Kleine lateinische und griechische Buchstaben sollen stets Elemente von  $A$  bzw.  $\Gamma$ , insbesondere  $e$  und  $\varepsilon$  das Einselement von  $A$  bzw.  $\Gamma$  bezeichnen. Wir betrachten die Menge der Paare  $(a, \alpha)$  ( $a \in A, \alpha \in \Gamma$ ); die Gleichheit der Elemente  $(a, \alpha), (a', \alpha')$  ist mit  $a = a', \alpha = \alpha'$  definiert. Die Menge dieser Paare machen wir zu einer Struktur, die wir mit  $A \circ \Gamma$  bezeichnen, so daß wir eine Multiplikation der Elemente durch

$$(1) \quad (a, \alpha)(b, \beta) = (ab, \alpha^b \beta)$$

definieren, wobei  $\alpha^b (\in \Gamma)$  eine Funktion der Argumente  $b (\in A), \alpha (\in \Gamma)$  bezeichnet, stets unterworfen der Anfangsbedingung

$$(2) \quad \alpha^e = \alpha \quad (\alpha \in \Gamma).$$

Wir beweisen den folgenden

HILFSSATZ A. Die Struktur  $A \circ \Gamma$  ist dann und nur dann eine Gruppe, wenn  $A$  eine Gruppe ist und

$$(3) \quad \alpha^{bc} = (\alpha^b)^c \quad (b, c \in A, \alpha \in \Gamma),$$

$$(4) \quad (\alpha\beta)^c = \alpha^c \beta^c \quad (c \in A, \alpha, \beta \in \Gamma)$$

gelten. Diese Gruppen  $A \circ \Gamma$  sind die sämtlichen faktorenfreien (also zerfallenden) Schreierschen Erweiterungen von  $\Gamma$  mit  $A$ .

Die Multiplikation (1) ist dann und nur dann assoziativ, wenn  $A$  assoziativ und

$$(\alpha^b \beta)^c = \alpha^{bc} \beta^c$$

erfüllt ist. Die letztere Bedingung ist mit (3) und (4) gleichwertig. Aus (4) folgt ferner

$$(5) \quad \varepsilon^a = \varepsilon \quad (a \in A).$$

Man sieht nach (2), (5), daß  $A \circ \Gamma$  ( $e, \varepsilon$ ) zum Einselement hat. Eine einfache Rechnung zeigt, daß die Existenz des Linksinverses von  $(a, \alpha)$  in  $A \circ \Gamma$  mit der Existenz des Linksinverses von  $a$  in  $A$  gleichwertig ist. Die letzte Behauptung ist eine Folgerung von Korollar 1 des Satzes 1 in [3].

Nun bezeichne  $D$  eine Struktur mit zwei Verknüpfungen (d. h. eine Menge, in der die Addition und Multiplikation ausführbar ist), die ein Nullelement  $\underline{0}$  hat. Ferner sei ein beliebiger Ring  $P$  gegeben. Wir bezeichnen die Elemente von  $D$  bzw.  $P$  mit  $\underline{0}, a, b, \dots$  bzw.  $0, \alpha, \beta, \dots$ . Mit  $D \circ P$  bezeichnen wir die Menge der Paare  $(a, \alpha)$  ( $a \in D, \alpha \in P$ ), in der die Gleichheit der Elemente  $(a, \alpha), (a', \alpha')$  mit  $a = a', \alpha = \alpha'$  definiert ist. Wir machen diese Menge



zu einer Struktur so, daß die Addition und Multiplikation durch

$$(6) \quad (a, \alpha) + (b, \beta) = (a + b, \alpha + \beta),$$

$$(7) \quad (a, \alpha)(b, \beta) = (ab, \alpha b + a\beta + \alpha\beta)$$

definiert werden, wobei Funktionen  $\alpha b, a\beta (\in P; a, b \in D; \alpha, \beta \in P)$  den Anfangsbedingungen

$$(8) \quad 0\alpha = \alpha 0 = 0 \quad (\alpha \in P)$$

unterworfen sind. (Diese Funktionen können wir als „Operatorprodukte“ auffassen, also ist  $D$  gleichzeitig zweiseitiger „Operatorbereich“ von  $P$ .)

Es gilt der folgende, im wesentlichen von L. RÉDEI [2], [3] herrührende

**HILFSSATZ B.** *Die Struktur  $D \circ P$  ist dann und nur dann ein Ring, wenn  $D$  ein Ring ist und*

$$(9) \quad (a + b)\gamma = a\gamma + b\gamma, \quad \alpha(b + c) = \alpha b + \alpha c,$$

$$(10) \quad ab\gamma = a(b\gamma), \quad \alpha(bc) = (\alpha b)c,$$

$$(11) \quad a(\beta + \gamma) = a\beta + a\gamma, \quad (\alpha + \beta)c = \alpha c + \beta c,$$

$$(12) \quad a\beta\gamma = (a\beta)\gamma, \quad \alpha\beta c = \alpha(\beta c),$$

$$(13) \quad (a\beta)c = a(\beta c),$$

$$(14) \quad (\alpha b)\gamma = \alpha(b\gamma)$$

für  $a, b, c \in D$  und  $\alpha, \beta, \gamma \in P$  gelten. Diese Ringe sind die sämtlichen faktorenfreien (also zerfallenden) Schreierschen Erweiterungen von  $P$  mit  $D$ .

Der Beweis ergibt sich nach RÉDEI [2] Satz 2 und [3] Satz 4 nebst Korollar mit einer leichten Modifikation.

4. Nun betrachten wir die Menge  $A = a, b, \dots$  aller Abbildungen einer Gruppe  $\Gamma$  in sich. Für ein beliebiges Element  $a (\in A)$  bezeichnen wir mit  $a^\alpha$  das entsprechende Bild von  $a$ . In der Menge  $A$  definieren wir das Produkt  $ab$  — wie üblich — als dasjenige Element in  $A$ , wofür (3) gilt.  $A$  ist assoziativ und hat auch ein Einselement  $e$ , die identische Abbildung von  $\Gamma$ . Die Menge aller Paare  $(a, \alpha) (a \in A, \alpha \in \Gamma)$  machen wir durch die Multiplikation (1) zu einer Struktur, die wir mit  $A \circ \Gamma$  bezeichnen.

**DEFINITION A<sub>1</sub>.** *Unter dem Holomorph der Gruppe  $\Gamma$  verstehen wir die maximale Untergruppe von der Struktur  $A \circ \Gamma$ , die die Gruppe  $\Gamma$  enthält.*

Eine Untergruppe von  $A \circ \Gamma$ , die  $\Gamma$  enthält — wie man aus Hilfssatz A sieht — ist eine faktorenfreie Schreiersche Erweiterung von  $\Gamma$  mit einer passenden Automorphismengruppe von  $\Gamma$ . Daraus folgt die Existenz einer einzigen maximalen Untergruppe von  $A \circ \Gamma$ , in der  $\Gamma$  enthalten ist. Diese maximale Untergruppe ist offenbar die der vollen Automorphismengruppe  $A^*$  von  $\Gamma$  zugehörige faktorenfreie Schreiersche Erweiterung  $A^* \circ \Gamma$  von der Gruppe  $\Gamma$ .

Es gilt also der folgende



SATZ A. Die Definitionen  $A_0$  und  $A_1$  sind äquivalent.

Betrachten wir nun die Menge  $D = a, b, \dots$  aller Doppelabbildungen eines Ringes  $P$  in sich. Unter einer Doppelabbildung  $\alpha (\in D)$  von  $P$  in sich verstehen wir nach RÉDEI [4] das System von zwei Abbildungen  $\alpha \rightarrow a\alpha$ ,  $\alpha \rightarrow \alpha a$  ( $\alpha \in P$ ) von  $P$  in sich, wo die Elemente  $a\alpha$  und  $\alpha a$  die Bilder von  $\alpha$  bezeichnen. Die triviale Doppelabbildung von  $P$  in sich, wobei alle Bild-elemente gleich 0 sind, wird mit  $\underline{0}$  bezeichnet. In der Menge  $D$  dieser Doppelabbildungen von  $P$  in sich definiert man die Summe  $a + b$  und das Produkt  $ab$  als diejenigen Elemente der Menge  $D$ , wofür (9), (10) gelten. In der Menge aller Paare  $(a, \alpha)$  ( $a \in D$ ,  $\alpha \in P$ ) definieren wir die Addition und Multiplikation durch (6), (7). Diese Struktur bezeichnen wir mit  $D \circ P$ .

Die neue Definition der Holomorphe eines Ringes  $P$  ist die folgende:

DEFINITION  $B_1$ . Unter den Holomorphen des Ringes  $P$  verstehen wir die maximalen Unterringe der Struktur  $D \circ P$ , die den Ring  $P$  enthalten.

Nach Hilfssatz B ist ein Unterring von  $D \circ P$ , die  $P$  enthält, eine faktorenfreie Schreiersche Erweiterung von  $P$  mit einem passenden Ring von befreundeten Doppelhomothetismen von  $P$ . Da jeder Ring von befreundeten Doppelhomothetismen von  $P$  nach RÉDEI [4] in mindestens einem ebensolchen maximalen Ring  $D^*$  enthalten ist, folgt die Existenz der maximalen Unterringe von  $D \circ P$ , die den Ring  $P$  enthalten. Diese maximalen Unterringe fallen also mit den maximalen Ringen  $D^*$  von befreundeten Doppelhomothetismen von  $P$  zugehörigen faktorenfreien Schreierschen Erweiterungen  $D^* \circ P$  des Ringes  $P$  zusammen.

Wir erhalten daraus den folgenden

SATZ B. Die Definitionen  $B_0$  und  $B_1$  sind äquivalent.

(Eingegangen am 29. April 1954.)

### Literaturverzeichnis

- [1] W. H. MILLS, On the non-isomorphism of certain holomorphs, *Trans. Amer. Math. Soc.*, **74** (1953), S. 428—443.
- [2] L. RÉDEI, Über gewisse Ringkonstruktionen durch schiefes Produkt, *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 185—188.
- [3] L. RÉDEI, Die Verallgemeinerung der Schreierschen Erweiterungstheorie, *Acta Sci. Math.*, **14** (1952), S. 252—273.
- [4] L. RÉDEI, Die Holomorphentheorie für Gruppen und Ringe, *Acta Math. Acad. Sci. Hung.*, **5** (1954), S. 169—195 (die vorstehende Arbeit).



## НОВОЕ ОПРЕДЕЛЕНИЕ ГОЛОМОРФА ГРУППЫ И ГОЛОМОРФОВ КОЛЬЦА

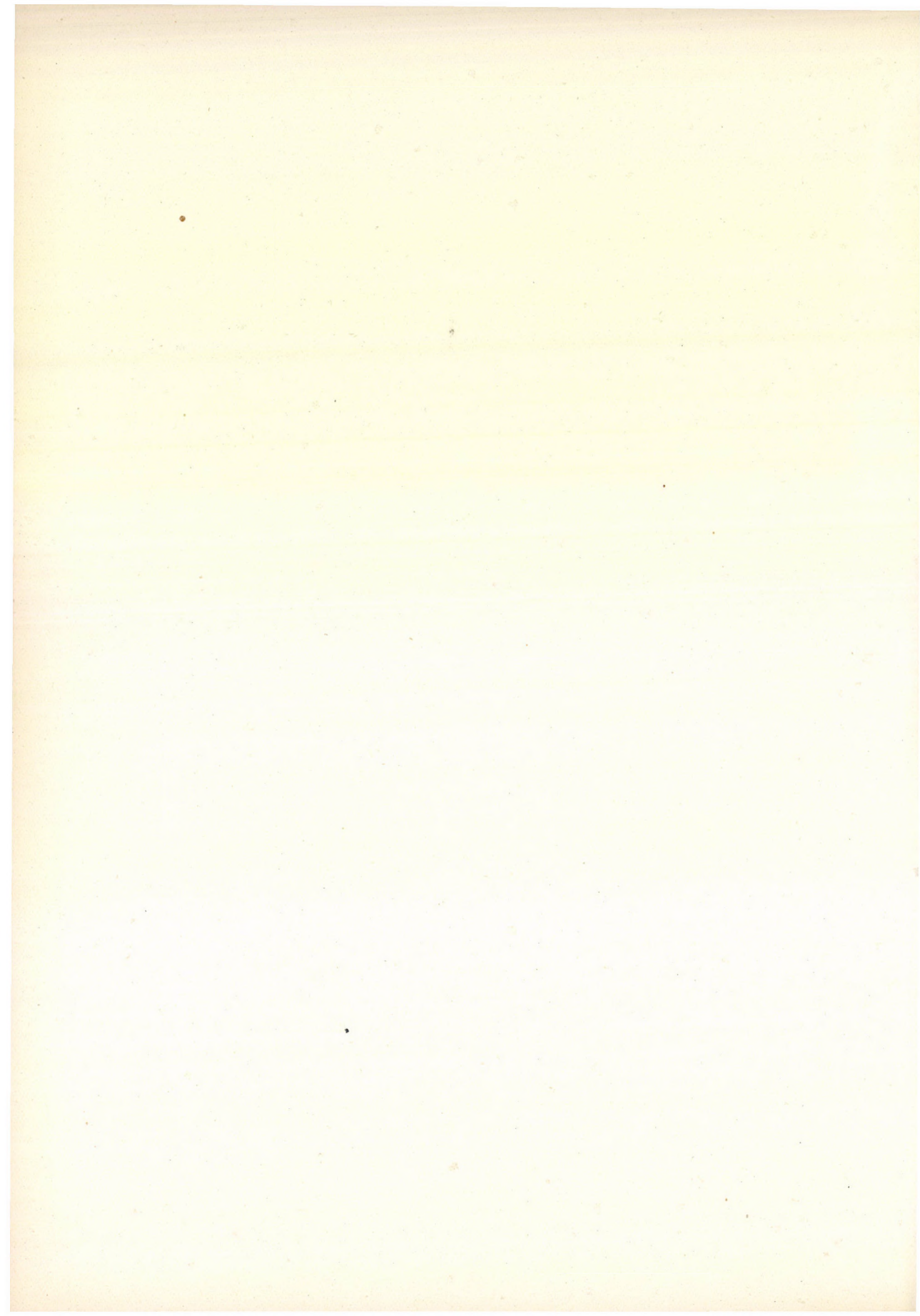
Й. СЕНДРЕИ (Cseredi)

(Резюме)

Исследуя аналог в теории колец понятия голоморфа группы, Л. Рэдэи в своей работе (4) ввел понятие голоморфов кольца.

В настоящей работе даются новые простые определения голоморфа группы и голоморфов кольца (определения  $A_1$  и  $B_1$ ), указывающие с новой точки зрения на аналогию между этими понятиями. Если сопоставить два способа определения голоморфа, то оказывается, что предлагаемые нами определения носят менее явный характер, однако легко доказать их эквивалентность известным явным определениям (теоремы  $A$  и  $B$ ).







# ON SECONDARY PROCESSES GENERATED BY A POISSON PROCESS AND THEIR APPLICATIONS IN PHYSICS

By  
LAJOS TAKÁCS (Budapest)  
(Presented by A. RÉNYI)

## § 1. Formulation of the problem

Let us consider a stochastic process of the Poisson type. Let the density of the occurrence of the events be a non-negative, continuous and bounded function  $\lambda(u)$  of the time parameter  $u$ . The average number of the events occurring in the time interval  $(0, t)$  is, as well known,

$$(1) \quad A(t) = \int_0^t \lambda(u) du$$

and the probability of exactly  $k$  events occurring in the time interval  $(0, t)$  is

$$(2) \quad P(t, k) = e^{-A(t)} \frac{[A(t)]^k}{k!} \quad (k = 0, 1, 2, \dots).$$

The present work is devoted to investigating the following problem. Suppose that every event in the Poisson process is the starting point of a signal (random pulse). Let the progress in time of the amplitude of a signal be  $f(u, \chi)$  where  $u$  is the time counted from the beginning of the signal and the parameter  $\chi$  is a random variable. We suppose that the parameters belonging to different events are mutually independent random variables with the common distribution function  $P(\chi \leq x) = H(x)$ . We suppose, finally, that the different signals linearly superpose and shall consider the secondary stochastic process which consists of the sum of these signals.

We shall distinguish the following cases:

1) Let us consider the Poisson process in the time interval  $0 \leq u < \infty$  and denote the moments of occurrence of the events of the Poisson process by  $u_1, u_2, \dots, u_k, \dots$ . Let us denote the values of the parameters belonging to the produced signals in order by  $\chi_1, \chi_2, \dots, \chi_k, \dots$ . The parameters  $\{\chi_k\}$  are supposed to be equidistributed, mutually independent, non-negative random variables. We denote the sum of the amplitudes of the signals in



the moment  $u$  by  $r_l(u)$ , that is

$$(3) \quad r_l(u) = \sum_{0 \leq u_k \leq u} f(u - u_k, \chi_k).$$

Clearly  $r_l(0) = 0$  and  $r_l(u)$  is defined for  $u \geq 0$ . Let  $P(r_l(u) \leq x) = G(u, x)$  denote the distribution function of the random variable  $r_l(u)$ .

2) Let us consider the case when  $\lambda(u) \equiv \lambda$ , i. e. the Poisson process is homogeneous in time, and suppose that each event occurring in the time interval  $-\infty < u < \infty$  is the starting point of a signal. Let us denote the sum of the signals produced by the events occurring in the time interval  $(-\infty, u)$  at the moment  $u$  by

$$(4) \quad r_l^*(u) = \sum_{-\infty < u_k \leq u} f(u - u_k, \chi_k).$$

In this case, as we shall see, under general conditions, the random function  $r_l^*(u)$  exists for all  $u$  and its distribution function is independent of  $u$ . Let us denote this by  $P(r_l^*(u) \leq x) = G^*(x)$ . It is easily seen that if in case 1) the underlying Poisson process is homogeneous in time, i. e.  $\lambda(u) \equiv \lambda$ , then we have  $G^*(x) = \lim_{u \rightarrow \infty} G(u, x)$ .

In the present paper we shall first discuss the special case  $f(u, \chi) = \chi e^{-\alpha u}$  where  $\alpha$  is a positive constant. This case is interesting from a theoretical point of view, because it is an interesting application of purely discontinuous Markov processes; on the other hand, this case is important also from practical point of view. We determine further the expected number of transition of  $r_l(u)$  and  $r_l^*(u)$ , respectively, through on a fixed threshold value  $a$  during some time interval.

Using a generalisation of a well-known theorem concerning the sum of independent random variables, we shall examine the stochastic processes mentioned in the more general case, when the functions  $f(u, x)$  are arbitrary. After this we shall investigate the auto-correlation function and the harmonic analysis of the process  $r_l^*(u)$ . Finally, we shall give a few examples in the field of particle counting.

The stochastic processes defined above have been in special cases the subject of a great deal of discussion in physical and mathematical literature. Without endeavouring to achieve completeness, we mention the papers of N. CAMPBELL [1], [2], E. N. ROWLAND [17], A. JA. KHINTCHINE [7], H. HURWITZ and M. KAC [6] and S. O. RICE [16]. Among these it is the paper of S. O. RICE [16] which is the most far-reaching from a practical point of view, however, RICE treated only the homogeneous process and supposed  $f(u, \chi) = \chi f(u)$  where  $\chi$  is a random variable. Furthermore RICE did not deal with the existence of the limiting process and the limiting distribution. We shall use a method more simple than those employed by the authors mentioned and by this method we shall obtain more general results.



We also mention a theorem by A. RÉNYI [15] concerning a process of happenings generated by a Poisson process. He has proved that the number of happenings which are just going on at a given instant follows a Poisson distribution. This result can be obtained from our discussion, by specializing  $f(u, \chi) = 1$  if  $0 \leq u \leq \chi$  and  $f(u, \chi) = 0$  elsewhere.

## § 2. The case of exponentially decreasing signals

In this and the following two sections we shall discuss the process in special cases when the progress of the signals in time is described by the function

$$(5) \quad f(u, \chi) = \begin{cases} \chi e^{-\alpha u} & \text{if } u \geq 0, \\ 0 & \text{if } u < 0 \end{cases}$$

where  $\alpha$  is a positive constant and  $\chi$  a positive random variable. In what follows we give the distribution function  $G(u, x)$  and we shall deal with the question: Under what conditions does the distribution function  $G^*(x)$  exist and how can it be determined? Furthermore, we shall also discuss the following problems. We shall say that at the time  $u$  the process is in the state  $A$  if at the moment  $u$  the value of the random function  $\eta(u)$  (resp.  $\eta^*(u)$ ) is  $< a$  and we shall say that it is in the state  $B$  if it is  $\geq a$  ( $\eta(u) \geq a$  or  $\eta^*(u) \geq a$ ). Here  $a$  denotes a fixed positive number. Now, the question arises: What will be the mean duration of the process staying in the state  $B$  during the time interval  $(0, t)$ ? We denote this mean duration by  $\tau(t, a)$  in case 1) and by  $\tau^*(t, a)$  in the case 2). An other question is: What will be the average number of transitions  $A \rightarrow B$  in the time interval  $(0, t)$ ? We denote this number by  $m(t, a)$  in the case 1) and by  $m^*(t, a)$  in the case 2).

Let us call the events occurring in the underlying Poisson process *real events* and the transitions  $A \rightarrow B$  *apparent events*. The question now arises as to how it is possible to conclude, at least in case 2), from the density of apparent events to the density of real events.

First of all, however, we shall mention an important field of application in experimental physics, illustrating what has been said.

AN EXAMPLE FROM EXPERIMENTAL PHYSICS. In order to concretize our discussion, let us think of the example of counting particles by means of an electron multiplier. The corpuscles or photons incident on the cathode of an electron multiplier are releasing electrons. Along a line of electrodes the number of electrons becomes multiplied in succession by means of secondary emission.

Let  $p_0(k)$  denote the probability of an incident particle releasing  $k$  electrons from the cathode, and let the number of multiplying plates of the electron multiplier be  $r$ , for each of which  $p(k)$  denotes the probability of an incident electron releasing  $k$  secondary



electrons. If we suppose that the multiplication process of each electron is independent of that of the others, it is easily possible to determine, with the aid of the theory of multiplicative processes, the probability that an avalanche of electrons released by a particle incident on the cathode, consists of  $k$  electrons at the moment when it reaches the anode. If the generating function of  $p_0(k)$  is denoted by  $f_0(s) = \sum_{k=0}^{\infty} p_0(k)s^k$ , that of  $p(k)$

by  $f(s) = \sum_{k=0}^{\infty} p(k)s^k$  and that of the number of electrons released on the  $n$ -th multiplier

plate by  $g_n(s)$ , then  $g_0(s) = f_0(s)$  and  $g_n(s) = f[g_{n-1}(s)]$  ( $n = 1, 2, \dots, r$ ). With the aid of this recursive formula it is possible to determine  $g_r(s)$ , the generating function of the probability of the number of electrons incident on the anode. When the generating function is known, it is possible to determine the probabilities themselves too.

If the number of electrons incident on the anode is  $k$ , these electrons will charge the condenser, possessing the capacity  $C$ , of the anode circuit to

a voltage  $x = \frac{ke}{C}$  proportional to the number of electrons ( $e = 1.6 \cdot 10^{-19}$  coul.

is the charge of an electron). Whilst the random variable  $k$  assumes the discrete values  $k = 0, 1, 2, \dots$ , the random variable  $x$  can, with a good approximation, be considered continuous. As a matter of fact, for instance in the

case of  $C = 1 \text{ pF}$  we have  $\Delta x = \frac{\Delta k \cdot e}{C} = 1.6 \cdot 10^{-7}$  volt. The condenser charged

to a voltage  $x$  will become discharged through the resistance  $R$ , its voltage decreasing in time in an exponential manner, in accordance with the function  $x e^{-t/RC}$ . The reciprocal value of the time constant  $RC$  is equal to the constant  $\alpha$ , whilst the distribution function of the magnitude of the voltage impulse is  $H(x)$ .

This phenomenon can be described by the stochastic processes mentioned above. In many cases, such as in case of disintegration of radioactive atoms, or in case of cosmic radiation, the time instants of the corpuscles or photons arriving at the cathode form a Poisson process. In this case the events are the particles arriving at the cathode. The multiplication processes of the electrons starting from the cathode are mutually independent and regarding the new particles arriving during the period of time of an impulse, the electron multiplier remains effective. The electron avalanche arriving at the anode will cause a voltage impulse of exponential progress in the external circuit, the amplitude of which voltage impulse is a random variable. The various impulses superpose linearly.

The phenomenon can be observed with the aid of a recording apparatus which is connected with a counter. The counter will indicate as soon as the value of the sum of the signals will exceed a given threshold voltage  $a$  and the relay remains connected as long as the value of the signals exceeds  $a$ . In this case the problem is to determine the average number and the mean duration of the connections of the relay during the period  $t$ .



We remark that the average number of connections of the relay will not be affected by the fact that the relay should remain connected during the whole time of duration of the state  $B$ . This serves only for illustrating the mean length of time of staying in the state  $B$ ; although this is a case not unrealisable in practice.

We distinguish two different kinds of problems.

**PROBLEM OF OBSERVATIONS.** In the case of the counting of particles effected with the aid of a single electron multiplier we shall say that an observation begins at a given moment, if the relay effects connection. The average number of observations during the period  $t$  will be equal to the average number of connections of the relay. In this case it is a problem how it is possible at different values of  $a$ , that is to say, with different threshold voltages, to conclude from the density of apparent events (density of observations) to the density of real events. Another interesting point is the question of how to select the most advantageous value of  $a$ .

**PROBLEM OF COINCIDENCES.** Using simultaneously a plurality of electron multipliers for the observation of random events, we obtain the problem of coincidence. Now the signals supplied by the totality of electron multipliers are examined by using a single coincidence apparatus. The counter indicates when the sum of the signals exceeds the adjusted threshold voltage  $a$ . Here the problem arises: of which value (for a given value of  $a$ ) the average number of chance coincidences during the period of time  $t$  will be. The knowledge of this average number is of particular importance in those cases when it is being examined whether any correlations exist between two or more processes. In this case the problem of finding the value of  $a$  which is the most advantageous to effect the examination of correlation possesses some interest.

This problem can be reduced to the preceding one, if we suppose that the processes of counting with the individual electron multipliers are mutually independent and that the magnitude of the voltage impulse produced by an incident particle presents the same probability distribution on the anode of each electron multiplier. In this case, if the density of the particles arriving at the various electron multipliers is  $\lambda_1(u), \lambda_2(u), \dots, \lambda_m(u)$ , the  $m$  electron multipliers can be replaced by a single one in which the density of the arrival of the particles is  $\lambda(u) = \lambda_1(u) + \lambda_2(u) + \dots + \lambda_m(u)$ .

From the point of view of how to construct the amplifier circuits, the harmonic analysis of our stochastic process possesses some importance.



### § 3. The determination of the distribution functions

Our main problem in this section is to determine the distribution functions  $P(\eta_1(u) \leq x) = G(u, x)$  and  $P(\eta_1^*(u) \leq x) = G^*(x)$  in case  $f(u, \chi) = \chi e^{-u}$ . The random functions  $\eta_1(u)$  and  $\eta_1^*(u)$  describe a Markov process. In fact, if we know in a moment the value of  $\eta_1(u)$  or  $\eta_1^*(u)$ , then this present state uniquely determines the future stochastic development of the process. If, however, the time direction is reversed, that is to say, if it is counted backwards from a given instant  $u$ , the random function

$$(6) \quad \zeta(u, z) = \sum_{u-z \leq u_k \leq u} f(u - u_k, \chi_k)$$

representing the sum of the values assumed at the time  $u$  by the signals occurring in the time interval  $(u-z, u)$  is introduced, then the process described by  $\zeta(u, z)$  ( $u$  is fixed and  $z$  is considered as a variable) will be a Markov process in  $z$  having independent increments, i. e. an *additive Markov process*. By this we understand that the "future" stochastic development of the process  $\zeta(u, z)$  depends neither on the present state nor the past history of the process. With this procedure we obtain a process which is more easily discussed than the preceding one.

Now let  $u$  be fixed and consider  $\zeta(u, z)$  as a function of  $z$ . Then it will be convenient to write the density of the events in the form  $p(z) = \lambda(u-z)$ . With the aid of  $\zeta(u, z)$ ,  $\eta_1(u)$  and  $\eta_1^*(u)$  can be expressed in a simple manner, namely  $\eta_1(u) = \zeta(u, u)$  when  $p(z) = \lambda(u-z)$  and  $\eta_1^*(u) = \lim_{z \rightarrow \infty} \zeta(u, z)$  when  $p(z) \equiv \lambda$ . Let  $P(\zeta(u, z) \leq x) = F(z, x)$ , then  $G(u, x) = F(u, x)$  provided that  $p(z) = \lambda(u-z)$  and  $G^*(x) = \lim_{z \rightarrow \infty} F(z, x)$  provided that  $p(z) \equiv \lambda$ .

So far, we have not dealt with the question of the existence of  $\eta_1^*(u)$ . Now we prove that  $\eta_1^*(u)$  exists with probability 1, when  $M(\chi_n) < \infty$ . Namely  $\eta_1^*(u)$  can be expressed as a sum of independent random variables, as follows:

$$\eta_1^*(u) = \sum_{n=0}^{\infty} \Delta \zeta(u, n \Delta z)$$

where

$$\Delta \zeta(u, n \Delta z) = \zeta(u, (n+1) \Delta z) - \zeta(u, n \Delta z) \quad (n = 0, 1, 2, \dots).$$

So we can apply the well-known three series theorem of A. N. KOLMOGOROV (Cf. P. R. HALMOS, *Measure Theory* (1950), p. 199) which yields necessary and sufficient conditions for the convergence, with probability 1, of series of independent random variables. If the series of the expected values of the absolute values of the random variables is convergent, then (choosing  $c=1$ ) these conditions are satisfied and the series in question converges with probability 1. In our case if  $M(\chi_n) < \infty$ , then  $\sum_{n=0}^{\infty} M(|\Delta \zeta(u, n \Delta z)|) < \infty$ , consequently  $\eta_1^*(u) = \lim_{z \rightarrow \infty} \zeta(u, z)$  exists with probability 1. In addition it is clear,



that  $G^*(x) = \lim_{z \rightarrow \infty} F(z, x)$  exists, when  $\nu_i^*(u)$  does, but we shall give also an extra proof of the existence of  $G^*(x)$ .

The process  $\zeta(u, z)$  is a so-called discontinuous Markov process in  $z$ . A. N. KOLMOGOROV [10] was the first to consider purely discontinuous Markov processes in which the state changes only in jumps and established the integro-differential equation of such a process. W. FELLER [4], [5] made this process the subject of a thorough investigation, has proved the existence and unicity statements and represented the general solution of the integro-differential equation in the form of a uniformly convergent infinite series. The process  $\zeta(u, z)$  constitutes a special case of the process discussed by W. FELLER, as in our case the probability of the occurrence and of the magnitude of the change is independent of the actual value of  $\zeta(u, z)$  and it is a generalisation of the case discussed by A. JA. KHINTCHINE [8] in which the probabilities of the change are, moreover, also independent of time. It would be possible for us to apply the general discussion of W. FELLER to our process, but it appears to be preferable, because it is simpler to give a separate discussion of this special process. Now we shall prove the following

**THEOREM 1.** *The distribution function  $P(\zeta(u, z) \leq x) = F(z, x)$  of the random variable  $\zeta(u, z)$  satisfies the following integro-differential equation:*

$$(7) \quad \frac{\partial F(z, x)}{\partial z} = p(z) \left[ \int_0^x H[(x-y)e^{\alpha z}] d_y F(z, y) - F(z, x) \right],$$

which possesses the only solution  $F(z, x)$  satisfying the initial condition  $F(0, x) \equiv 1$  ( $x \geq 0$ ). Let the characteristic function of  $F(z, x)$  be

$$(8) \quad \Phi(z, \omega) = \int_0^{\infty} e^{i\omega x} d_x F(z, x)$$

and that of  $H(x)$

$$(9) \quad \varphi(\omega) = \int_0^{\infty} e^{i\omega x} dH(x).$$

Then we have

$$(10) \quad \log \Phi(z, \omega) = - \int_0^z p(v) [1 - \varphi(\omega e^{-\alpha v})] dv.$$

If  $\Phi(z, \omega)$  is known, the distribution function  $F(z, x)$  can be determined uniquely.

**PROOF.** For  $0 < \Delta z$  we get

$$(11) \quad \begin{aligned} & F(z + \Delta z, x) = \\ & = p(z) \Delta z \int_0^x H[(x-y)e^{\alpha z}] d_y F(z, y) + [1 - p(z) \Delta z] F(z, x) + o(\Delta z). \end{aligned}$$



As a matter of fact, the event that at the moment  $z + \Delta z$ ,  $\zeta(u, z + \Delta z) \leq x$ , can be realised in such a manner that  $\zeta(u, z) = y$  where  $0 \leq y < x$  and that in the time interval  $(z, z + \Delta z)$  there occurs a change whose probability is  $p(z)\Delta z + o(\Delta z)$  and, the probability of this change being  $\leq x - y$ , it is equal to  $H[(x - y)e^{\alpha z}]$ ; or at the moment  $z$ ,  $\zeta(u, z) \leq x$  the probability whereof is  $F(z, x)$ , and no change takes place in the interval  $(z, z + \Delta z)$ , the probability whereof is  $1 - p(z)\Delta z + o(\Delta z)$ . The probability of more than one event occurring in time interval  $(z, z + \Delta z)$  is  $o(\Delta z)$ . It is easily seen that it does not matter, in which point of the time interval  $(z, z + \Delta z)$  the event takes place and so we can consider the moment  $z$  as the occurring point.

From (11) it follows

$$(12) \quad \frac{F(z + \Delta z, x) - F(z, x)}{\Delta z} = p(z) \left[ \int_0^x H[(x - y)e^{\alpha z}] d_y F(z, y) - F(z, x) \right] + \frac{o(\Delta z)}{\Delta z}.$$

Letting  $\Delta z \rightarrow 0$  we obtain the equation

$$(13) \quad \frac{\partial F(z, x)}{\partial z} = p(z) \left[ \int_0^x H[(x - y)e^{\alpha z}] d_y F(z, y) - F(z, x) \right].$$

Thus we have shown that the right-hand derivative of  $F(z, x)$  with respect to  $z$  exists and is equal to the right-hand side of (13). Similarly, it can be shown that the left-hand derivative of  $F(z, x)$  also exists and is equal to the right-hand side of (13), that is to say, the derivative of  $F(z, x)$  with respect to  $z$  exists and satisfies equation (7).

In consequence of (9) we get

$$(14) \quad \int_0^{\infty} e^{i\omega x} d_x H(xe^{\alpha z}) = \int_0^{\infty} e^{i\omega x e^{-\alpha z}} dH(x) = \varphi(\omega e^{-\alpha z}).$$

Thus, passing from equation (7) to the characteristic functions, we obtain easily

$$(15) \quad \frac{\partial \Phi(z, \omega)}{\partial z} = -p(z) [1 - \varphi(\omega e^{-\alpha z})] \Phi(z, \omega).$$

Considering that  $\lambda(u)$  has been supposed to be a continuous function, and since a characteristic function is known to be uniformly continuous for  $-\infty < \omega < \infty$ ,  $\varphi(\omega e^{-\alpha z})$  is also continuous. Under these conditions our differential equation possesses a single solution satisfying the initial condition  $\Phi(0, \omega) = 1$  which follows from the condition concerning  $F(0, x)$  and with this initial condition the solution of the linear differential equation (15), written in logarithmic form, is

$$(16) \quad \log \Phi(z, \omega) = - \int_0^z p(r) [1 - \varphi(\omega e^{-\alpha r})] dr.$$



That our integro-differential equation (7) possesses the single solution  $F(z, x)$  satisfying the initial condition, and that it is a distribution function for  $x$ , follows from the theorem concerning the general case due to W. FELLER [4], [5]. As its characteristic function  $\Phi(z, \omega)$  is known, the distribution function  $F(z, x)$  can be determined uniquely [3, p. 93]:

Now let us suppose that  $\lambda(u) \equiv \lambda$  (constant), i. e.  $p(z) \equiv \lambda$ . Then the following question arises: under what conditions does the limiting distribution  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  exist? We shall prove the following

THEOREM 2. *In case the mean value*

$$(17) \quad M = \int_0^{\infty} x dH(x)$$

of the distribution function  $H(x)$  is finite, the solution  $F(z, x)$  satisfying our initial condition of the integro-differential equation (7) converges for  $z \rightarrow \infty$  to a limiting distribution  $F(x)$  whose characteristic function  $\Phi(\omega)$  satisfies

$$(18) \quad \log \Phi(\omega) = -\frac{\lambda}{\alpha} \int_0^1 \frac{1 - \varphi(\omega v)}{v} dv.$$

As  $\Phi(\omega)$  is known,  $F(x)$  can be determined uniquely.

PROOF. AS it is well known (see [3], p. 96), from  $\lim_{z \rightarrow \infty} \Phi(z, \omega) = \Phi(\omega)$ , if  $\Phi(\omega)$  is continuous at  $\omega = 0$ , it follows that  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  exists,  $F(x)$  is a distribution function and  $\Phi(\omega)$  is its characteristic function. Thus the limiting distribution function  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  will exist, if the limit

$$(19) \quad \lim_{z \rightarrow \infty} \int_0^z [1 - \varphi(\omega e^{-\alpha v})] dv$$

exists and if this function is continuous at the point  $\omega = 0$ . In this case, denoting by  $\Phi(\omega)$  the characteristic function of  $F(x)$ , i. e.

$$(20) \quad \Phi(\omega) = \int_0^{\infty} e^{i\omega x} dF(x),$$

the following equation will hold:

$$(21) \quad \log \Phi(\omega) = -\lambda \int_0^{\infty} [1 - \varphi(\omega e^{-\alpha v})] dv.$$

We supposed that the mean value

$$(22) \quad M = \int_0^{\infty} x dH(x)$$



is finite and in this case it is easy to conclude that

$$(23) \quad |1 - \varphi(\omega)| = \left| \int_0^{\infty} (1 - e^{i\omega v}) dH(x) \right| \leq M|\omega|.$$

By making use of this result, we obtain

$$(24) \quad \left| \int_0^{\infty} [1 - \varphi(\omega e^{-\alpha v})] dv \right| \leq \frac{M|\omega|}{\alpha}$$

whence it follows that the limit (19) exists for every finite value of  $\omega$  and is continuous at the point  $\omega = 0$ . So we have proved the stated theorem.

Now let us denote the characteristic function of  $G(u, x)$  by  $\Psi(u, \omega)$ , i. e.

$$(25) \quad \Psi(u, \omega) = \int_0^{\infty} e^{i\omega x} d_x G(u, x)$$

and that of  $G^*(x)$  by  $\Psi^*(\omega)$ , i. e.

$$(26) \quad \Psi^*(\omega) = \int_0^{\infty} e^{i\omega x} dG^*(x);$$

then we have

$$(27) \quad \log \Psi(u, \omega) = - \int_0^u \lambda(u-z) [1 - \varphi(\omega e^{-\alpha z})] dz$$

and in case  $M < \infty$

$$(28) \quad \log \Psi^*(\omega) = - \frac{\lambda}{\alpha} \int_0^1 \frac{1 - \varphi(\omega v)}{v} dv.$$

REMARK. If the  $m$ -th moment of the distribution function  $H(x)$  exists, the moments

$$(29) \quad M_j = \int_0^{\infty} x^j dH(x) \quad (j = 0, 1, \dots, m)$$

will, as well known, also exist, and for the characteristic function  $\varphi(\omega)$  of  $H(x)$

$$(30) \quad \varphi(\omega) = \sum_{j=0}^m \frac{M_j}{j!} (i\omega)^j + o(\omega^m)$$

will hold, and thus, by making use of (28), we have

$$(31) \quad \log \Psi^*(\omega) = \frac{\lambda}{\alpha} \sum_{j=1}^m \frac{M_j}{j!} (i\omega)^j + o(\omega^m),$$

i. e., the semi-invariants of the distributions  $G^*(x)$  (which we shall denote by  $A_j$ ) also exist up to the  $m$ -th semi-invariant inclusively, and these are the following:

$$(32) \quad A_j = (-i)^j \left[ \frac{d^j}{d\omega^j} \log \Psi^*(\omega) \right]_{\omega=0} = \frac{\lambda}{\alpha} \frac{M_j}{j} \quad (j = 1, 2, \dots, m).$$

The relation (32) is particularly suitable for the solution of practical problems in which the distribution  $H(x)$  is given empirically.



#### § 4. Determination of the mean values

The determination of  $\tau(t, a)$  and  $\tau^*(t, a)$ . At the time  $u$ , the system is in the state  $B$ , if at the same time the sum of the values of the signals is greater than  $a$ . The probability of this is  $1 - G(u, a)$  in case 1) and  $1 - G^*(a)$  in case 2). It can then easily be seen that in the time interval  $(0, t)$  the mean duration of staying in the state  $B$  will be in case 1)

$$(33) \quad \tau(t, a) = \int_0^t [1 - G(u, a)] du$$

and in case 2)

$$(34) \quad \tau^*(t, a) = [1 - G^*(a)]t.$$

The determination of  $m(t, a)$  and  $m^*(t, a)$ . Between the moments  $u$  and  $u + \Delta u$  a transition  $A \rightarrow B$  occurs, if at the moment  $u$  the sum of the values of the signals is less than  $a$  and a signal begins in the interval  $(u, u + \Delta u)$ , the probability whereof is  $\lambda(u)\Delta u + o(\Delta u)$  and its amplitude exceeds  $a - x$  whose probability is  $1 - H(a - x)$ . The probability of more than one transition  $A \rightarrow B$  occurring in the interval  $(u, u + \Delta u)$  is, as can easily be seen,  $o(\Delta u)$ .

On the basis of what has been said, it results easily that the probability of a transition  $A \rightarrow B$  taking place in the interval  $(u, u + \Delta u)$  is in case 1)

$$(35) \quad \lambda(u)K(u, a)\Delta u + o(\Delta u)$$

where

$$(36) \quad K(u, a) = \int_0^a [1 - H(a - x)] dG(u, x)$$

and in case 2)

$$(37) \quad \lambda K^*(a)\Delta u + o(\Delta u)$$

where

$$(38) \quad K^*(a) = \int_0^a [1 - H(a - x)] dG^*(x).$$

As can readily be seen, the average number of transitions  $A \rightarrow B$  occurring in the time interval  $(u, u + \Delta u)$  will be given in case 1) by (35) and in case 2) by (37). Thus the average number of transitions  $A \rightarrow B$  occurring in the interval  $(0, t)$  is obtained by dividing the interval  $(0, t)$  into  $n$  subintervals of length  $\Delta u = t/n$ , forming the mean value in each subinterval, adding these up and then taking the limit for  $n \rightarrow \infty$ . From the proved continuity of the distribution function  $G(u, x)$  in the variable  $u$ , it follows the continuity of  $K(u, a)$  in the same variable. Thus it is assured the existence of the limit



and as a result we obtain

$$(39) \quad m(t, a) = \int_0^t \lambda(u) K(u, a) du,$$

$$(40) \quad m^*(t, a) = \lambda K^*(a)t$$

in cases 1) and 2), respectively.

REMARK 1. In case 2) it appears from (38) that  $K^*(a)$  can be represented as the difference of two distribution functions and thus its Fourier—Stieltjes transform exists. Let us denote this by

$$(41) \quad \Gamma(\omega) = \int_0^{\infty} e^{i\omega a} dK^*(a).$$

On the basis of (38) for this transform we have

$$(42) \quad \Gamma(\omega) = [1 - \varphi(\omega)] \Psi^*(\omega)$$

and thence  $K^*(a)$  can be determined uniquely.

$K(u, a)$  can also be determined in a similar manner.

REMARK 2. In case 2) the mean length of time of stay in the state  $B$  during the period  $t$  is  $\tau^*(t, a) = [1 - G^*(a)]t$  and the average number of transitions  $A \rightarrow B$  is  $m^*(t, a) = \lambda K^*(a)t$ . Let us denote by  $\mathcal{G}$  the average length of the period of time which has elapsed between the consecutive transitions  $A \rightarrow B$  and  $B \rightarrow A$ .

For the mean value  $\mathcal{G}$  we have

$$(43) \quad \mathcal{G} = \lim_{t \rightarrow \infty} \frac{\tau^*(t, a)}{m^*(t, a)} = \frac{1 - G^*(a)}{K^*(a)}.$$

This statement can be proved with the aid of the following theorem due to A. N. KOLMOGOROV and JU. V. PROHOROV [11]. We recall this theorem in the special case which is interesting for us:

Let us consider the random variables  $\xi_1, \xi_2, \dots, \xi_n, \dots$  with the common expected values  $M(\xi_n) = \mathcal{G} < \infty$ . We form the sums  $\zeta_r = \xi_1 + \xi_2 + \dots + \xi_r$  where  $r$ , the number of terms, itself is again a random variable. If the variables  $\xi_{m+1}, \xi_{m+2}, \dots$  and the event  $r = m$  are independent for each  $m$ , then in case  $M(r) < \infty$  we have

$$M(\zeta_r) = \mathcal{G}M(r).$$

Now let us denote by  $\xi_n$  ( $n = 1, 2, \dots$ ) the distances between two consecutive transitions  $A \rightarrow B$  and  $B \rightarrow A$  in moments  $t \geq 0$  and by  $r$  the number of transitions  $A \rightarrow B$  occurring in the time interval  $(0, t)$ . Then we have  $M(\zeta_r) = \mathcal{G}m^*(t, a)$ . It is easily seen that  $\lim_{t \rightarrow \infty} M(\zeta_r)/\tau^*(t, a) = 1$  and hence

(43) follows. We can apply the cited theorem if we show that the variables  $\xi_n$  ( $n = 1, 2, \dots$ ) are mutually independent variables having the same distri-



bution. This is, indeed, true because the process examined is a Markov process and the future stochastic development of the process counted from the moment of a transition  $B \rightarrow A$  does not depend on the past history of the process and the process reproduces itself from such moments onwards.

REMARK 3. In the case 2) the real density of events is  $\lambda$  and the apparent density of events (the density of transitions  $A \rightarrow B$ ) is  $\lambda'(a) = \lambda K^*(a)$  where  $K^*(a)$  is defined by (38).

### § 5. The case of arbitrary signals

Now we shall discuss the processes 1) and 2) mentioned in the first section in the more general case when the progress in time of a signal is described by an arbitrary function  $f(u, \chi)$ . We suppose that  $f(u, \chi) = 0$  if  $u < 0$  and that  $\chi$  is a positive random variable having the distribution function  $H(x)$ . The random functions  $\eta_i(u)$  and  $\eta_i^*(u)$  connected with an arbitrary function  $f(u, \chi)$  will in general describe a non-Markov process, except in the case  $f(u, \chi) = \chi e^{-au}$ . If, however, the time direction is reversed, i. e. if counting backwards from a given instant  $u$ , further the random function

$$(44) \quad \zeta(u, z) = \sum_{u-z \equiv u_k \equiv u} f(u - u_k, \chi_k)$$

representing the sum of the values assumed at the time  $u$  by the signals occurring in the time interval  $(u - z, u)$  is introduced, the process described by  $\zeta(u, z)$  ( $u$  is fixed and  $z$  is a variable) will already be a Markov process, moreover, it will be a Markov process having independent increments, i. e. an additive Markov process. With the help of  $\zeta(u, z)$ , the functions  $\eta_i(u)$  and  $\eta_i^*(u)$  can be expressed in a simple manner:  $\eta_i(u) = \zeta(u, u)$  and  $\eta_i^*(u) = \lim_{z \rightarrow \infty} \zeta(u, z)$  provided  $\lambda(u) \equiv \lambda$  (constant).

Now  $\eta_i^*(u) = \lim_{z \rightarrow \infty} \zeta(u, z)$  exists with probability 1, if the expression (62) is bounded. Namely  $\eta_i^*(u)$  has the same properties which have been mentioned in § 3 and so it follows from the theorem of A. N. KOLMOGOROV, that  $\eta_i^*(u)$  exists with probability 1 if  $\sum_{n=0}^{\infty} M(|\Delta \zeta(u, n \Delta z)|) < \infty$ . Letting  $\Delta z \rightarrow 0$  in this last expression, we get (62). If (62) is finite, then the above series also converges for sufficiently small  $\Delta z$ ; this proves our statement. In this case the limiting distribution  $G^*(x)$  evidently exists, but we shall give an extra proof of this fact.

This general problem has also an interesting physical application, namely to the discussion of the anode current fluctuation of vacuum tubes. As well known, the moments of electrons emitted from the cathode of vacuum tubes



follow a Poisson process. The electrons starting from the cathode will, during the time whilst they reach the anode, influence a current in the external circuit. Let the progress in time of the current-pulse influenced by an electron emitted at the time  $u=0$  be  $f(u, \chi)$  where  $\chi$  means the initial velocity of the electron which is a random variable. According to classical statistics,  $\chi$  follows a Maxwell distribution. The current-pulses influenced by the individual electrons will linearly superpose and their sum  $\eta(u)$  will give the instantaneous value of the anode current. This problem will be treated by us in a separate paper.

The process  $\zeta(u, z)$  will be treated by making use of a theorem stated in the following section. In the same way in which certain stochastic processes can be reduced to the determination of the distribution function of the sum of independent random variables, the processes discussed up to now can be reduced to the determination of the integral of a certain random function.

### § 6. The fundamental lemma

Now we shall prove the following lemma concerning an additive Markov process  $\zeta(z)$ .

LEMMA. Let us consider the random function  $\zeta(z)$  defined for  $z \geq 0$  for which

1)  $\zeta(0) \equiv 0$ ,

2) if  $z_1 < z_2 \leq z_3 < z_4$ , the random variables  $\zeta(z_2) - \zeta(z_1)$  and  $\zeta(z_4) - \zeta(z_3)$  are independent.

Let the characteristic function of  $\zeta(z + \Delta z) - \zeta(z)$  be denoted by

$$M \{ e^{i\omega [\zeta(z+\Delta z) - \zeta(z)]} \} = \psi_{\Delta z}(z, \omega).$$

3) Let us suppose that the following relation holds uniformly in  $\omega$ :

$$(45) \quad \log \psi_{\Delta z}(z, \omega) = \Delta z \log \psi(z, \omega) + o(\Delta z).$$

Let  $F(z, x)$  denote the distribution function of  $\zeta(z)$  and  $\Phi(z, \omega)$  the characteristic function of  $\zeta(z)$ . Then we have

$$(46) \quad \log \Phi(z, \omega) = \int_0^z \log \psi(u, \omega) du,$$

provided that the integral on the right-hand side exists. If  $\Phi(z, \omega)$  is known,  $F(u, x)$  can be determined uniquely.

We remark that in case when the random function  $\zeta(z)$  is differentiable and its differential quotient is equal to the random function  $\xi(z)$ , i. e.

$\xi(z) = \lim_{h \rightarrow 0} \frac{\zeta(z+h) - \zeta(z)}{h}$  (in probability sense), then condition 3) is satisfied and  $\psi(z, \omega)$  is the characteristic function of  $\xi(z)$ . In this case we may



write symbolically

$$(47) \quad \zeta(z) = \int_0^z \xi(u) du.$$

PROOF. It is well known that the characteristic function of the sum of independent random variables is equal to the product of the characteristic functions of the individual variables. This theorem remains valid also if in the sum of the random variables each variable is independent of the sum of the preceding variables. In this case we shall call the random variables *independent in the wider sense* (it is this fact that enables us to suppose instead of the independence of any number of variable increments only the independence of two such increments). Our theorem is the direct generalisation of this theorem for the case in which the number of the random variables to be added is replaced by a continuously varying parameter. The generalisation can be carried out directly, if the theorem referred to is formulated in such a form, that in case of the sum of independent random variables, or of random variables which are independent in the wider sense, the logarithms of the characteristic functions of the individual variables are added up.

Let us divide the interval  $(0, z)$  by means of the points  $z_0 = 0, z_1, z_2, \dots, z_n = z$  into  $n$  subintervals, each of length  $\Delta z = z/n$  and put  $\zeta(z_i) - \zeta(z_{i-1}) = \xi_i$  ( $i = 1, 2, \dots, n$ ). Then we have

$$(48) \quad \zeta(z) = \xi_1 + \xi_2 + \dots + \xi_n.$$

The random variables  $\xi_1, \xi_2, \dots, \xi_n$  are independent in the wider sense and their characteristic functions are  $\psi_{\Delta z}(z_{i-1}, \omega)$  ( $i = 1, 2, \dots, n$ ). Thus

$$(49) \quad \Phi(z, \omega) = \prod_{i=1}^n \psi_{\Delta z}(z_{i-1}, \omega).$$

Taking the logarithm of both sides, we obtain

$$(50) \quad \log \Phi(z, \omega) = \sum_{i=1}^n \log \psi_{\Delta z}(z_{i-1}, \omega) = \sum_{i=1}^n [\log \psi_{\Delta z}(z_{i-1}, \omega) \Delta z + o(\Delta z)]$$

where the second equality is a consequence of condition 3).

If we suppose that  $\log \psi(z, \omega)$  is integrable in the interval  $(0, z)$ , it will follow that the right-hand side of (50) converges for  $n \rightarrow \infty$  to the integral

$$(51) \quad \int_0^z \log \psi(u, \omega) du$$

and thereby we have justified our statement.

Let us note that if the limit of the integral (46) exists for  $z \rightarrow \infty$  and is continuous at the point  $\omega = 0$ , then the limiting distribution  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  also exists, and if the characteristic function of this limiting distribu-



tion is denoted by  $\Phi(\omega)$ , we have

$$(52) \quad \log \Phi(\omega) = \int_0^{\infty} \log \psi(u, \omega) du.$$

REMARK. If the random function  $\xi(z) = \zeta'(z)$  exists and the  $m$ -th semi-invariant of  $\xi(z)$  is finite, the semi-invariants

$$(53) \quad \lambda_j(z) = (-i)^j \left[ \frac{d^j \log \psi(z, \omega)}{d\omega^j} \right]_{\omega=0} \quad (j = 1, 2, \dots, m)$$

are also finite and in this case the semi-invariants of  $\zeta(z)$  also exist up to the  $m$ -th semi-invariant inclusively,

$$(54) \quad A_j(z) = (-i)^j \left[ \frac{d^j \log \Phi(z, \omega)}{d\omega^j} \right]_{\omega=0} \quad (j = 1, 2, \dots, m).$$

Hence

$$(55) \quad A_j(z) = \int_0^z \lambda_j(u) du \quad (j = 1, 2, \dots, m).$$

As a matter of fact, in this case

$$(56) \quad \log \psi(z, \omega) = \sum_{j=1}^m \frac{\lambda_j(z)}{j!} (i\omega)^j + o(\omega^m)$$

and hence (55) is obtained using (46).

## § 7. The determination of the distribution functions

### $G(u, x)$ and $G^*(x)$

Let us consider the random function  $\zeta(u, z)$ . As  $u$  is fixed, it will be convenient to denote the density of the events by  $p(z) = \lambda(u-z)$ . If we know  $F(z, x)$ , the distribution function of the random variable  $\zeta(u, z)$ , then we can easily get the distribution functions  $G(u, x)$  and  $G^*(x)$  too, since  $G(u, x) = F(u, x)$  provided that  $p(z) = \lambda(u-z)$  and  $G^*(x) = \lim_{z \rightarrow \infty} F(z, x)$  provided that  $p(z) \equiv \lambda$ .

Now we shall prove the following

THEOREM 3. For the characteristic function  $\Phi(z, \omega)$  of  $F(z, x)$  we have

$$(57) \quad \log \Phi(z, \omega) = - \int_0^z p(v) [1 - q(v, \omega)] d\omega$$

where

$$(58) \quad q(u, \omega) = \int_0^{\infty} e^{i\omega f(u, x)} dH(x).$$



PROOF. Let us denote the characteristic function of the random variable  $\zeta(u, z + \Delta z) - \zeta(u, z)$  by  $\psi_{\Delta z}(z, \omega)$ . Then

$$(59) \quad \psi_{\Delta z}(z, \omega) = (1 - p(z)\Delta z) + p(z)\Delta z \int_0^{\infty} e^{i\omega f(z, x)} dH(x) + o(\Delta z).$$

As a matter of fact,  $\zeta(u, z)$  does not vary in the time interval  $(z, z + \Delta z)$ , if no event occurs in this interval, the probability whereof is  $1 - p(z)\Delta z + o(\Delta z)$  and in this case  $\zeta(u, z + \Delta z) - \zeta(u, z) = 0$ . Or, it does vary, the probability whereof is  $p(z)\Delta z + o(\Delta z)$  and if  $\zeta(u, z + \Delta z) - \zeta(u, z) = f(z, y)$  with  $y \leq x$ , the probability of this will be  $H(x)$ . Thus equation (59) is obtained. Introducing by (58) the function  $\varphi(z, \omega)$ , (59) can be written in the form

$$(60) \quad \psi_{\Delta z}(z, \omega) = (1 - p(z)\Delta z) + p(z)\Delta z \varphi(z, \omega) + o(\Delta z)$$

and hence

$$(61) \quad \log \psi(z, \omega) = \lim_{\Delta z \rightarrow 0} \frac{\log \psi_{\Delta z}(z, \omega)}{\Delta z} = -p(z)[1 - \varphi(z, \omega)].$$

Consequently, we can apply our lemma to the random function  $\zeta(u, z)$ . Indeed  $\zeta(u, z)$  evidently satisfies conditions 1) and 2) and we have now shown that it satisfies also condition 3). If the integral (58) exists, then (57) exists too and so the theorem is proved.

Now we are interested under what conditions the limiting distribution function  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  exists. We suppose  $p(z) \equiv \lambda$  (constant). We shall prove the following

THEOREM 4. *If the integral*

$$(62) \quad \int_0^{\infty} \int_0^{\infty} |f(z, x)| dH(x) dz$$

*exists and is finite, then the limiting distribution function  $\lim_{z \rightarrow \infty} F(z, x) = F(x)$  also exists, provided  $\lambda(u) \equiv \lambda$ , and its characteristic function  $\Phi(\omega)$  is given by*

$$(63) \quad \log \Phi(\omega) = -\lambda \int_0^{\infty} [1 - \varphi(z, \omega)] dz.$$

PROOF. According to the mentioned theorem of LÉVY and CRAMÉR it is sufficient to show that

$$(64) \quad \lim_{z \rightarrow \infty} \int_0^z [1 - \varphi(v, \omega)] dv$$

exists and is continuous at  $\omega = 0$ . Let

$$(65) \quad M(z) = \int_0^{\infty} |f(z, x)| dH(x),$$



then we have

$$(66) \quad |1 - \varphi(z, \omega)| \leq |\omega| M(z),$$

and so

$$(67) \quad \left| \int_0^z [1 - \varphi(v, \omega)] dv \right| \leq |\omega| \int_0^z M(v) dv.$$

It follows that, if (62) exists, then the limit (64) will also exist for any fixed  $\omega$ , and will be continuous at the point  $\omega = 0$ .

Let  $\Psi(u, \omega)$  denote the characteristic function of  $G(u, x)$ , i. e.

$$(68) \quad \Psi(u, \omega) = \int_{-\infty}^{\infty} e^{i\omega x} d_x G(u, x),$$

and  $\psi^*(\omega)$  the characteristic function of  $G^*(x)$ , i. e.

$$(69) \quad \psi^*(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} dG^*(x),$$

then we have

$$(70) \quad \log \Psi(u, \omega) = - \int_0^u \lambda(u-z) [1 - \varphi(\omega, z)] dz,$$

provided that (58) exists for every  $z$ , further

$$(71) \quad \log \Psi^*(\omega) = - \lambda \int_0^{\infty} [1 - \varphi(\omega, z)] dz,$$

provided that (62) exists and is finite.

REMARK 1. If  $f(u, \chi) = \chi e^{-u\chi}$  and we suppose that the mean value of  $\chi$  is finite, we obtain as a special case the results of § 3.

REMARK 2. Let  $A_j(u)$  be the  $j$ -th semi-invariant of  $G(u, x)$ , i. e.

$$(72) \quad A_j(u) = (-i)^j \left[ \frac{d^j \log \Psi(u, \omega)}{d\omega^j} \right]_{\omega=0} \quad (j = 1, 2, \dots).$$

It follows from our lemma that in case

$$(73) \quad \lambda_j(z) = \lambda(u-z) \int_0^{\infty} [f(z, x)]^j dH(x)$$

exists and is integrable in the interval  $(0, u)$ , then  $A_j(u)$  also exists and we have

$$(74) \quad A_j(u) = \int_0^u \lambda_j(z) dz.$$

In particular, if  $f(z, \chi) = \chi f(z)$  and the  $j$ -th moment

$$(75) \quad M_j = \int_0^{\infty} x^j dH(x)$$



of  $\chi$  exists, we have

$$(76) \quad A_j(u) = M_j \int_0^u \lambda(u-z)[f(z)]^j dz.$$

If we denote by  $A_j$  the  $j$ -th semi-invariant of  $G^*(x)$  and suppose that it exists, then we get

$$(77) \quad A_j = \lambda \left[ \int_0^\infty \left( \int_0^\infty (f(z, x))^j dH(x) \right) dz \right]$$

and, in the particular case  $f(z, \chi) = \chi f(z)$ ,

$$(78) \quad A_j = \lambda M_j \int_0^\infty [f(z)]^j dz.$$

This is CAMPBELL'S formula well-known in physical literature. We remark that the original Campbell formula, concerning the mean value and the dispersion, has in the course of time been greatly generalised by E. N. ROWLAND, A. JA. KHINTCHINE, S. O. RICE and others.

REMARK 3. It follows from the above theorems that the mean value of  $\eta^*(u)$  is

$$(79) \quad M\{\eta^*(u)\} = \lambda \int_0^\infty \left[ \int_0^\infty f(z, x) dH(x) \right] dz,$$

provided that the integral exists. As a matter of fact, the mean value of  $\eta^*(u)$  is identical with the first semi-invariant.

The mean square deviation of  $\eta^*(u)$  is

$$(80) \quad D^2\{\eta^*(u)\} = \lambda \int_0^\infty \left[ \int_0^\infty (f(z, x))^2 dH(x) \right] dz,$$

provided that this integral exists. As a matter of fact, the mean square deviation of  $\eta^*(u)$ , namely,  $D^2[\eta^*(u)] = M[\eta^{*2}(u)] - M^2[\eta^*(u)]$  is identical with the second semi-invariant.

### § 8. The harmonic analysis of the stationary process $\eta^*(u)$

It can be shown that if the integral

$$(81) \quad \int_0^\infty \left[ \int_0^\infty f(z, x) dH(x) \right] dx$$

exists,  $\eta^*(u)$  defines a process which is stationary in the stricter sense. We shall, however, not require this but only that the process  $\eta^*(u)$  should be stationary in the wider sense, i. e. that the average and the dispersion of  $\eta^*(u)$  should be constant and that its auto-correlation function should not depend on  $u$ .



As has been shown in the preceding section,

$$(82) \quad M\{\eta_1^*(u)\} = \lambda \int_0^\infty \int_0^\infty f(z, x) dH(x) dz = m$$

and

$$(83) \quad D^2\{\eta_1^*(u)\} = \lambda \int_0^\infty \int_0^\infty (f(z, x))^2 dH(x) dz = \sigma^2,$$

if these integrals exist. These quantities are constants, independent of  $u$ .

We define the auto-correlation function of  $\eta_1^*(u)$  in the usual way:

$$(84) \quad R(\tau) = \frac{M\{\eta_1^*(u) \eta_1^*(u-\tau)\} - m^2}{\sigma^2}.$$

**THEOREM 4.** *In the case of the stationary stochastic process  $\eta_1^*(u)$  defined above, for the auto-correlation function defined by (84) we have*

$$(85) \quad R(\tau) = \frac{\lambda}{\sigma^2} \int_0^\infty \int_0^\infty f(u, x) f(u-\tau, x) dH(x) du$$

where  $R(\tau)$  is an even function of  $\tau$ .

**PROOF.** Let us consider the process  $\theta(u) = \eta_1^*(u) + \eta_1^*(u-\tau)$ . This process is again stationary and differs from the preceding one only in that it is instead of  $f(u, \chi)$  the expression  $f(u, \chi) + f(u-\tau, \chi)$  that supplies the progress in time of the signal which began at  $u=0$ . Therefore

$$(86) \quad \begin{aligned} D^2\{\theta(u)\} &= \lambda \int_0^\infty \int_0^\infty (f(u, x) + f(u-\tau, x))^2 dH(x) du = \\ &= 2\sigma^2 + 2\lambda \int_0^\infty \int_0^\infty f(u, x) f(u-\tau, x) dH(x) du. \end{aligned}$$

On the other hand, it is generally true that the dispersion of the sum of the random variables  $\eta_{11}$  and  $\eta_{12}$  is

$$(87) \quad D^2\{\eta_{11} + \eta_{12}\} = \sigma_1^2 + 2\sigma_1\sigma_2R + \sigma_2^2$$

where  $\sigma_1^2$  and  $\sigma_2^2$  are the dispersions of  $\eta_{11}$  and  $\eta_{12}$ , respectively, and  $R$  is the correlation coefficient of  $\eta_{11}$  and  $\eta_{12}$ . Putting  $\eta_{11} = \eta_1^*(u)$  and  $\eta_{12} = \eta_1^*(u-\tau)$ , we obtain

$$(88) \quad D^2\{\theta(u)\} = 2\sigma^2 + 2\sigma^2R(\tau).$$

From the comparison of (86) and (88) it results (85).

For stationary stochastic processes, A. JA. KHINTCHINE [9] using BOCHNER'S theorem proved that a necessary and sufficient condition for  $R(\tau)$  being the correlation function of a stationary process is that it should be of the form

$$(89) \quad R(\tau) = \int_{-\infty}^{\infty} \cos \tau \omega dF(\omega)$$



where  $F(\omega)$  is a probability distribution function.  $F(\omega)$  is called *the spectral function of the process*.

In physics, we mean by the spectral function of such a stochastic process the function

$$(90) \quad G(\omega) = m^2 + \sigma^2[F(\omega) - F(-\omega)],$$

the meaning of which is the following: if  $r_t^*(u)$  is considered to represent a current passed through a unit resistance, the average power dissipated is

$$(91) \quad \lim_{T \rightarrow \infty} M \left\{ \frac{1}{T} \int_0^T [r_t^*(u)]^2 du \right\} = M \{ [r_t^*(u)]^2 \} = m^2 + \sigma^2$$

and the distribution over the frequencies  $0 \leq \omega < \infty$  ( $\omega = 2\pi f$ ) of this power is supplied by  $G(\omega)$  which gives the power dissipated in the frequency band  $(0, \omega)$ .

Next we shall discuss the determination of  $F(\omega)$ .

Let us suppose that for  $p = 1, 2$

$$(92) \quad \int_0^{\infty} |f(u, x)|^p du < \infty$$

for all values of  $x$ . In this case, as well known,  $f(u, x)$  can be represented in the form of a Fourier integral

$$(93) \quad f(u, x) = \int_{-\infty}^{\infty} A(\omega, x) e^{i\omega u} d\omega$$

where

$$(94) \quad A(\omega, x) = \frac{1}{2\pi} \int_0^{\infty} f(u, x) e^{-i\omega u} du.$$

In physical literature it is  $2|A(\omega, x)|$  that is considered as the amplitude density function of the component possessing the frequency  $\omega$  in the spectral representation of  $f(u, x)$ .

Now, from PLANCHEREL'S theorem it follows

$$(95) \quad \int_0^{\infty} [f(u, x)]^2 du = 4\pi \int_0^{\infty} |A(\omega, x)|^2 d\omega.$$

If the square of  $f(u, x)$  is integrable, then  $f(u, x) + f(u - \tau, x)$  has the same property and in that case, if PLANCHEREL'S theorem is applied to it, it results in a simple manner that the following equation holds:

$$(96) \quad \int_0^{\infty} [f(u, x) + f(u - \tau, x)]^2 du = 4\pi \int_0^{\infty} |A(\omega, x)|^2 (2 + 2 \cos \omega \tau) d\omega.$$



Hence we conclude, taking (95) into consideration, that

$$(97) \quad \int_0^{\infty} f(u, x) f(u - \tau, x) du = 4\pi \int_0^{\infty} |A(\omega, x)|^2 \cos \omega \tau d\omega.$$

We take the Stieltjes integral of both sides with respect to  $H(x)$ :

$$(98) \quad \int_0^{\infty} \left[ \int_0^{\infty} f(u, x) f(u - \tau, x) du \right] dH(x) = 4\pi \int_0^{\infty} \left[ \int_0^{\infty} |A(\omega, x)|^2 \cos \omega \tau d\omega \right] dH(x).$$

If we suppose that the correlation function (84) exists for  $\tau = 0$ , these integrals will also exist, and it is easy to see that the order of integration may be reversed. Thus

$$(99) \quad \begin{aligned} R(\tau) &= \frac{\lambda}{\sigma^2} \int_0^{\infty} \left[ \int_0^{\infty} f(u, x) f(u - \tau, x) dH(x) \right] du = \\ &= \frac{4\pi\lambda}{\sigma^2} \int_0^{\infty} \left[ \int_0^{\infty} |A(\omega, x)|^2 dH(x) \right] \cos \omega \tau d\omega. \end{aligned}$$

This representation also shows that the function  $F(\omega)$  figuring in the Khintchine representation (89) is differentiable with respect to  $\omega$  and that  $F'(\omega)$  is an even function,

$$(100) \quad F'(\omega) = \frac{2\pi\lambda}{\sigma^2} \int_0^{\infty} |A(\omega, x)|^2 dH(x).$$

This representation is valid for all values of  $\omega$ , since  $|A(\omega, x)|$  is an even function of  $\omega$ .

We remark that, in accordance with a theorem on characteristic functions proved by G. PÓLYA [14], if  $R(\tau)$  is supposed to be convex for  $\tau > 0$ , it follows that the distribution function  $F(\omega)$  figuring in the Khintchine representation (89) is for  $\omega \neq 0$  differentiable in all cases and that its differential quotient is an even function of  $\omega$ .

The following theorem summarizes the above results.

**THEOREM 6.** *Consider the stochastic process  $\eta_1^*(u)$  defined above, and suppose that its correlation function (85) exists. Let us suppose further that  $|f(u, x)|$  and  $[f(u, x)]^2$  are integrable for the values of  $x$  in consideration; in this case the distribution function  $F(\omega)$  figuring in the Khintchine representation (89) of the correlation function  $R(\tau)$  is differentiable with respect to  $\omega$  and*

$$(101) \quad F'(\omega) = \frac{2\pi\lambda}{\sigma^2} \int_0^{\infty} |A(\omega, x)|^2 dH(x).$$

The above mentioned spectral distribution  $G(\omega)$  defined for  $0 \leq \omega < \infty$  is composed of a jump having the magnitude  $m^2$  at the point  $\omega = 0$  and



a continuous part whose density function is

$$(102) \quad G'(\omega) = 4\pi\lambda \int_0^{\infty} |A(\omega, x)|^2 dH(x).$$

REMARK. If  $f(u, x) = xf(u)$  and the Fourier transform of  $f(u)$  is

$$(103) \quad A(\omega) = \frac{1}{2\pi} \int_0^{\infty} f(u) e^{-i\omega u} du,$$

further if

$$(104) \quad M_2 = \int_0^{\infty} x^2 dH(x),$$

then it follows that  $A(\omega, x) = xA(\omega)$  and

$$(105) \quad G'(\omega) = 4\pi\lambda M_2 |A(\omega)|^2;$$

on the basis of (85) we have

$$(106) \quad R(\tau) = \frac{\lambda M_2}{\sigma^2} \int_0^{\infty} f(u) f(u-\tau) du$$

where

$$(107) \quad \sigma^2 = \lambda M_2 \int_0^{\infty} [f(u)]^2 du.$$

## § 9. Examples

In this chapter we shall discuss only the stationary process  $\eta^*(u)$  and we suppose  $f(u, \chi) = \chi e^{-\alpha u}$  where the distribution function of  $\chi$  is  $H(x)$ . We shall apply this process to the problems arising in the theory of particle counting effected with the aid of electron multipliers.

Let us denote by  $\lambda$  the temporal density of the corpuscles arriving at the cathode of the electron multiplier and by  $\chi$  the amplitudes of the signals produced by the individual particles. Now  $\alpha = 1/RC$  where  $RC$  is the time constant of the recording apparatus. If we denote by  $a$  the threshold voltage of the amplifier, then the apparent density of the events will be

$$(108) \quad \lambda'(a) = \lambda K^*(a).$$

$K^*(a)$  can be determined with the help of its Fourier transform

$$(109) \quad \Gamma^*(\omega) = \int_0^{\infty} e^{i\omega a} dK^*(a) = [1 - \varphi(\omega)] \Psi^*(\omega).$$

Here

$$(110) \quad \varphi(\omega) = \int_0^{\infty} e^{i\omega x} dH(x)$$



and

$$(111) \quad \Psi^*(\omega) = \int_0^{\infty} e^{i\omega x} dG^*(x) = \exp \left\{ -\frac{\lambda}{\alpha} \int_0^1 \frac{1 - \varphi(\omega v)}{v} dv \right\}.$$

Knowing  $G^*(x)$  and  $K^*(a)$ , the functions  $\tau^*(t, a)$ ,  $m^*(t, a)$  and  $\mathcal{G}^*(a)$  may easily be determined with the aid of the formulae (34), (40) and (43), respectively.

1. OBSERVATION PROBLEM. Let the random variable  $\chi$  possess an exponential distribution, i. e. let  $H(x) = 1 - e^{-\frac{x}{\mu}}$  with a positive constant  $\mu$ . Then

$$(112) \quad \varphi(\omega) = \frac{1}{1 - i\mu\omega}$$

whence

$$(113) \quad \Psi^*(\omega) = (1 - i\mu\omega)^{-\frac{\lambda}{\alpha}}$$

and

$$(114) \quad \Gamma(\omega) = (1 - i\mu\omega)^{-\frac{\lambda}{\alpha}} - (1 - i\mu\omega)^{-\frac{\lambda}{\alpha} - 1}.$$

Thus we have

$$(115) \quad G^*(x) = \int_0^x e^{-\frac{y}{\mu}} \frac{\left(\frac{y}{\mu}\right)^{\frac{\lambda}{\alpha} - 1}}{\Gamma\left(\frac{\lambda}{\alpha}\right)} \frac{dy}{\mu}$$

and

$$(116) \quad K^*(a) = \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}}}{\Gamma\left(1 + \frac{\lambda}{\alpha}\right)}$$

whence

$$(117) \quad m^*(t, a) = \lambda t \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}}}{\Gamma\left(1 + \frac{\lambda}{\alpha}\right)},$$

$$(118) \quad \tau^*(t, a) = \frac{t}{\Gamma\left(\frac{\lambda}{\alpha}\right)} \int_0^{\infty} e^{-y} y^{\frac{\lambda}{\alpha} - 1} dy$$

and

$$(119) \quad \mathcal{G}(a) = \alpha \frac{\int_0^{\infty} e^{-y} y^{\frac{\lambda}{\alpha} - 1} dy}{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}}}.$$



For the quotient of the incomplete and complete gamma functions figuring here, tables can be found in K. PEARSON'S book [13].

Now the apparent density of events in case the threshold voltage is  $a$  will be:

$$(120) \quad \lambda'(a) = \lambda \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}}}{\Gamma\left(1 + \frac{\lambda}{\alpha}\right)}.$$

Here  $\mu$  denotes the mean value of the amplitudes of the signals.

The maximum of the apparent density of events  $\lambda'(a)$  is situated at  $a_{\max} = \frac{\lambda\mu}{\alpha}$ . Let  $\lambda/\alpha = 0.25$  (for instance  $\lambda = 10^5 \text{ sec}^{-1}$  and  $\alpha = 4 \cdot 10^5 \text{ sec}$ , i. e.  $RC = 2.5 \cdot 10^{-6} \text{ sec}$ ) and let us introduce  $a = z\mu$ . We can construct the following Table 1 showing the quotient  $\lambda'/\lambda$  as a function of  $z$ .

TABLE 1

$z$	$\lambda'/\lambda$	$z$	$\lambda'/\lambda$	$z$	$\lambda'/\lambda$	$z$	$\lambda'/\lambda$
0,00	0,0000	0,50	0,5627	1,00	0,4059	1,50	0,2724
0,05	0,4963	0,55	0,5482	1,05	0,3908	1,55	0,2613
0,10	0,5614	0,60	0,5329	1,10	0,3761	1,60	0,2505
0,15	0,5910	0,65	0,5171	1,15	0,3618	1,65	0,2401
0,20	0,6041	0,70	0,5019	1,20	0,3478	1,70	0,2301
0,25	0,6076	0,75	0,4850	1,25	0,3342	1,75	0,2205
0,30	0,6049	0,80	0,4688	1,30	0,3211	1,80	0,2112
0,35	0,5980	0,85	0,4528	1,35	0,3083	1,85	0,2023
0,40	0,5881	0,90	0,4369	1,40	0,2959	1,90	0,1937
0,45	0,5762	0,95	0,4212	1,45	0,2840	1,95	0,1855

2. PROBLEM OF COINCIDENCE.

a) *Chance coincidences.* Let the distribution function of  $\chi$  again be

$H(x) = 1 - e^{-\frac{x}{\mu}}$ . Let us now count two corpuscle processes arriving with a density  $\lambda$ , by means of two similar electron multipliers. Then, with a threshold voltage of  $a$ , the apparent density of coincidences results from (120) by replacing  $\lambda$  in this formula by  $2\lambda$ :

$$(121) \quad \lambda'_c(a) = 2\lambda e^{-\frac{a}{\mu}} \frac{\left(\frac{a}{\mu}\right)^{\frac{2\lambda}{\alpha}}}{\Gamma\left(1 + \frac{2\lambda}{\alpha}\right)}.$$

The maximum of this is situated at  $a_{\max} = \frac{2\lambda\mu}{\alpha}$ . Let again  $\lambda/\alpha = 0,25$  as in



the preceding example, and let  $z = a/\mu$ . In this case Table 2 shows the quotient  $\lambda'_k/\lambda$  as a function of  $z$ .

TABLE 2

$z$	$\lambda'_k/\lambda$	$z$	$\lambda'_k/\lambda$	$z$	$\lambda'_k/\lambda$	$z$	$\lambda'_k/\lambda$	$z$	$\lambda'_k/\lambda$
0,000	0,0000	0,05	0,4800	0,55	0,9656	1,05	0,8092	1,55	0,5963
0,001	0,0713	0,10	0,6457	0,60	0,9594	1,10	0,7879	1,60	0,5763
0,002	0,1007	0,15	0,7523	0,65	0,9498	1,15	0,7663	1,65	0,5567
0,003	0,1232	0,20	0,8263	0,70	0,9376	1,20	0,7446	1,70	0,5375
0,004	0,1422	0,25	0,8788	0,75	0,9232	1,25	0,7229	1,75	0,5188
0,005	0,1588	0,30	0,9157	0,80	0,9070	1,30	0,7013	1,80	0,5005
0,010	0,2234	0,35	0,9408	0,85	0,8893	1,35	0,6798	1,85	0,4826
0,020	0,3128	0,40	0,9567	0,90	0,8704	1,40	0,6585	1,90	0,4653
0,030	0,3793	0,45	0,9653	0,95	0,8507	1,45	0,6374	1,95	0,4484
0,040	0,4337	0,50	0,9679	1,00	0,8302	1,50	0,6176	2,00	0,4319

b) *Artificial coincidences.* The preceding result was obtained under the hypothesis that the corpuscles arriving at the cathodes of the electron multipliers are mutually independent. Now let us suppose that a complete connection exists between the two processes; i. e. the corpuscles incident on the cathodes of both electron multipliers are arriving exactly simultaneously. The question arises in this case of which value the density of apparent coincidences will be. These two processes can be reduced to a single process, in which the density of events is  $\lambda$  and the distribution function  $H(x)$  of the amplitudes of the signals represent the composition of the former exponential distribution with itself, i. e.

$$(122) \quad H(x) = \int_0^x e^{-\frac{u}{\mu}} \frac{u}{\mu} \frac{du}{\mu} = 1 - \left(1 + \frac{x}{\mu}\right) e^{-\frac{x}{\mu}}.$$

Now

$$(123) \quad \varphi(\omega) = (1 - i\mu\omega)^{-2}$$

and

$$(124) \quad \log \Psi^*(\omega) = -\frac{\lambda}{\alpha} \int_0^1 \frac{1 - \varphi(\omega v)}{v} dv = \frac{\lambda}{\alpha} \left[ \frac{i\mu\omega}{1 - i\mu\omega} - \log(1 - i\mu\omega) \right],$$

i. e.

$$(125) \quad \Psi^*(\omega) = (1 - i\mu\omega)^{-\frac{\lambda}{\alpha}} e^{\frac{\lambda}{\alpha} \frac{i\mu\omega}{1 - i\mu\omega}} = e^{-\frac{\lambda}{\alpha} \sum_{j=0}^{\infty} \frac{\left(\frac{\lambda}{\alpha}\right)^j}{j!} \left(\frac{1}{1 - i\mu\omega}\right)^{\frac{\lambda}{\alpha} + j}}$$

and

$$(126) \quad \Gamma(\omega) = \frac{\mu\omega}{i} \sum_{j=0}^{\infty} \frac{e^{-\frac{\lambda}{\alpha} \left(\frac{\lambda}{\alpha}\right)^j}}{j!} \left[ \left(\frac{1}{1 - i\mu\omega}\right)^{\frac{\lambda}{\alpha} + j + 1} + \left(\frac{1}{1 - i\mu\omega}\right)^{\frac{\lambda}{\alpha} + j + 2} \right]$$



whence

$$(127) \quad G^*(x) = \sum_{j=0}^{\infty} \frac{e^{-\frac{\lambda}{\alpha}} \left(\frac{\lambda}{\alpha}\right)^j}{j!} \int_0^x e^{-\frac{y}{\mu}} \frac{\left(\frac{y}{\mu}\right)^{\frac{\lambda}{\alpha}+j-1}}{\Gamma\left(j+\frac{\lambda}{\alpha}\right)} \frac{dy}{\mu}$$

and

$$(128) \quad K^*(a) = \sum_{j=0}^{\infty} \frac{e^{-\frac{\lambda}{\alpha}} \left(\frac{\lambda}{\alpha}\right)^j}{j!} \left[ \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}+j}}{\Gamma\left(\frac{\lambda}{\alpha}+1+j\right)} + \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}+j+1}}{\Gamma\left(\frac{\lambda}{\alpha}+2+j\right)} \right].$$

In view of the definition of the Bessel function

$$(129) \quad J_{\rho}(ix) = \left(\frac{ix}{2}\right)^{\rho} \sum_{\nu=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2\nu}}{\nu! \Gamma(\nu+\rho+1)},$$

we may write

$$(130) \quad K^*(a) = e^{-\left(\frac{\lambda}{\alpha}+\frac{a}{\mu}\right)} \left(\frac{a}{\mu}\right)^{\frac{\lambda}{\alpha}} \left[ \frac{J_{\frac{\lambda}{\alpha}}\left(2i\sqrt{\frac{\lambda a}{\alpha\mu}}\right)}{\left(i\sqrt{\frac{\lambda a}{\alpha\mu}}\right)^{\frac{\lambda}{\alpha}}} + \frac{a}{\mu} \frac{J_{\frac{\lambda}{\alpha}+1}\left(2i\sqrt{\frac{\lambda a}{\alpha\mu}}\right)}{\left(i\sqrt{\frac{\lambda a}{\alpha\mu}}\right)^{\frac{\lambda}{\alpha}+1}} \right].$$

If  $\lambda/\alpha = n$  is an integer, the calculation of  $K^*(a)$  can be reduced to the determination of the distribution function of the difference of independent random variables representing a Poisson distribution, namely,

$$(131) \quad K^*(a) = P(\xi_1 - \xi_2 = n) + P(\xi_1 - \xi_2 = n + 1)$$

where

$$(132) \quad P(\xi_1 = k) = \frac{e^{-\frac{a}{\mu}} \left(\frac{a}{\mu}\right)^k}{k!} \quad (k = 0, 1, 2, \dots)$$

and

$$(133) \quad P(\xi_2 = k) = \frac{e^{-n} n^k}{k!} \quad (k = 0, 1, 2, \dots).$$

For Poisson distributions a good table can be found in E. C. MOLINA's book [12].

In this case the apparent density of the coincidences will be

$$(134) \quad \lambda'_k(a) = \lambda K^*(a)$$

where  $K^*(a)$  is given by (130). Let again, as formerly,  $\lambda/\alpha = 0,25$  and  $z = a/\mu$ ; then we have the following table:



TABLE 3

$z$	$\lambda'_k/\lambda$	$z$	$\lambda''_k/\lambda$	$z$	$\lambda'_k/\lambda$	$z$	$\lambda''_k/\lambda$	$z$	$\lambda''_k/\lambda$
0,000	0,0000	0,05	0,4059	0,55	0,6749	1,05	0,6588	1,55	0,5707
0,001	0,1528	0,10	0,4814	0,60	0,6793	1,10	0,6520	1,60	0,5603
0,002	0,1817	0,15	0,5303	0,65	0,6820	1,15	0,6446	1,65	0,5497
0,003	0,2011	0,20	0,5664	0,70	0,6831	1,20	0,6367	1,70	0,5391
0,004	0,2161	0,25	0,5944	0,75	0,6828	1,25	0,6283	1,75	0,5284
0,005	0,2285	0,30	0,6167	0,80	0,6813	1,30	0,6194	1,80	0,5176
0,010	0,2717	0,35	0,6345	0,85	0,6786	1,35	0,6102	1,85	0,5068
0,020	0,3231	0,40	0,6487	0,90	0,6750	1,40	0,6007	1,90	0,4960
0,030	0,3575	0,45	0,6599	0,95	0,6704	1,45	0,5909	1,95	0,4852
0,040	0,3840	0,50	0,6685	1,00	0,6650	1,50	0,5809	2,00	0,4744

Fig. 1 shows the values of  $\lambda'_k/\lambda$ ,  $\lambda''_k/\lambda$  and  $(\lambda'_k - \lambda''_k)/\lambda$  as functions of  $z = a/\mu$ . It is desired to investigate the question whether or not there is any connection between the times of arrival of the corpuscles incident on the cathodes of two electron multipliers. It has been shown, that, in the case of independence, the apparent density of chance coincidences is  $\lambda'_k(a)$ ; and in the case of complete connection, however,  $\lambda''_k(a)$ . It will be advisable to select  $a$  in such a manner as to ensure that  $|\lambda''_k(a) - \lambda'_k(a)|$  should be a maximum.

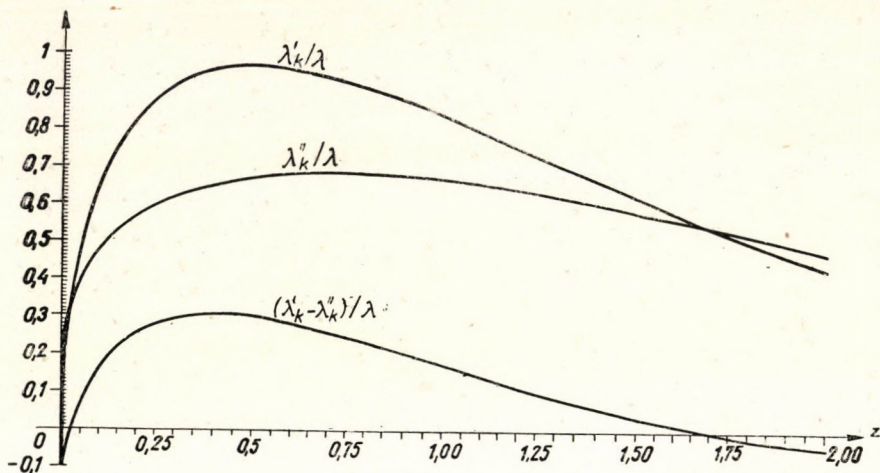


Fig. 1

3. PROBLEM OF OBSERVATION. Let  $H(x)$  be a gamma distribution belonging to the index  $n$ , that is, equal to the  $n$ -fold composition of the exponential distribution  $1 - e^{-x/\mu}$  with itself,

$$(135) \quad H(x) = \int_0^x e^{-\frac{y}{\mu}} \frac{\left(\frac{y}{\mu}\right)^{n-1}}{(n-1)!} \frac{dy}{\mu} = 1 - \sum_{j=0}^{n-1} e^{-\frac{x}{\mu}} \frac{\left(\frac{x}{\mu}\right)^j}{j!}.$$



We conclude

$$(136) \quad \varphi(\omega) = (1 - i\mu\omega)^{\lambda/\alpha}$$

and

$$(137) \quad \begin{aligned} \log \Psi^*(\omega) &= -\frac{\lambda}{\alpha} \int_0^1 \frac{1 - \varphi(v\omega)}{v} dv = \\ &= -\frac{\lambda}{\alpha} \left[ \log(1 - i\mu\omega) + \sum_{j=1}^{n-1} \left( 1 - \frac{1}{(1 - i\mu\omega)^j} \right) \right], \end{aligned}$$

$$(138) \quad \Psi^*(\omega) = \frac{e^{-\frac{\lambda}{\alpha}(n-1)}}{(1 - i\mu\omega)^{\frac{\lambda}{\alpha}}} e^{\frac{\lambda}{\alpha} \sum_{j=1}^{n-1} \left( \frac{1}{1 - i\mu\omega} \right)^j} = e^{-\frac{\lambda}{\alpha}(n-1)} \sum_{l=0}^{\infty} C_l \left( \frac{1}{1 - i\mu\omega} \right)^{l + \frac{\lambda}{\alpha}}$$

where

$$(139) \quad C_l = \sum_{\nu_1 + 2\nu_2 + \dots + (n+1)\nu_{n-1} = l} \frac{1}{\nu_1! \nu_2! \dots \nu_{n-1}!} \left( \frac{\lambda}{\alpha} \right)^{\nu_1 + \nu_2 + \dots + \nu_{n-1}}$$

In this case

$$(140) \quad G^*(x) = e^{-\frac{\lambda}{\alpha}(n-1)} \sum_{l=0}^{\infty} C_l \int_0^x e^{-\frac{y}{\mu}} \frac{\left( \frac{y}{\mu} \right)^{\frac{\lambda}{\alpha} + l + 1}}{\Gamma\left( l + \frac{\lambda}{\alpha} \right)} \frac{dy}{\mu},$$

moreover

$$(141) \quad \begin{aligned} \Gamma(\omega) &= [1 - \varphi(\omega)] \Psi^*(\omega) = \frac{\omega\mu}{i} \left[ \frac{1}{(1 - i\mu\omega)} + \dots + \frac{1}{(1 - i\mu\omega)^n} \right] \Psi^*(\omega) = \\ &= \frac{\omega\mu}{i} e^{-\frac{\lambda}{\alpha}(n-1)} \sum_{r=1}^{\infty} B_r \left( \frac{1}{1 - i\mu\omega} \right)^{r + \frac{\lambda}{\alpha}}, \end{aligned}$$

with

$$(142) \quad B_r = \sum_{l=r-n}^{r-1} C_l.$$

Here  $C_l$  is defined for  $l \geq 0$  by (139), while for  $l < 0$  we put  $C_l = 0$ . Thus we obtain

$$(143) \quad K^*(a) = e^{-\frac{\lambda}{\alpha}(n-1)} \sum_{r=1}^{\infty} B_r \frac{e^{-\frac{a}{\mu}} \left( \frac{a}{\mu} \right)^{\frac{\lambda}{\alpha} + r + 1}}{\Gamma\left( \frac{\lambda}{\alpha} + r \right)}.$$

In the calculation of  $G^*(x)$  and  $K^*(a)$  a good assistance is given by the table of PEARSON [13]. In case  $\lambda/\alpha$  is an integer, the calculation of these series can be carried out with the aid of a table concerning Poisson distributions.



For the coincidence problem discussed in connection with the preceding example it is possible in the case of a complete connection (artificial coincidences) to utilize the formulae as above, as the composition of two gamma distributions is again a gamma distribution.

4. THE HARMONIC ANALYSIS OF THE PROCESS  $\eta^*(u)$ . Let the distribution function  $H(x)$  be arbitrary, but let its first and second moments  $M_1$  and  $M_2$  be finite. If  $f(u, x) = xe^{-\alpha u}$  when  $u \geq 0$  and otherwise  $f(u, x) = 0$ , we have on account of (85)

$$(144) \quad R(\tau) = e^{-\alpha|\tau|},$$

independently of the distribution function  $H(x)$ .

The spectral distribution  $G(\omega)$  possesses at the point  $\omega = 0$  a jump of size  $m^2 = M_1^2/\alpha^2$ , is differentiable for  $0 < \omega < \infty$  and its differential quotient is

$$(145) \quad G'(\omega) = \frac{2\lambda M_2}{\alpha^2 + \omega^2},$$

since

$$(146) \quad A(\omega) = \frac{1}{2\pi} \int_0^{\infty} e^{-\alpha u - i\omega u} du = \frac{1}{2\pi} \frac{1}{\alpha + i\omega}.$$

### Appendix\*

Let us consider the random function

$$\eta(u) = \sum_{0 \leq u_k < u} f(u - u_k, \chi_k)$$

defined earlier by (3). In what follows we give a new proof of the determination of the distribution function  $G(u, x) = P(\eta(u) \leq x)$ . We use only one elementary property of the Poisson processes.

Concerning the underlying Poisson process we suppose only, more generally than before, that  $A(u)$ , the expected number of events occurring in the time interval  $(0, u)$ , is a non-decreasing function of  $u$ . Then the probability of exactly  $n$  events occurring in the time interval  $(0, u)$  will be given by

$$P(\eta_u = n) = \frac{e^{-A(u)} [A(u)]^n}{n!} \quad (n = 0, 1, 2, \dots)$$

where the random variable  $\eta_u$  denotes the number of events occurring in the time interval  $(0, u)$ .

\* Received 15 February 1953.



Next we shall determine

$$\Psi(u, \omega) = \int_{-\infty}^{\infty} e^{i\omega x} d_u G(u, x),$$

the characteristic function of the variable  $\eta(u)$ . Then the distribution function  $G(u, x)$  can be determined uniquely by its characteristic function.

**THEOREM.** *If  $f(u, x)$  is a real valued function, measurable in both variables, and*

$$\varphi(z, \omega) = \int_{-\infty}^{\infty} e^{i\omega f(z, x)} dH(x),$$

then

$$\Psi(u, \omega) = \exp \left\{ - \int_0^u [1 - \varphi(u - z, \omega)] dA(z) \right\}.$$

**PROOF.** Our proof is based on the following

**LEMMA.** *Suppose in the Poisson process mentioned above there occur exactly  $n$  events in the time interval  $(0, u)$ ; then the moments of occurrence of the  $n$  events may be considered as a distribution of  $n$  independent random points in the time interval  $(0, u)$  where the probability that a point lies in the interval  $(0, x)$  ( $0 \leq x \leq u$ ) is  $A(x)/A(u)$ .*

Let the random variable  $\eta_u$  be the number of the events in the Poisson process occurring in the time interval  $(0, u)$  and  $0 \leq \xi_1 \leq \xi_2 \leq \dots \leq \xi_n \leq \dots$  the moments of the consecutive events. Then it is sufficient to show for the values  $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq \dots$  the validity of the following formula:

$$\begin{aligned} & P(\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n | \eta_u = n) = \\ &= \sum_{j_1, j_2, \dots, j_n} \frac{n!}{j_1! j_2! \dots j_n!} \left[ \frac{A(x_1)}{A(u)} \right]^{j_1} \left[ \frac{A(x_2) - A(x_1)}{A(u)} \right]^{j_2} \dots \left[ \frac{A(x_n) - A(x_{n-1})}{A(u)} \right]^{j_n} \end{aligned}$$

where the summation is extended over those values of  $j$  for which  $j_1 \geq 1, j_1 + j_2 \geq 2, \dots, j_1 + j_2 + \dots + j_{n-1} \geq n - 1$  and  $j_1 + j_2 + \dots + j_n = n$ . Indeed, the right-hand side gives the joint distribution function of the coordinates rearranged in increasing order of magnitude of  $n$  independent points distributed in the time interval  $(0, u)$  according to the mentioned law, i. e. the distribution function of the  $n$  random variables  $(\xi_1, \xi_2, \dots, \xi_n)$ .

By the definition of the conditional probability we obtain

$$P(\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n | \eta_u = n) = \frac{P(\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n; \eta_u = n)}{P(\eta_u = n)}.$$

Now

$$\begin{aligned} & P(\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_n \leq x_n; \eta_u = n) = \\ &= e^{-A(u)} \sum_{j_1, j_2, \dots, j_n} \frac{[A(x_1)]^{j_1}}{j_1!} \frac{[A(x_2) - A(x_1)]^{j_2}}{j_2!} \dots \frac{[A(x_n) - A(x_{n-1})]^{j_n}}{j_n!} \end{aligned}$$



where the summation must be extended over the values  $j_1 \geq 1$ ,  $j_1 + j_2 \geq 2, \dots$ ,  $j_1 + j_2 + \dots + j_{n-1} \geq n-1$  and  $j_1 + j_2 + \dots + j_n = n$ . The right-hand side of this equation agrees with the probability that in the Poisson process  $j_1, j_2, \dots, j_n, 0$  events occur in the intervals  $(0, x_1), (x_1, x_2), \dots, (x_{n-1}, x_n), (x_n, u)$  where the admissible values of  $j$  are defined by the preceding inequalities.

Furthermore we have

$$P(\eta_u = n) = e^{-\Lambda(u)} \frac{[\Lambda(u)]^n}{n!},$$

and thus our lemma is proved.

Now we begin the proof the theorem. First we determine the characteristic function of the value of the amplitude of a random signal in the moment  $u$  (the signal started from a random point in the interval  $(0, u)$  having the distribution law  $A(x)/A(u)$ ). If we know the signal started in the moment  $z$ , then, by this condition, we have the characteristic function

$$\varphi(u-z, \omega) = \int_{-\infty}^{\infty} e^{i\omega f(u-z, x)} dH(x).$$

Since the distribution function of  $z$  is  $A(z)/A(u)$ , therefore, using the theorem on conditional expected values, we get for the characteristic function of the  $u$ -moment value of the amplitude of a signal started by a random point

$$\psi_u(\omega) = \frac{1}{A(u)} \int_0^u \varphi(u-z, \omega) dA(z).$$

If  $n$  events occur in the interval  $(0, u)$ , then, by our lemma, the sum of the values of the signals in the moment  $u$  equals the sums of  $n$  independent random variables, each of them having the characteristic function  $\psi_u(\omega)$ . Therefore, the characteristic function of the sum is  $[\psi_u(\omega)]^n$ .

The probability of the occurrence of  $n$  events is given by formula  $P(\eta_u = n)$ ; consequently,

$$\Psi(u, \omega) = \sum_{n=0}^{\infty} e^{-\Lambda(u)} \frac{[\Lambda(u)]^n}{n!} [\psi_u(\omega)]^n = e^{-\Lambda(u)[1 - \psi_u(\omega)]}.$$

This completes the proof.

Finally, I would express my best thanks to Prof. A. RÉNYI for his valuable remarks.



## Bibliography

- [1] N. CAMPBELL, The study of discontinuous phenomena, *Proceedings of the Cambridge Philosophical Society*, **15** (1909), pp. 117—137.
- [2] N. CAMPBELL, Discontinuities in light emission, *Proc. Camb. Phil. Soc.*, **15** (1909), pp. 310—328.
- [3] H. CRAMÉR, *Mathematical methods of statistics* (Princeton, 1946).
- [4] W. FELLER, On the integro-differential equations of purely discontinuous Markov processes, *Trans. Amer. Math. Soc.*, **48** (1940), pp. 488—515.
- [5] W. FELLER, Zur Theorie der stochastischen Prozesse, *Math. Annalen*, **113** (1936), pp. 113—160.
- [6] H. HURWITZ and M. KAC, Statistical analysis of certain types of random functions, *Annals of Math. Stat.*, **15** (1944), pp. 173—181.
- [7] A. KHINTCHINE, Theorie des abklingenden Spontaneffektes, *Bull. Acad. Sci URSS, Ser. Math.*, **3** (1938), pp. 312—322.
- [8] A. KHINTCHINE, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung* (Berlin, 1933), p. 23.
- [9] A. KHINTCHINE, Korrelationstheorie der stationären stochastischen Prozesse, *Math. Annalen*, **109** (1934), pp. 604—615.
- [10] A. N. KOLMOGOROV, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Annalen*, **104** (1931), pp. 415—458.
- [11] А. Н. КОЛМОГОРОВ и Ю. В. ПРОХОРОВ, О суммах случайного числа случайных слагаемых, *Усп. Матем. Наук*, **4** (1949), pp. 168—172.
- [12] E. C. MOLINA, *Poisson's exponential binomial limit* (New York, 1945).
- [13] K. PEARSON, *Tables of incomplete  $\Gamma$ -function* (London, 1922).
- [14] G. PÓLYA, Remarks on characteristic functions, *Proc. Berkeley Symposium*, 1949, pp. 115—123.
- [15] A. RÉNYI, On some problems concerning Poisson processes, *Publicationes Mathematicae (Debrecen)*, **2** (1951), pp. 66—73.
- [16] S. O. RICE, Mathematical analysis of random noise I, *Bell. System Technical Journal*, **23** (1944), pp. 282—332; II, **24** (1945), pp. 46—156.
- [17] E. N. ROWLAND, The theory of mean square variation of a function formed by adding known functions with random phases and applications to the theories of shot effect and of light, *Proc. Camb. Phil. Soc.*, **32** (1936), pp. 580—597.



## О СТОХАСТИЧЕСКИХ ПРОЦЕССАХ, ПОРОЖДАЕМЫХ ПРОЦЕССОМ ПУАССОНА, И ОБ ИХ ПРИМЕНЕНИЯХ В ФИЗИКЕ

Л. ТАКАЧ (Будапешт)

(Резюме)

В настоящей работе рассматриваются случайные величины типов

$$\eta(t) = \sum_{0 < t_n \leq t} f(t - t_n, \chi_n)$$

и

$$\eta^*(t) = \sum_{-\infty < t_n \leq t} f(t - t_n, \chi_n),$$

где  $f(u, x)$  есть некоторая функция двух переменных, равная нулю для отрицательных значений  $u$ , моменты времени  $\{t_n\}$  являются моментами наступления событий в процессе Пуассона с плотностью событий  $\lambda(u)$  ( $-\infty < u < \infty$ ), параметры же  $\{\chi_n\}$  суть одинаково распределенные случайные величины.

Процессы такого рода встречаются в физике в теории счетчиков и при исследовании колебаний анодного тока электронных ламп.

В первой части работы исследуется специальный случай  $f(u, x) = xe^{-ax}$ , играющий важную роль в теории счетчиков. В этом случае стохастический процесс  $\eta(t)$  есть т. н. дискретный марковский процесс, допускающий, по теории В. Феллера, исследование с помощью интегро-дифференциального уравнения.

Во второй части работы ведется исследование случая произвольной функции  $f(u, x)$ . Этот случай имеет важность в связи с проблемой описания колебаний анодного тока электронных ламп. В этом общем случае процесс  $\eta(t)$  уже не является процессом марковского типа, однако можно из него сделать аддитивный марковский процесс, производя исследование при обратно направленной оси  $t$ . Этим приемом исследование процесса  $\eta(t)$  по существу приводится к суммированию независимых случайных величин.

В работе для обоих случаев исследованы распределения и предельные распределения случайных величин  $\eta(t)$  и  $\eta^*(t)$ ; вычисляются также некоторые вероятности перехода. Для однородного случая  $\lambda(u) \equiv \lambda$  приводится гармонический анализ процесса  $\eta^*(t)$ .

Полученные результаты применяются к решению проблем, связанных с совпадениями при употреблении электронного умножителя в качестве счетчика.

Наконец, в приложении излагается метод, дающий решение проблемы о распределении  $\eta(t)$  применением простой теоремы о процессах Пуассона.



# ÜBER DIE CESÀROSCHES SUMMIERBARKEIT DER ORTHOGONALEN POLYNOMREIHEN. II.

Von

KÁROLY TANDORI (Szeged)

(Vorgelegt von G. ALEXITS)

## Einleitung

In einer früheren Arbeit<sup>1</sup> habe ich mich mit der  $(C, 1)$ - und  $H_2$ -Summierbarkeit<sup>2</sup> der Entwicklungen von  $L^2$ -integrierbaren Funktionen nach orthogonalen Polynomen beschäftigt. In dieser Arbeit gebe ich gewisse Verallgemeinerungen meiner damaligen Resultate, indem ich nun für orthogonale Polynomreihen das Analogon des allgemeinen starken Summationssatzes von G. H. HARDY und J. E. LITTLEWOOD<sup>3</sup> beweise. Unter gewissen Annahmen für das orthogonale Polynomsystem gelingt es dadurch die  $H_\sigma$ - und  $(C, r > 0)$ -Summierbarkeit der Entwicklungen  $L^p$  ( $p > 1$ )-integrierbarer Funktionen nachzuweisen. Ich beweise weiterhin auch einen Satz bezüglich der  $(C, r > 0)$ -Summierbarkeit von orthogonalen Polynomreihen  $L$ -integrierbarer Funktionen. In Gegensatz zu der vorangehenden Arbeit betrachte ich jetzt orthonormierte Polynomsysteme relativ zu einer allgemeinen Verteilung  $d\alpha(x)$ ; ist  $\alpha(x)$  absolut stetig, so ist das Polynomsystem bezüglich der Gewichtsfunktion  $w(x) = \alpha'(x)$  orthonormiert.

<sup>1</sup> K. TANDORI, Über die Cesàrosche Summierbarkeit der orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, 3 (1952), S. 73–82.

<sup>2</sup> Eine Reihe  $\sum_{\nu=0}^{\infty} u_\nu$  heißt  $H_\sigma$ -summierbar, oder stark summierbar  $\sigma$ -ter Ordnung zu dem Werte  $S$ , wenn

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\nu=0}^n |s_\nu - S|^\sigma = 0$$

ist, wo  $s_\nu = u_0 + u_1 + \dots + u_\nu$  die  $\nu$ -te Partialsumme der Reihe bezeichnet.

<sup>3</sup> G. H. HARDY—J. E. LITTLEWOOD, Notes on the theory of series (IV): On the strong summability of Fourier-series, *Proceedings of the London Math. Society*, 26 (1927), S. 273–286.



Bekanntlich ist die Frage der  $(C, r > 0)$ -Summation der allgemeinen orthogonalen Polynomreihen durch jenen Umstand erschwert, daß keine, der Christoffel-Darboux'schen Formel ähnliche, geschlossene Darstellung für die Kernfunktionen der  $(C, r > 0)$ -Summation bekannt ist. Das Interessante in der nachfolgenden Behandlung besteht darin, daß man bloß aus der Struktur der Kernfunktion der Partialsummen auf die starke Summierbarkeit, und daraus auf die  $(C, r > 0)$ -Summierbarkeit schließen kann.

### § 1. Bezeichnungen

Es sei  $\alpha(x)$  eine im endlichen Grundintervall  $[a, b]$  definierte, nicht abnehmende, beschränkte Funktion mit unendlichvielen Zuwachspunkten. Bezeichne  $\{p_n(x)\}$  das zur Verteilung  $d\alpha(x)$  gehörende orthonormierte Polynomsystem, wo  $p_n(x)$  ein Polynom genau  $n$ -ten Grades mit positivem Hauptkoeffizient ist; für diese Polynome besteht also die Relation

$$\int_a^b p_m(x)p_n(x)d\alpha(x) = \begin{cases} 0 & (m \neq n), \\ 1 & (m = n), \end{cases}$$

wo das Integral im Lebesgue-Stieltjes'schen Sinne zu verstehen ist. Es wird im folgenden die Menge der Funktionen, für welche

$$\int_a^b |f(x)|^p d\alpha(x) < \infty$$

ist, mit  $L_\alpha^p[a, b]$ , oder wenn kein Mißverständnis zu befürchten ist, mit  $L_\alpha^p$  bezeichnet.

Ist  $f(x) \in L_\alpha[a, b]$ , so existiert die Entwicklung von  $f(x)$  nach dem System  $\{p_n(x)\}$ :

$$f(x) \sim \sum_{v=0}^{\infty} c_v p_v(x),$$

wo

$$c_v = \int_a^b f(t)p_v(t)d\alpha(t) \quad (v=0, 1, \dots)$$

den  $v$ -ten verallgemeinerten Fourierkoeffizient von  $f(x)$  bedeutet. Wir bezeichnen die  $n$ -te Partialsumme dieser Entwicklung mit  $s_n(f; x)$  und ihr  $n$ -tes  $(C, r)$ -Mittel mit  $\sigma_n^r(f; x)$ :

$$s_n(f; x) = \sum_{v=0}^n c_v p_v(x), \quad \sigma_n^r(f; x) = \frac{1}{A_n^r} \sum_{v=0}^n A_{n-r}^r c_v p_v(x),$$

wo

$$(1) \quad A_n^r = \binom{n+r}{n} \sim n^r \quad (r \neq -1, -2, \dots)$$



ist. Es gelten die nachstehenden Formeln:

$$(2) \quad s_\nu(f; x) - f(x) = \int_a^b [f(t) - f(x)] K_\nu(x, t) d\alpha(t),$$

$$(3) \quad \sigma_n^r(f; x) - f(x) = \frac{1}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{s_\nu(f; x) - f(x)\},$$

$$(4) \quad \sigma_n^r(f; x) - f(x) = \int_a^b [f(t) - f(x)] K_n^r(x, t) d\alpha(t),$$

wo

$$(5) \quad K_\nu(x, t) = \sum_{k=0}^{\nu} p_k(x) p_k(t)$$

die Kernfunktion der  $\nu$ -ten Partialsumme und

$$(6) \quad K_n^r(x, t) = \frac{1}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} K_\nu(x, t)$$

die zu der  $(C, r)$ -Summation gehörende  $n$ -te Kernfunktion ist. Das Integral

$$L_n^{(r)}(x) = \int_a^b |K_n^r(x, t)| d\alpha(t)$$

heißt die  $n$ -te Lebesguesche Funktion der  $(C, r)$ -Summation.

Im folgenden werden wir die bekannte Christoffel-Darboux'sche Formel

$$K_\nu(x, t) = \gamma_\nu \frac{p_{\nu+1}(t)p_\nu(x) - p_\nu(t)p_{\nu+1}(x)}{t-x} \quad (\gamma_\nu = O(1))$$

verwenden.

### § 2. Hilfssätze

Aus  $f(x) \in L^p[a, b]$  folgt bekanntlich, daß

$$\int_0^h |f(x \pm v) - f(x)|^p dx = o(h)^4$$

in  $[a, b]$  fast überall gilt. Eine ähnliche Behauptung gilt auch im Fall eines Lebesgue-Stieltjes'schen Integrals.

HILFSSATZ 1. *Es sei  $E \subseteq [a, b]$  die Menge der Punkte  $x$ , in welchen  $\alpha'(x)$  existiert und endlich ist. Ist  $f(x) \in L_a^p[a, b]$  ( $p \geq 1$ ), so gilt auf der Menge  $E$   $\alpha$ -fast überall<sup>5</sup>*

$$(7) \quad \int_0^h |f(x \pm v) - f(x)|^p d\alpha(x \pm v) = o(h).$$

<sup>4</sup> S. z. B. A. ZYGMUND, *Trigonometrical series* (Warszawa—Lwów, 1935), S. 237—238.

<sup>5</sup> Die Ausdrücke „ $\alpha$ -fast überall“, „vom  $\alpha$ -Maße Null“ und „ $\alpha$ -meßbar“ sind die Abkürzungen der Ausdrücke: fast überall bezüglich des Maßes  $\alpha$ , usw. Die Ausdrücke „fast überall“, „vom Maß Null“ werden im Fall des gewöhnlichen Lebesgueschen Maßes gebraucht.



Ist in einem Intervall  $[c, d] \subseteq [a, b]$  fast überall  $\alpha'(x) > 0$ , so ist (7) in  $[c, d]$  fast überall erfüllt.

Für  $F(x) \in L_\alpha[a, b]$  besteht im Intervall  $[a, b]$   $\alpha$ -fast überall<sup>6</sup>

$$\lim_{h \rightarrow 0} \frac{1}{\alpha(x+h+0) - \alpha(x-0)} \int_0^h F(x \pm v) d\alpha(x \pm v) = F(x).$$

Daraus folgt, daß die Menge  $E_1$  der Punkte  $x \in E$ , in welchen für irgendeine rationale Zahl  $u$  die Relation

$$(8) \quad \lim_{h \rightarrow 0} \frac{1}{\alpha(x+h+0) - \alpha(x-0)} \int_0^h |f(x \pm v) - u|^p d\alpha(x \pm v) = |f(x) - u|^p$$

nicht erfüllt wird, oder  $f(x)$  nicht endlich ist, oder aber die Ungleichung  $\alpha(x-\eta) < \alpha(x) < \alpha(x+\eta)$  ( $\eta > 0$ ) nicht besteht, eine Menge vom  $\alpha$ -Maße Null ist. Es bleibt zu zeigen, daß (7) für  $x \in E - E_1$  gültig ist.

Nehmen wir  $x \in E - E_1$  an und sei  $\varepsilon$  eine beliebige positive Zahl. Wählen wir die rationale Zahl  $u$  derart, daß

$$(9) \quad |f(x) - u|^p < \varepsilon$$

sei. Durch Anwendung der Ungleichung  $|z_1 + z_2|^p \leq 2^{p-1} \{|z_1|^p + |z_2|^p\}$  erhalten wir die Ungleichung

$$\begin{aligned} & \frac{1}{h} \int_0^h |f(x \pm v) - f(x)|^p d\alpha(x \pm v) \leq \\ & \leq \frac{2^{p-1}}{h} \left\{ \int_0^h |f(x \pm v) - u|^p d\alpha(x \pm v) + \int_0^h |f(x) - u|^p d\alpha(x \pm v) \right\} = \\ & = 2^{p-1} \left\{ \frac{\alpha(x+h+0) - \alpha(x)}{h} \frac{1}{\alpha(x+h+0) - \alpha(x)} \int_0^h |f(x \pm v) - u|^p d\alpha(x \pm v) + \right. \\ & \quad \left. + \frac{\alpha(x+h+0) - \alpha(x)}{h} |f(x) - u|^p \right\}. \end{aligned}$$

Da  $\alpha(x)$  im Punkte  $x$  differenzierbar ist, ergibt sich auf Grund von (8) und (9) die Abschätzung

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_0^h |f(x \pm v) - f(x)|^p d\alpha(x \pm v) \leq 2^p \alpha'(x) \varepsilon.$$

Damit ist der erste Teil des Hilfssatzes 1 bewiesen.

<sup>6</sup> S. z. B. F. RIESZ, Sur les fonctions des transformations hermitiennes dans l'espace de Hilbert, *Acta Sci. Math., Szeged*, 7 (1934-35), S. 147-159.



Der zweite Teil geht daraus hervor, daß die Menge, auf welcher  $\alpha(x)$  keine endliche Ableitung hat, vom Maße Null ist, wenn ferner im Intervall  $[c, d]$  fast überall  $\alpha'(x) > 0$  gilt, so ist eine Teilmenge vom  $\alpha$ -Maße Null des Intervalls  $[c, d]$  auch im Lebesgueschen Sinne vom Maße Null.<sup>7</sup>

Der zweite Hilfssatz ist eine lokalisierte Form des Satzes von F. RIESZ über die Koeffizienten der orthogonalen Reihenentwicklungen in  $L^p$ .<sup>8</sup>

HILFSSATZ 2. *Es sei  $\{\varphi_n(x)\}$  ( $n = 0, 1, \dots$ ) ein im Grundintervall  $[a, b]$  bezüglich der Verteilung  $d\alpha(x)$  orthonormiertes Funktionssystem und  $|\varphi_n(x)| \leq M$  ( $n = 0, 1, \dots$ ) in einem Teilintervall  $[c, d]$  von  $[a, b]$ . Ist die Funktion  $f(x)$  außerhalb  $[c, d]$  gleich 0 und*

$$\int_c^d |f(x)|^p d\alpha(x) < \infty \quad (1 < p \leq 2),$$

so gilt die Ungleichung

$$\left\{ \sum_{v=0}^{\infty} |c_v|^q \right\}^{\frac{1}{q}} \leq M^{\frac{(2-p)}{p}} \left\{ \int_c^d |f(x)|^p d\alpha(x) \right\}^{\frac{1}{p}},$$

mit  $\frac{1}{p} + \frac{1}{q} = 1$  und

$$c_v = \int_c^d f(x) \varphi_v(x) d\alpha(x) \quad (v = 0, 1, \dots).$$

Der Beweis dieses Hilfssatzes ergibt sich aus dem Konvexitätssatz von M. RIESZ auf bekannter Weise.<sup>9</sup>

Später wird der folgende reihentheoretische Satz zur Anwendung kommen:

HILFSSATZ 3. *Es sei  $\{s_v\}$  eine beliebige Zahlenfolge. Besteht für jedes positive  $\sigma$  die Beziehung*

$$(10) \quad \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{v=0}^n |s_v - S|^\sigma = 0,$$

so gilt für jedes Zahlenpaar  $r > 0, s > 0$

$$(11) \quad \lim_{n \rightarrow \infty} \frac{1}{A_n^r} \sum_{v=0}^n A_{n-v}^{r-1} |s_v - S|^s = 0.$$

Sind  $s_v$  und  $S$  von einem Parameter abhängig und ist (10) gleichmäßig erfüllt, so besteht auch (11) gleichmäßig.

<sup>7</sup> S. z. B. B. SZ.-NAGY, Über die Konvergenz von Reihen orthogonaler Polynome, *Math. Nachrichten*, 4 (1950-51), S. 50-55.

<sup>8</sup> S. z. B. A. ZYGMUND, a. a. O., S. 190.

<sup>9</sup> S. z. B. A. ZYGMUND, a. a. O., S. 197-201.



Es ist klar, daß es den Hilfssatz nur für den Fall  $s=1$  zu beweisen genügt. Sei alsdann  $r$  eine beliebige positive Zahl ( $r < 1$ ) und wählen wir den Exponenten  $\sigma$  so, daß die Ungleichung

$$(12) \quad \sigma'(r-1) > -1 \quad \left( \frac{1}{\sigma} + \frac{1}{\sigma'} = 1 \right)$$

erfüllt wird. Aus der Hölderschen Ungleichung ergibt sich dann die Abschätzung

$$\frac{1}{A_n^r} \sum_{r=0}^n A_{n-r}^{r-1} |s_r - S| \leq \frac{1}{A_n^r} \left\{ \sum_{r=0}^n (A_{n-r}^{r-1})^{\sigma'} \right\}^{\frac{1}{\sigma'}} \left\{ \sum_{r=0}^n |s_r - S|^\sigma \right\}^{\frac{1}{\sigma}}.$$

Nach der asymptotischen Formel (1) ist

$$\frac{1}{A_n^r} \left\{ \sum_{r=0}^n (A_{n-r}^{r-1})^{\sigma'} \right\}^{\frac{1}{\sigma'}} = O \left( \left\{ \frac{1}{n^{r\sigma'}} \sum_{r=1}^n r^{\sigma'(r-1)} \right\}^{\frac{1}{\sigma'}} \right);$$

auf Grund von (12) ergibt sich

$$\frac{1}{n^{r\sigma'}} \sum_{r=1}^n r^{\sigma'(r-1)} = O \left( \frac{1}{n^{r\sigma'} n^{\sigma'r - \sigma' + 1}} \right) = O \left( \frac{1}{(n+1)^{\sigma'-1}} \right)$$

und somit erhalten wir die Abschätzung

$$\frac{1}{A_n^r} \sum_{r=0}^n A_{n-r}^{r-1} |s_r - S| = O \left( \left\{ \frac{1}{n+1} \sum_{r=0}^n |s_r - S|^\sigma \right\}^{\frac{1}{\sigma}} \right),$$

woraus (11) folgt. Damit ist der Hilfssatz 3 bewiesen, da unsere Behauptung bezüglich der Gleichmäßigkeit in diesem Beweis offenbar enthalten ist.

### § 3. Starke Summation orthogonaler Polynomreihen

Das Analogon für orthogonale Polynomreihen des in der Einleitung erwähnten Hardy-Littlewoodschen starken Summationssatzes ist der folgende

SATZ I. a) Das Polynomsystem  $\{p_n(x)\}$  sei im Intervall  $[c, d] \subseteq [a, b]$  gleichmäßig beschränkt. Ist

$$(13) \quad f(x) \in L_\alpha^2[a, c_1], f(x) \in L_\alpha^p[c_1, d_1] \quad (1 < p \leq 2), f(x) \in L_\alpha^2[d_1, b] \\ (c \leq c_1 < d_1 \leq d),$$

so gilt für ein beliebiges positives  $\sigma$  auf der Menge  $E \cap [c, d]$   $\alpha$ -fast überall

$$(14) \quad \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{r=0}^n |s_r(f; x) - f(x)|^\sigma = 0,$$

wo  $E$  die Menge der Punkte von  $[a, b]$  bezeichnet, in welchen  $\alpha'(x)$  existiert und endlich ist. Ist in  $[c, d]$  fast überall  $\alpha'(x) > 0$ , so gilt (14) in  $[c, d]$  fast überall.



Genauer: (14) besteht in jedem Punkte  $x \in (c, d)$ , wo

$$(15) \quad \int_0^h |f(x \pm v) - f(x)|^p d\alpha(x \pm v) = o(h)$$

erfüllt ist. Ist  $c = a$  bzw.  $d = b$ , so besteht (13) auch in den Punkten  $x = a$  bzw.  $x = b$ , wenn nur

$$(16) \quad \int_0^h |f(a+h) - f(a)|^p d\alpha(a+v) = o(h)$$

bzw.

$$\int_0^h |f(b-v) - f(b)|^p d\alpha(b-v) = o(h)$$

gilt.

b) Nehmen wir an, daß die Polynome im Intervall  $[c, d]$  gleichmäßig beschränkt sind und  $\alpha(x)$  im Intervall  $[c_1, d_1]$  eine Lipschitz-Bedingung erster Ordnung befriedigt.<sup>10</sup> Genügt  $f(x)$  der Bedingung (13), so ist (14) in jedem Stetigkeitspunkte  $x \in (c_1, d_1)$  von  $f(x)$  erfüllt. Ist  $f(x)$  im ganzen Intervall  $(c_1, d_1)$  stetig, so besteht (14) in jedem Teilintervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Eine einfache Folgerung des Satzes I ist folgendes:

Ist das Polynomsystem in jedem inneren Teilintervall von  $[a, b]$  gleichmäßig beschränkt und gilt für ein positives  $\nu$

$$f(x) \in L_\alpha^2[a, a + \nu], f(x) \in L_\alpha^p[a + \nu, b - \nu] \quad (1 < p), f(x) \in L_\alpha^2[b - \nu, b],$$

so ist (14) auf der Menge  $E$   $\alpha$ -fast überall erfüllt; gilt in  $[a, b]$  fast überall  $\alpha'(x) > 0$ , so ist (14) in  $[a, b]$  fast überall erfüllt.

Wenn die Polynome in jedem inneren Teilintervall gleichmäßig beschränkt sind und  $\alpha(x)$  in jedem inneren Teilintervall eine Lipschitz-Bedingung erster Ordnung befriedigt, so ist (14) für jede stetige Funktion in jedem inneren Teilintervall von  $[a, b]$  gleichmäßig erfüllt.

Es sei noch der folgende Spezialfall des Satzes I besonders erwähnt:

Ist das Polynomsystem in dem ganzen Grundintervall  $[a, b]$  gleichmäßig beschränkt und  $f(x) \in L_\alpha^p[a, b]$  ( $p > 1$ ), so ist (14) auf der Menge  $E$   $\alpha$ -fast überall erfüllt, und wenn in  $[a, b]$  fast überall  $\alpha'(x) > 0$  gilt, so ist (14) in  $[a, b]$  fast überall erfüllt.

Zum Beweise des Satzes I bemerken wir, daß im Falle  $\sigma \leq \sigma'$  die Ungleichung

$$\left\{ \frac{1}{n+1} \sum_{r=0}^n |s_r(f; x) - f(x)|^\sigma \right\}^{\frac{1}{\sigma}} \leq \left\{ \frac{1}{n+1} \sum_{r=0}^n |s_r(f; x) - f(x)|^{\sigma'} \right\}^{\frac{1}{\sigma'}}$$

<sup>10</sup> Ist  $\alpha(x)$  absolut stetig, so ist die Gewichtsfunktion in dem Intervall  $[c_1, d_1]$  wesentlich beschränkt.



gilt, daher genügt es (14) für beliebig große Exponenten  $\sigma'$  nachzuweisen. Wählen wir  $\sigma'$  so groß, daß der zugehörige konjugierte Exponent  $p'$  nicht größer als  $p$  wird, so folgt  $f(x) \in L_\alpha^{p'}$  aus der Annahme (13).

Sei nun  $x \in (c, d)$  ein Punkt, in welchem (15) gilt, und wählen wir  $n$  so groß, daß die Ungleichungen  $c \leq x - \frac{1}{n}$ ,  $x + \frac{1}{n} \leq d$  erfüllt werden. Im Fall  $0 \leq \nu \leq n$  schreiben wir die Differenz  $s_\nu(f; x) - f(x)$  laut (2) in folgender Form auf:

$$\begin{aligned} s_\nu(f; x) - f(x) &= \left( \int_a^c + \int_c^{x-\frac{1}{n}} + \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} + \int_{x+\frac{1}{n}}^d + \int_d^b \right) [f(t) - f(x)] K_\nu(x, t) d\alpha(t) = \\ &= \sum_{i=1}^5 I_{i, \nu}^{(n)}(x). \end{aligned}$$

Durch Anwendung der Minkowskischen Ungleichung ergibt sich daraus die Abschätzung

$$(17) \quad \left\{ \frac{1}{n+1} \sum_{\nu=0}^n |s_\nu(f; x) - f(x)|^{\sigma'} \right\}^{\frac{1}{\sigma'}} \leq \sum_{i=1}^5 \left\{ \frac{1}{n+1} \sum_{\nu=0}^n |I_{i, \nu}^{(n)}(x)|^{\sigma'} \right\}^{\frac{1}{\sigma'}} = \sum_{i=1}^5 S_i^{(n)}(x).$$

Wegen der Beschränktheit der Polynome und der Formel (5) erhalten wir

$$\begin{aligned} (18) \quad |I_{3, \nu}^{(n)}(x)| &= \left| \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} [f(t) - f(x)] K_\nu(x, t) d\alpha(t) \right| = O(n) \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} |f(t) - f(x)| d\alpha(t) = \\ &= O(n) \left\{ \int_0^{\frac{1}{n}} |f(x+\nu) - f(x)| d\alpha(x+\nu) + \int_0^{\frac{1}{n}} |f(x-\nu) - f(x)| d\alpha(x-\nu) \right\}, \end{aligned}$$

bei festem  $\nu$  gilt also  $\lim_{n \rightarrow \infty} I_{3, \nu}^{(n)}(x) = 0$ , d. h.  $S_3^{(n)}(x) = o(1)$ .

Wir zeigen nun, daß  $S_1^{(n)}(x) = o(1)$  und  $S_5^{(n)}(x) = o(1)$  ist. Aus der Christoffel-Darboux'schen Formel ergibt sich nämlich die Beziehung

$$\begin{aligned} (19) \quad |I_{1, \nu}^{(n)}(x)| &= \left| \int_a^c [f(t) - f(x)] K_\nu(x, t) d\alpha(t) \right| = \\ &= O(1) \left| p_\nu(x) \int_a^c \frac{f(t) - f(x)}{t - x} p_{\nu+1}(t) d\alpha(t) - p_{\nu+1}(x) \int_a^c \frac{f(t) - f(x)}{t - x} p_\nu(t) d\alpha(x) \right|. \end{aligned}$$

Wegen der gleichmäßigen Beschränktheit der Polynome ist

$$(20) \quad |I_{1, \nu}^{(n)}(x)| = O(1) (|\gamma_{\nu+1}(x)| + |\gamma_\nu(x)|),$$



wo  $\gamma_r(x)$  der  $r$ -te verallgemeinerte Fourierkoeffizient der Funktion

$$g_x(t) = \begin{cases} \frac{f(t)-f(x)}{t-x} & (a \leq t \leq c), \\ 0 & (c < t \leq b) \end{cases}$$

bedeutet. Nach (13) ist  $g_x(t) \in L^2_\alpha$ , daher  $\lim_{r \rightarrow \infty} \gamma_r(x) = 0$ . Hieraus folgt  $S_1^{(n)}(x) = o(1)$ . Auf derselben Weise ergibt sich auch  $S_5^{(n)}(x) = o(1)$ .

Wir haben noch nachzuweisen, daß auch die Beziehungen  $S_2^{(n)}(x) = o(1)$  und  $S_4^{(n)}(x) = o(1)$  bestehen. Mit Hilfe der Christoffel-Darboux'schen Formel erhalten wir die Abschätzung

$$\begin{aligned} |I_{2,r}^{(n)}(x)| &= \left| \int_c^{x-\frac{1}{n}} [f(t)-f(x)] K_r(x,t) d\alpha(t) \right| = \\ &= O(1) \left| p_r(x) \int_c^{x-\frac{1}{n}} \frac{f(t)-f(x)}{t-x} p_{r+1}(t) d\alpha(t) - p_{r+1}(x) \int_c^{x-\frac{1}{n}} \frac{f(t)-f(x)}{t-x} p_r(t) d\alpha(t) \right|. \end{aligned}$$

Wegen der gleichmäßigen Beschränktheit der Polynome ergibt sich also die Abschätzung

$$(21) \quad |I_{2,r}^{(n)}(x)| = O(1) (|\chi_{r+1}^{(n)}(x)| + |\chi_r^{(n)}(x)|),$$

wo  $\chi_r^{(n)}(x)$  den  $r$ -ten verallgemeinerten Fourierkoeffizienten der Funktion

$$h_x^{(n)}(t) = \begin{cases} \frac{f(t)-f(x)}{t-x} & \left( c \leq t \leq x - \frac{1}{n} \right), \\ 0 & \left( t < c, \text{ oder } x - \frac{1}{n} < t \right) \end{cases}$$

bedeutet. Nach (21) ist

$$|I_{2,r}^{(n)}(x)|^{\sigma'} = O(1) \{ |\chi_{r+1}^{(n)}(x)|^{\sigma'} + |\chi_r^{(n)}(x)|^{\sigma'} \},$$

mithin gilt

$$S_2^{(n)}(x) = \left\{ \frac{1}{n+1} \sum_{r=0}^n |I_{2,r}^{(n)}(x)|^{\sigma'} \right\}^{\frac{1}{\sigma'}} = O(1) \left\{ \frac{1}{n+1} \sum_{r=0}^{\infty} |\chi_r^{(n)}(x)|^{\sigma'} \right\}^{\frac{1}{\sigma'}}.$$

Da die Polynome in dem Intervall  $[c, d]$  gleichmäßig beschränkt sind, ferner  $h_x^{(n)}(t)$  außerhalb des Intervalls  $[c, d]$  gleich Null ist und  $h_x^{(n)}(t) \in L^p_\alpha[c, d]$  gilt, kann der Hilfssatz 2 angewendet werden. Daraus folgt die Abschätzung

$$\begin{aligned} (22) \quad S_2^{(n)}(x) &= O(1) \frac{1}{(n+1)^{\frac{1}{\sigma'}}} \left\{ \int_c^{x-\frac{1}{n}} \left| \frac{f(t)-f(x)}{t-x} \right|^{p'} d\alpha(t) \right\}^{\frac{1}{p'}} = \\ &= O(1) \frac{1}{(n+1)^{\frac{1}{\sigma'}}} \left\{ \int_{\frac{1}{n}}^{x-c} \frac{|f(x-v)-f(x)|^{p'}}{v^{p'}} d\alpha(x-v) \right\}^{\frac{1}{p'}}. \end{aligned}$$



Da nach Annahme (15) im Punkte  $x$  erfüllt ist, gilt auch die Relation

$$\Phi_x(h) = \int_0^h |f(x-v) - f(x)|^{p'} d\alpha(x-v) = o(h)$$

und damit erhalten wir die Abschätzung

$$\begin{aligned} & \int_{\frac{1}{n}}^{x-c} \frac{|f(x-v) - f(x)|^{p'}}{v^{p'}} d\alpha(x-v) = \\ (23) \quad & = \frac{\Phi_x(x-c+0)}{(x-c)^{p'}} - \frac{\Phi_x\left(\frac{1}{n}-0\right)}{\left(\frac{1}{n}\right)^{p'}} + p' \int_{\frac{1}{n}}^{x-c} \frac{\Phi_x(v)}{v^{p'+1}} dv = \\ & = o(n^{p'-1}) + \int_{\frac{1}{n}}^{b-a} o(v^{-p'}) dv = o(n^{p'-1}).^{11} \end{aligned}$$

Nach (22) folgt daraus  $S_2^{(n)}(x) = o(1)$  und auf derselben Weise auch  $S_4^{(n)}(x) = o(1)$ ; die Beziehung (14) besteht also im Punkte  $x$ . Da (15) nach dem Hilfssatz 1 auf der Menge  $E \cap [c, d]$   $\alpha$ -fast überall bzw. für  $\alpha'(x) > 0$  in  $[c, d]$  fast überall erfüllt ist, haben wir damit den Teil a) des Satzes I vollständig bewiesen. Die Behauptung bezüglich der starken Summierbarkeit in den Endpunkten des Grundintervalls bedarf nach dem Vorangehenden keinen eingehenden Beweis, es sei nur bemerkt, daß man z. B. im Falle  $x = a$  aus der Zerlegung

$$s_r(f; x) - f(x) = \left( \int_a^{a+\frac{1}{n}} + \int_{a+\frac{1}{n}}^t + \int_d^b \right) [f(t) - f(x)] K_r(x, t) d\alpha(t)$$

ausgehen muß.

Der Teil b) des Satzes I kann ähnlich bewiesen werden. Ist  $f(x)$  im Punkte  $x \in (c_1, d_1)$  stetig, so ist (15) wegen unserer Annahme über  $\alpha(x)$  im Punkte  $x$  erfüllt, folglich besteht (14) gemäß des vorangehenden Beweises.

Ist  $f(x)$  im ganzen Intervall  $(c_1, d_1)$  stetig, so ist (15) in einem beliebigen Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig erfüllt und daher bestehen in diesem Intervall nach (18) bzw. (22) und (23) die Beziehungen  $S_3^{(n)}(x) = o(1)$ ,  $S_2^{(n)}(x) = o(1)$  und  $S_4^{(n)}(x) = o(1)$  gleichmäßig.

Zum Beweis des Teiles b) muß noch nachgewiesen werden, daß in einem beliebigen Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig  $S_1^{(n)}(x) = o(1)$

<sup>11</sup> Bezüglich der Möglichkeit dieser partiellen Integration s. z. B. S. SAKS, *Theory of the integral* (Warszawa—Lwów, 1937), S. 37, Satz (15.1) und S. 102, Satz (14.1).



gilt. Da die Polynome im Intervall  $[c, d]$  gleichmäßig beschränkt sind, ferner  $f(x)$  im Intervall  $[c_1 + \delta, d_1 - \delta]$  beschränkt und wegen  $t \in [a, c], x \in [c_1 + \delta, d_1 - \delta], |t-x| > \delta$  ist, kann auf Grund von (19)  $|I_{1,r}^{(n)}(x)| \leq 1$  für jedes  $n$  und  $r$  angenommen werden. (Im entgegengesetzten Fall betrachten wir nämlich die Funktion  $mf(x)$ , wo  $m$  eine geeignet gewählte Konstante ist.) Da  $\sigma' \geq 2$  ist, folgt

$$|S_1^{(n)}(x)|^{\sigma'} = \frac{1}{n+1} \sum_{r=0}^n |I_{1,r}^{(n)}(x)|^{\sigma'} \leq \frac{1}{n+1} \sum_{r=0}^n [I_{1,r}^{(n)}(x)]^2.$$

Auf Grund von (20) erhalten wir durch Anwendung der Besselschen Ungleichung die Abschätzung

$$\begin{aligned} |S_1^{(n)}(x)|^{\sigma'} &= \frac{O(1)}{n+1} \int_a^c \left[ \frac{f(t)-f(x)}{t-x} \right]^2 d\alpha(t) = \\ &= \frac{O(1)}{n+1} \frac{1}{\delta^2} \left\{ \int_a^c f^2(t) d\alpha(t) + \max_{c_1+\delta \leq x \leq d_1-\delta} f^2(x) \int_a^c d\alpha(t) \right\}, \end{aligned}$$

woraus sich das gleichmäßige Bestehen von  $S_1^{(n)}(x) = o(1)$  in einem beliebigen Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) ergibt. Da dasselbe auch für  $S_5^{(n)}(x)$  auf derselben Weise folgt, ist damit der Satz I vollständig bewiesen.

#### § 4. (C, r > 0)-Summation orthogonaler Polynomreihen

Durch Anwendung des Hilfssatzes 3 ergibt sich aus dem Satze I unmittelbar der

SATZ II. a) Ist das Polynomsystem  $\{p_n(x)\}$  im Intervall  $[c, d] \subseteq [a, b]$  gleichmäßig beschränkt und befriedigt die Funktion  $f(x)$  die Bedingung (13), so ist bei beliebigem  $r > 0, s > 0$  auf der Menge  $E \cap [c, d]$   $\alpha$ -fast überall

$$(24) \quad \lim_{n \rightarrow \infty} \frac{1}{A_n^r} \sum_{r=0}^n A_{n-r}^{r-1} |s_r(f; x) - f(x)|^s = 0,^{12}$$

wo  $E$  die Menge der Punkte  $x \in [a, b]$  bezeichnet, wo  $\alpha'(x)$  existiert und endlich ist. Für  $\alpha'(x) > 0$  in  $[c, d]$  fast überall ist (24) in  $[c, d]$  fast überall erfüllt. Genauer: (24) besteht in jedem Punkte  $x \in (c, d)$ , in welchem die Bedingung (15) erfüllt ist. Wenn  $c = a$  bzw.  $d = b$  ist, so besteht (24) im Punkte  $x = a$  bzw.  $x = b$ , wenn nur die entsprechende Bedingung (16) erfüllt ist.

b) Nehmen wir an, daß die Polynome im Intervall  $[c, d]$  gleichmäßig beschränkt sind und  $\alpha(x)$  im Intervall  $[c_1, d_1]$  eine Lipschitz-Bedingung erster Ordnung befriedigt. Genügt  $f(x)$  der Bedingung (13), so ist (24) in jedem Stetigkeits-

<sup>12</sup> Wir bemerken, daß (24) nach dem starken Summationssatz von HARDY—LITTLEWOOD und dem Hilfssatz 3 auch für die Fourierreihen von Funktionen in  $L^p (p > 1)$  fast überall erfüllt wird.



punkte  $x \in (c_1, d_1)$  von  $f(x)$  erfüllt. Ist  $f(x)$  im ganzen Intervall  $(c_1, d_1)$  stetig, so besteht (24) in jedem Teilintervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Es ist klar, daß die  $(C, r > 0)$ -Summierbarkeit der Entwicklung von  $f(x)$  aus (24) hervorgeht. Für  $s = 1$  folgt nämlich

$$\lim_{n \rightarrow \infty} \frac{1}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{s_\nu(f; x) - f(x)\} = 0,$$

m. a. W.  $\sigma_n^r(f; x) \rightarrow f(x)$ .

Das folgende Korollar des Satzes II soll besonders erwähnt werden:

Sei  $f(x)$  stetig im ganzen Grundintervall  $[a, b]$ . Ist das Polynomsystem in einem Intervall  $[c, d] \subseteq [a, b]$  gleichmäßig beschränkt und erfüllt  $\alpha(x)$  in  $[c, d]$  eine Lipschitz-Bedingung erster Ordnung, so gilt in einem beliebigen Intervall  $[c + \delta, d - \delta]$  ( $\delta > 0$ ) gleichmäßig

$$(25) \quad \sigma_n^r(f; x) \rightarrow f(x) \quad (r > 0).$$

Sind die Polynome in einem beliebigen inneren Teilintervall gleichmäßig beschränkt und erfüllt  $\alpha(x)$  in jedem inneren Teilintervall eine Lipschitz-Bedingung erster Ordnung, so besteht (25) in einem beliebigen Intervall  $[a + \delta, b - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Aus dieser Folgerung kann mit bekannten Methoden<sup>13</sup> bewiesen werden der

SATZ III. Ist das Polynomsystem in einem Intervall  $[c, d] \subseteq [a, b]$  gleichmäßig beschränkt und befriedigt  $\alpha(x)$  in  $[c, d]$  eine Lipschitz-Bedingung erster Ordnung, so gilt in einem beliebigen Intervall  $[c + \delta, d - \delta]$  ( $\delta > 0$ ) gleichmäßig

$$(26) \quad L_n^{(r)}(x) = \int_a^b |K_n^r(x, t)| d\alpha(t) = O(1).$$

Sind die Polynome in einem beliebigen inneren Teilintervall gleichmäßig beschränkt und befriedigt  $\alpha(x)$  in jedem inneren Teilintervall eine Lipschitz-Bedingung erster Ordnung, so besteht (26) in einem beliebigen Intervall  $[c + \delta, d - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Durch Anwendung dieses Satzes beweisen wir den

SATZ IV. Das Polynomsystem  $\{p_n(x)\}$  sei in einem Intervall  $[c, d] \subseteq [a, b]$  gleichmäßig beschränkt und  $\alpha(x)$  erfülle in  $[c, d]$  eine Lipschitz-Bedingung erster Ordnung. Gilt

$$f(x) \in L_\alpha^2[a, c], \quad f(x) \in L_\alpha[c, d], \quad f(x) \in L_\alpha^2[d, b],$$

so ist (25) in jedem Stetigkeitspunkt von  $f(x)$  erfüllt. Ist  $f(x)$  im ganzen Intervall  $(c_1, d_1) \subseteq [c, d]$  stetig, so gilt (25) in einem beliebigen Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig.

<sup>13</sup> S. z. B. H. LEBESGUE, Sur les intégrales singulières, *Annales de Toulouse*, 1 (1909), S. 25–117.



Als Spezialfall des Satzes IV soll besonders erwähnt werden:

Nehmen wir an, das Polynomsystem sei im ganzen Grundintervall  $[a, b]$  (oder auf jedem inneren Teilintervall von  $[a, b]$ ) gleichmäßig beschränkt und  $\alpha(x)$  erfülle in jedem inneren Teilintervall von  $[a, b]$  eine Lipschitz-Bedingung erster Ordnung. Ist  $f(x) \in L_\alpha[a, b]$  (oder für irgendein positives  $\eta$ :

$$f(x) \in L_\alpha^2[a, a + \eta], \quad f(x) \in L_\alpha[a + \eta, b - \eta] \quad \text{und} \quad f(x) \in L_\alpha[b - \eta, b],$$

so gilt (25) in jedem Stetigkeitspunkte  $x \in (a, b)$  von  $f(x)$ . Ist  $f(x)$  in einem ganzen Intervall  $(c_1, d_1) \subseteq [a, b]$  stetig, so besteht (25) in jedem Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Sei in der Tat  $x \in (c, d)$  ein Stetigkeitspunkt von  $f(x)$  und schreiben wir die Differenz  $\sigma_n^r(f; x) - f(x)$  auf Grund der Formel (4) in folgender Form:

$$\begin{aligned} \sigma_n^r(f; x) - f(x) &= \left( \int_a^c + \int_c^{x-h} + \int_{x-h}^{x+h} + \int_{x+h}^d + \int_d^b \right) [f(t) - f(x)] K_n^r(x, t) d\alpha(t) = \\ &= I_1(x) + I_2(x) + I_3(x) + I_4(x) + I_5(x), \end{aligned}$$

wo  $h$  eine noch zu bestimmende positive Zahl bedeutet. Ist  $h$  genügend klein, so folgt auf Grund des Satzes III  $|I_3(x)| < \varepsilon$  aus der Abschätzung

(27)

$$|I_3(x)| = \left| \int_{x-h}^{x+h} [f(t) - f(x)] K_n^r(x, t) d\alpha(t) \right| \leq \sup_{x-h \leq t \leq x+h} |f(t) - f(x)| \int_a^b |K_n^r(x, t)| d\alpha(t).$$

Halten wir  $h$  solcherweise fest und schreiben wir das Integral  $I_2(x)$  mit Hilfe der Formeln (4) und (6) folgenderweise:

$$\begin{aligned} (28) \quad |I_2(x)| &= \left| \frac{1}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \int_c^{x-h} [f(t) - f(x)] K_\nu(x, t) d\alpha(t) \right| = \\ &= \frac{O(1)}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}(x)| + |\chi_\nu(x)| \}, \end{aligned}$$

wo  $\chi_\nu(x)$  den  $\nu$ -ten verallgemeinerten Fourierkoeffizienten der Funktion

$$h_\nu(t) = \begin{cases} \frac{f(t) - f(x)}{t - x} & (c \leq t \leq x - h), \\ 0 & (t < c \text{ oder } x - h < t) \end{cases}$$

bedeutet. Es ist leicht zu zeigen, daß  $\lim_{\nu \rightarrow \infty} \chi_\nu(x) = 0$  ist. Sei nun  $\eta$  eine beliebige positive Zahl und  $f_1(x)$  eine beschränkte,  $\alpha$ -meßbare Funktion, welche außerhalb des Intervalls  $[c, x - h]$  gleich 0 ist und für welche

$$\int_c^{x-h} |h_\nu(t) - f_1(t)| d\alpha(t) < \eta$$



gilt. Dann ist

$$\begin{aligned} |\chi_r(x)| &= \left| \int_c^{x-h} h_x(t) p_r(t) d\alpha(t) \right| \leq \left| \int_c^{x-h} [h_x(t) - f_1(t)] p_r(t) d\alpha(t) \right| + \\ &+ \left| \int_c^{x-h} f_1(t) p_r(t) d\alpha(t) \right| \leq \sup_{\substack{c \leq x \leq d \\ n=0, 1, \dots}} |p_n(x)| r_1 + \left| \int_a^b f_1(t) p_r(t) d\alpha(t) \right|. \end{aligned}$$

Da die Polynome  $\{p_n(x)\}$  im Intervall  $[c, d]$  gleichmäßig beschränkt sind und  $f_1(t) \in L_\alpha^2$  ist, ergibt sich hieraus die Beziehung  $\lim_{r \rightarrow \infty} \chi_r(x) = 0$ . Somit erhalten wir aus (28) durch Anwendung der Identität

$$(29) \quad A_n = \sum_{r=0}^n A_{n-r}^{r-1}$$

für genügend große  $n$  die Abschätzung  $|I_2(x)| < \varepsilon$ .

Einso, wie im Fall von  $I_2(x)$  erhalten wir die Abschätzung

$$(30) \quad |I_1(x)| = \frac{O(1)}{A_n^r} \sum_{r=0}^n A_{n-r}^{r-1} \{ |\gamma_{r+1}(x)| + |\gamma_r(x)| \},$$

wo  $\gamma_r(x)$  der  $r$ -te verallgemeinerte Fourierkoeffizient der  $L_\alpha^2$ -integrierbaren Funktion

$$g_x(t) = \begin{cases} \frac{f(t) - f(x)}{t - x} & (a \leq t \leq c), \\ 0 & (c < t) \end{cases}$$

ist. Wegen  $\lim_{r \rightarrow \infty} \gamma_r(x) = 0$  und (30) ergibt sich  $|I_1(x)| < \varepsilon$  für genügend große  $n$ . Da ähnliche Abschätzungen auch für  $I_4(x)$  und  $I_5(x)$  gelten, ist damit (25) vollständig bewiesen.

Der Beweis der Behauptung bezüglich der Gleichmäßigkeit ist dem obigen ähnlich, nur etwas langwieriger. Sei  $f(x)$  im Intervall  $(c_1, d_1)$  stetig, so ergibt sich aus Satz III und aus (27), daß  $|I_3(x)| < \varepsilon$  für ein beliebig kleines  $h$  im Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig gilt. Es soll nun nachgewiesen werden, daß  $I_2(x) = o(1)$  im Intervall  $[c_1 + \delta, d_1 - \delta]$  gleichmäßig erfüllt wird. Sei zu diesem Zweck  $r_1$  eine beliebige positive Zahl und  $f_2(t)$  eine beschränkte,  $\alpha$ -meßbare Funktion, welche außerhalb des Intervalls  $[c, d]$  gleich 0 ist und für welche

$$(31) \quad \int_c^d |f(t) - f_2(t)| d\alpha(t) < r_1$$

gilt. Betrachten wir folgende Darstellung von  $\chi_r(x)$ :

$$\begin{aligned} \chi_r(x) &= \int_c^{x-h} \frac{f_2(t)}{t-x} p_r(t) d\alpha(t) + \int_c^{x-h} \frac{f(t) - f_2(t)}{t-x} p_r(t) d\alpha(t) - f(x) \int_c^{x-h} \frac{p_r(t)}{t-x} d\alpha(t) = \\ &= \chi_r^{(1)}(x) + \chi_r^{(2)}(x) + \chi_r^{(3)}(x). \end{aligned}$$



Hieraus ergibt sich nach (28) die Abschätzung

$$(32) \quad |I_2(x)| = O(1) \left\{ \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(1)}(x)| + |\chi_{\nu}^{(1)}(x)| \} + \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(2)}(x)| + |\chi_{\nu}^{(2)}(x)| \} + \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(3)}(x)| + |\chi_{\nu}^{(3)}(x)| \} \right\}.$$

Da  $\chi_{\nu}^{(1)}(x)$  der verallgemeinerte Fourierkoeffizient der Funktion  $\frac{f_2(t)}{t-x} \in L_{\alpha}^2$  ist, erhalten wir die Abschätzung

$$(33) \quad \left| \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(1)}(x)| + |\chi_{\nu}^{(1)}(x)| \} \right| = \frac{O(1)}{A_n} \sqrt{\sum_{\nu=0}^n (A_{n-\nu}^{r-1})^2} \sqrt{\sum_{\nu=0}^{\infty} [\chi_{\nu}^{(1)}(x)]^2} = \\ = O(1) \left\{ \frac{1}{n^r} \sqrt{\sum_{\nu=0}^n \nu^{2r-2}} \right\} \int_c^{x-h} \left[ \frac{f_2(t)}{t-x} \right]^2 d\alpha(t) \left\{ \frac{1}{\sqrt{n}} \right\} = \frac{O(1)}{\sqrt{n}} \left\{ \frac{1}{h^2} \int_c^d f_2^2(t) d\alpha(t) \right\}^{\frac{1}{2}}.$$

Ähnlicherweise ergibt sich auch folgendes:

$$(34) \quad \left| \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(3)}(x)| + |\chi_{\nu}^{(3)}(x)| \} \right| = \frac{O(1)}{\sqrt{n}} \left\{ \max_{c_1+\delta \leq x \leq d_1-\delta} f^2(x) \int_c^{x-h} \frac{d\alpha(t)}{(t-x)^2} \right\}^{\frac{1}{2}} = \\ = \frac{O(1)}{\sqrt{n}} \left\{ \frac{1}{h^2} \int_c^d d\alpha(t) \right\}^{\frac{1}{2}}.$$

Ferner, da die Polynome im Intervall  $[c, d]$  gleichmäßig beschränkt sind, besteht nach (31) die Ungleichung

$$|\chi_{\nu}^{(2)}(x)| = \left| \int_c^{x-h} \frac{f(t) - f_2(t)}{t-x} p_{\nu}(t) d\alpha(t) \right| \leq \\ \leq \sup_{\substack{c \leq x \leq d \\ n=0, 1, \dots}} |p_n(x)| \frac{1}{h} \int_c^d |f(t) - f_2(t)| d\alpha(t) = O(1) \eta_1,$$

wo  $\eta_1 (> 0)$  beliebig klein ist. Somit gilt wegen (29):

$$(35) \quad \frac{1}{A_n} \sum_{\nu=0}^n A_{n-\nu}^{r-1} \{ |\chi_{\nu+1}^{(2)}(x)| + |\chi_{\nu}^{(2)}(x)| \} = O(1) \eta_1.$$

Aus (32), (33), (34) und (35) folgt  $I_2(x) = o(1)$  im Intervall  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ) gleichmäßig.

Da  $g_x(t) \in L_{\alpha}^2$  ist, ergibt sich endlich aus (30) die Abschätzung

$$|I_1(x)| = \frac{O(1)}{\sqrt{n}} \left\{ \frac{1}{\theta^2} \int_c^d [f(t) - f(x)]^2 d\alpha(t) \right\}^{\frac{1}{2}} = \\ = \frac{O(1)}{\sqrt{n}} \left\{ \int_c^d f^2(t) d\alpha(t) + \max_{c_1+\delta \leq x \leq d_1-\delta} f^2(x) \int_c^d d\alpha(t) \right\}^{\frac{1}{2}}.$$



und daher gilt auch  $I_1(x) = o(1)$  im Intervall  $[c_1 + \delta, d_1 - \delta]$  gleichmäßig. Die Integrale  $I_4(x)$  und  $I_5(x)$  können ähnlicherweise abgeschätzt werden, womit der Satz IV vollständig bewiesen ist.

### § 5. Bemerkung

In den Beweisen der Sätze I—IV wurde ausschließlich jene Eigenschaft der orthogonalen Polynome ausgenutzt, daß der Kern  $K_n(x, t)$  durch die Christoffel-Darboux'sche Formel dargestellt werden kann. Daraus folgt, wie es von G. ALEXITS bemerkt wurde, daß diese Sätze für ein beliebiges orthonormiertes Funktionssystem  $\{\varphi_n(x)\}$  gelten, dessen Kern  $K_n(x, t)$  folgenderweise darstellbar ist:

$$K_r(x, t) = \frac{1}{\Phi(t) - \Phi(x)} \sum_{i, j=r-k}^{r+k} a_{ij}^{(r)} \varphi_i(x) \varphi_j(t),$$

wo  $k$  eine feste natürliche Zahl ist, die Zahlen  $a_{ij}^{(r)}$  beschränkt sind und die Funktion  $\Phi(t)$  der Bedingung  $|\Phi(t) - \Phi(x)| > m|t - x|$  ( $m = m(x) > 0$ ) genügt.

Ich möchte der Gelegenheit benutzen, Prof. G. ALEXITS meinen tiefsten Dank auszusprechen für die wertvollen Ratschläge bei der Zusammenstellung der vorliegenden Arbeit.

*(Eingegangen am 27. Februar 1954.)*



## О СУММИРУЕМОСТИ МЕТОДОМ ЧЕЗАРО РЯДОВ ОРТОГОНАЛЬНЫХ ПОЛИНОМОВ. II

К. ТАНДОРИ (Серед)

(Резюме)

Пусть дана функция  $\alpha(x)$ , определенная на конечном интервале  $[a, b]$ , и там же неубывающая и ограниченная. Обозначим через  $\{p_n(x)\}$  систему полиномов, ортонормированную на основном интервале  $[a, b]$  относительно распределения  $d\alpha(x)$ .

Главные результаты работы состоят в следующем:

Предположим систему полиномов  $\{p_n(x)\}$  равномерно ограниченной на подинтервале  $[c, d] \subseteq [a, b]$ . Если

$$(*) \quad f(x) \in L_\alpha^2[a, c_1], \quad f(x) \in L_\alpha^p[c_1, d_1] \quad (1 < p \leq 2; c \leq c_1 < d_1 \leq d), \quad f(x) \in L_\alpha^2[d_1, b],$$

то при любых  $r > 0, s > 0$  на множестве  $E \cap [c, d]$   $\alpha$ -почти всюду

$$(**) \quad \lim_{n \rightarrow \infty} \frac{1}{A_n^r} \sum_{\nu=0}^n A_{n-\nu}^{r-1} |s_\nu(f; x) - f(x)|^s = 0 \quad (A_n^r = \binom{n+r}{n}),$$

где  $E \subset [a, b]$  есть множество точек  $x$ , в которых существует конечная производная  $\alpha'(x)$ , а  $s_\nu(f; x)$  означает  $\nu$ -ую частичную сумму разложения  $f(x)$  по системе  $\{p_n(x)\}$ .

(\*\*) выполняется во всех точках  $x \in (c, d)$ , в которых

$$\int_0^h |f(x \pm v) - f(x)|^p d\alpha(x \pm v) = o(h).$$

Предположим, что система полиномов на подинтервале  $[c, d]$  равномерно ограничена, и на подинтервале  $[c_1, d_1]$   $\alpha(x)$  удовлетворяет условию Липшица первого порядка. Если (\*) имеет место, то (\*\*) выполняется во всякой точке  $x \in (c_1, d_1)$ , в которой  $f(x)$  непрерывна. Если  $f(x)$  непрерывна на всем интервале  $(c_1, d_1)$ , то (\*\*) равномерно выполняется на всяком отрезке  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ).

Из (\*\*) очевидно следует:

$$(***) \quad \lim_{n \rightarrow \infty} \sigma_n^r(f; x) = f(x),$$

где  $\sigma_n^r(f; x)$  есть  $n$ -ое  $(C, r)$ -среднее разложения функции  $f(x)$  по системе  $\{p_n(x)\}$ .

Предположим, что система полиномов  $\{p_n(x)\}$  равномерно ограничена на подинтервале  $[c, d]$ , а функция  $\alpha(x)$  удовлетворяет условию Липшица первого порядка на  $[c, d]$ . Если при таких условиях

$$f(x) \in L_\alpha^2[a, c], \quad f(x) \in L_\alpha[c, d], \quad f(x) \in L_\alpha^2[d, b],$$

то (\*\*\*) выполняется во всякой точке  $x \in (c, d)$ , в которой  $f(x)$  непрерывна. В случае непрерывности  $f(x)$  во всем подинтервале  $[c_1, d_1] \subseteq [c, d]$ , (\*\*\*) выполняется равномерно в любом интервале  $[c_1 + \delta, d_1 - \delta]$  ( $\delta > 0$ ).

Как это заметил Г. Алексич, приведенные теоремы справедливы для любой системы функций  $\{\varphi_n(x)\}$ , ортонормированной относительно распределения  $d\alpha(x)$ , ядро  $\nu$ -ой частичной суммы которой представляется в виде

$$K_\nu(x, t) = \frac{1}{\Phi(t) - \Phi(x)} \sum_{i, j = \nu-k}^{\nu+k} a_{ij}^{(\nu)} \varphi_i(x) \varphi_j(t),$$

где  $k$  есть фиксированное натуральное число, коэффициенты  $a_{ij}^{(\nu)}$  ограничены в совокупности по абсолютной величине, и  $|\Phi(t) - \Phi(x)| > m(x)|t - x|$ ,  $m(x) > 0$ .







ELEMENTARGEOMETRISCHER BEWEIS  
 DER WIDERSPRUCHSFREIHEIT  
 DER HYPERBOLISCHEN RAUMGEOMETRIE  
 MIT HILFE DES POINCARÉSCHEM HALBRAUMES

Von  
 PAUL SZÁSZ (Budapest)  
 (Vorgelegt von G. HAJÓS)

Zur Zentenarfeier der Geburt HENRI POINCARÉ'S

H. POINCARÉ<sup>1</sup> hat als eine Verwirklichung der hyperbolischen Raumgeometrie folgendes Modell, wir möchten sagen *Bildgeometrie* oder *Pseudogeometrie* angeben, wenn auch in etwas anderer Form.

Es sei im euklidischen Raume eine Ebene  $\gamma$  als *Fundamentalebene* vorgelegt. Die inneren Punkte eines der durch  $\gamma$  bestimmten zwei Halbräume sollen *Pseudopunkte*, die im Innern dieses Halbraumes liegenden und zu  $\gamma$  rechtwinkligen Halbkreise bzw. Halbgeraden (die wir *Orthogonalhalbkreise* resp. *Orthogonalhalbgeraden* nennen wollen) *Pseudogeraden*, die zu  $\gamma$  rechtwinkligen Halbkugeln bzw. Halbebenen (die *Orthogonalhalbkugeln* resp. *Orthogonalhalbebenen*) *Pseudoebenen* heißen. Liegen die Pseudopunkte  $P_1, P_2$  nicht auf einer Orthogonalhalbgeraden, so soll der Bogen  $\overline{P_1 P_2}$  des durch sie gelegten Orthogonalhalbkreises, im entgegengesetzten Falle die Strecke  $P_1 P_2$ , als die *Pseudostrecke*  $\overline{P_1 P_2}$  erklärt werden. Als *Charakteristik* der Pseudostrecke  $\overline{P_1 P_2}$  erklären wir im ersten Falle das Doppelverhältnis

$$(\Xi H P_2 P_1) = \frac{\Xi P_2}{P_2 H} : \frac{\Xi P_1}{P_1 H},$$

wobei  $\Xi, H$  die Endpunkte des durch  $P_1$  und  $P_2$  gelegten Orthogonalhalbkreises sind, so bezeichnet, daß  $P_2$  auf diesem Halbkreisbogen  $\overline{\Xi H}$  zwischen  $P_1$  und  $H$  liegt (Fig. 1). Im zweiten Falle sei diese Charakteristik das Teilungsverhältnis

$$(\Xi P_2 P_1) = \Xi P_2 : \Xi P_1,$$

<sup>1</sup> H. POINCARÉ, Mémoire sur les groupes kleinéens, *Acta Mathematica*, 3 (1883), S. 49–92, besonders S. 55–56, oder *Oeuvres de Henri Poincaré*, t. II (Paris, 1916), S. 258–299, insbesondere S. 264–265.



wenn die Bezeichnung so gewählt ist, daß  $P_1$  zwischen  $P_2$  und dem Anfangspunkt  $\Xi$  der  $P_1$  und  $P_2$  enthaltenden Orthogonalhalbgeraden liegt (Fig. 2). Zwei Pseudostrecken sollen einander *pseudokongruent* oder *pseudogleich* heißen, falls sie dieselbe Charakteristik haben. Dem gegenüber soll der

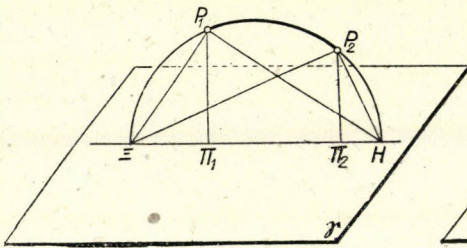


Fig. 1.

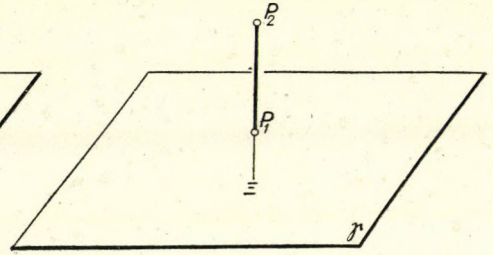


Fig. 2.

*Pseudowinkel* zweier Pseudogeraden  $l, m$  einfach durch den Winkel der sie erzeugenden Orthogonalhalbkreise (deren einer in eine Orthogonalhalbgerade entarten kann) charakterisiert sein (Fig. 3), und zwei Pseudowinkel von gleicher Charakteristik sollen als einander *pseudokongruent* (*pseudogleich*) angesehen werden.

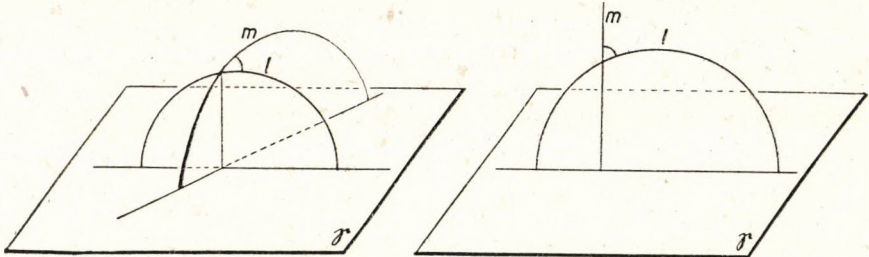


Fig. 3.

In der durch diese Festsetzungen erklärten Bildgeometrie von H. POINCARÉ sind in der Tat alle Axiome der hyperbolischen Raumgeometrie gültig, mit anderen Worten ist dieser sogenannte *Poincarésche Halbraum* wahrlich ein Modell des hyperbolischen Raumes. Unser Ziel ist dafür einen *elementargeometrischen Beweis* zu geben. Dabei soll als Axiomensystem der hyperbolischen Geometrie der Inbegriff folgender Axiome zugrunde gelegt werden: die Axiome der Verknüpfung, der Anordnung und der Kongruenz von D. HILBERT,<sup>2</sup> die Negation des euklidischen Parallelenaxioms, endlich das Archimedische Axiom und das Cantorsche Stetigkeitsaxiom.

<sup>2</sup> D. HILBERT, *Grundlagen der Geometrie*, 7. Aufl. (Leipzig und Berlin, 1930), S. 3—5, 11—14.



Wir können uns auf den Beweis des Kongruenzaxioms<sup>3</sup> III<sub>5</sub> beschränken, da das Bestehen aller übrigen Axiome unmittelbar einleuchtet. Zu beweisen ist also für die obige Pseudogeometrie folgender Satz:

III<sub>5</sub>. Wenn für zwei Pseudodreiecke  $ABC$  und  $A_1B_1C_1$  die Pseudokongruenzen

$$(1) \quad \overline{AB} \equiv \overline{A_1B_1}, \quad \overline{AC} \equiv \overline{A_1C_1}, \quad \sphericalangle BAC \equiv \sphericalangle B_1A_1C_1$$

gelten, so ist auch stets die Pseudokongruenz

$$(2) \quad \sphericalangle ABC \equiv \sphericalangle A_1B_1C_1$$

erfüllt.

Zum Beweise dieses Satzes schicken wir drei Hilfssätze voran.

HILFSSATZ 1. Liegen die Pseudopunkte  $P_1, P_2$  nicht auf einer Orthogonalhalbgeraden und sind  $II_1, II_2$  ihre Orthogonalprojektionen auf der Fundamentalebene  $\gamma$ , so gilt für die Charakteristik der Pseudostrecke  $\overline{P_1P_2}$  die Formel

$$(3) \quad (\Xi P_2 P_1)^2 = (\Xi H II_2 II_1).$$

Da nämlich nach dem Satze des THALES  $\sphericalangle \Xi P_1 H$  ein rechter ist, so gilt bekanntlich

$$\Xi P_1^2 = \Xi II_1 \cdot \Xi H, \quad P_1 H^2 = II_1 H \cdot \Xi H$$

(Fig. 1), also ist

$$\left( \frac{\Xi P_1}{P_1 H} \right)^2 = \frac{\Xi II_1}{II_1 H}.$$

Aus ähnlichem Grunde besteht

$$\left( \frac{\Xi P_2}{P_2 H} \right)^2 = \frac{\Xi II_2}{II_2 H},$$

und aus diesen beiden Formeln folgt (3) durch Division.

HILFSSATZ 2. Liegen die Pseudopunkte  $P_1, P_2$  nicht auf einer Orthogonalhalbgeraden und ist  $\Sigma$  ein Punkt des Randkreises  $k$  einer durch  $P_1, P_2$  gelegten Orthogonalhalbkugel, so ist bei einer stereographischen Projektion dieser Halbkugel von  $\Sigma$  aus, das Bild der Pseudostrecke  $\overline{P_1P_2}$  eine Pseudostrecke  $\overline{P'_1P'_2}$ , für die die Pseudokongruenz

$$(4) \quad \overline{P_1P_2} \equiv \overline{P'_1P'_2}$$

gilt.

Der erste Teil dieses Hilfssatzes, daß nämlich  $\overline{P_1P_2}$  wieder in eine Pseudostrecke  $\overline{P'_1P'_2}$  übergeführt wird, folgt aus dem elementargeometrischen Satze, laut welchem bei einer stereographischen Projektion der Kugel Kreise in Kreise übergehen (die Gerade als eine Ausartung des Kreises betrachtend) und

<sup>3</sup> D. HILBERT, a. a. O. <sup>2</sup>, S. 14.



die Winkel von Kreisen unverändert bleiben.<sup>4</sup> Ist nämlich  $\tau$  die Bildebene und  $g$  die Schnittgerade von  $\tau$  und  $\gamma$ , so ist das Bild des Kreises  $k$  diese Gerade  $g$ , also das vom Orthogonalhalbkreis  $\overline{\Xi H}$ , der  $k$  rechtwinklig schneidet, ein zu  $g$  orthogonaler Kreisbogen  $\overline{\Xi' H'}$ , d. h. ein Orthogonalhalbkreis (der in eine Orthogonalhalbgerade ausartet, falls  $\Sigma$  mit  $\Xi$  oder  $H$  zusammenfällt), und somit das Bild von  $\overline{P_1 P_2}$  eine Pseudostrecke  $\overline{P'_1 P'_2}$  (Fig. 4). |

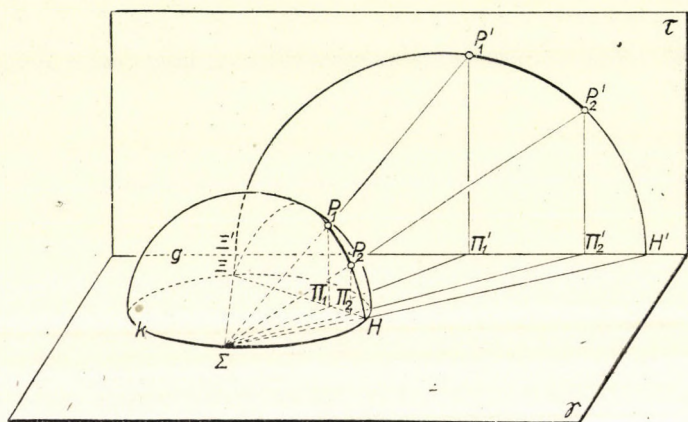


Fig. 4.

Die Pseudokongruenz (4) ergibt sich so. Ist zunächst  $\Sigma$  von  $\Xi$  und  $H$  verschieden, so sind die Bilder  $\Pi'_1, \Pi'_2$  von  $\Pi_1, \Pi_2$  offenbar die Orthogonalprojektionen von  $P'_1, P'_2$  auf der Fundamentelebene  $\gamma$ , also neben (3) besteht nach demselben Hilfssatz 1 auch

$$(5) \quad (\Xi' H' P'_2 P'_1)^2 = (\Xi' H' \Pi'_2 \Pi'_1),$$

wobei  $\Xi', H'$  die Bilder von  $\Xi, H$  sind. Da aber nach dem Satze des PAPPUS<sup>5</sup>

$$(\Xi H \Pi_2 \Pi_1) = (\Xi' H' \Pi'_2 \Pi'_1)$$

ausfällt, so folgt aus (3) und (5)

$$(\Xi H P_2 P_1) = (\Xi' H' P'_2 P'_1).$$

Die Pseudostrecken  $\overline{P_1 P_2}$  und  $\overline{P'_1 P'_2}$  haben demnach dieselbe Charakteristik, und somit ist (4) mit Rücksicht auf die obige Erklärung, für den betrachteten Fall erwiesen.

Fällt  $\Sigma$  z. B. mit  $\Xi$  zusammen, so ist das Bild von  $\overline{\Xi H}$  eine Orthogonalhalbgerade mit dem Anfangspunkt  $H'$ , und wir haben (Fig. 5)

$$\triangle \Xi H P_1 \sim \triangle \Xi P'_1 H' \quad \text{und} \quad \triangle \Xi H P_2 \sim \triangle \Xi P'_2 H',$$

<sup>4</sup> Siehe z. B. M. ZACHARIAS, *Elementargeometrie der Ebene und des Raumes* (Berlin und Leipzig, 1930), § 56.

<sup>5</sup> Siehe z. B. M. ZACHARIAS, a. a. O. <sup>4</sup>, S. 95.



also

$$\frac{P_1 H}{\Xi P_1} = \frac{H' P'_1}{\Xi H'}, \quad \frac{P_2 H}{\Xi P_2} = \frac{H' P'_2}{\Xi H'},$$

woher

$$(\Xi H P_2 P_1) = \frac{\Xi P_2}{P_2 H} : \frac{\Xi P_1}{P_1 H} = \frac{H' P'_1}{H' P'_2} = (H' P'_1 P'_2)$$

folgt. Das bedeutet definitionsgemäß wieder das Bestehen von (4). In ähnlicher Weise ergibt sich diese Pseudokongruenz, falls  $\Sigma$  mit  $H$  zusammenfällt.

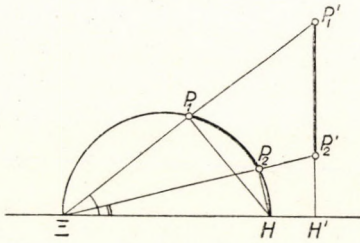


Fig. 5.

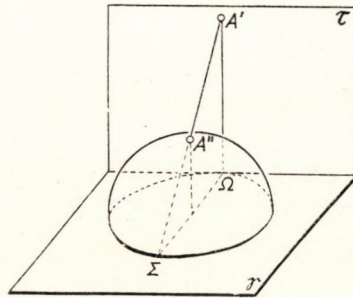


Fig. 6.

**HILFSSATZ 3.** *Ist die Pseudoebene eines Pseudodreiecks  $ABC$  eine Orthogonalhalbkugel, so ist bei einer stereographischen Projektion dieser Halbkugel von einem Punkte ihres Randkreises aus, das Bild von  $ABC$  wieder ein Pseudodreieck  $A'B'C'$ , welches ihm pseudokongruent ausfällt.*

Das ist eine unmittelbare Folge des Hilfssatzes 2, da doch bei dieser stereographischen Projektion die Winkel von Kreisen unverändert bleiben, also die Pseudowinkel von  $ABC$  in solche Pseudowinkel übergehen, die mit diesen gemäß der Definition der Pseudokongruenz der Reihe nach pseudokongruent sind.

Der obige Satz III<sub>5</sub> läßt sich nun leicht beweisen. Ist die Pseudoebene des Pseudodreiecks  $ABC$  eine Orthogonalhalbkugel, so werde sie von einem Punkte ihres Randkreises aus, stereographisch auf eine Ebene  $\tau$  projiziert (die also zur Berührungsebene an die Vollkugel im Gegenpunkte des Projektionszentrums parallel ist). Das Pseudodreieck  $ABC$  hat dabei laut Hilfssatz 3 ein ihm pseudokongruentes Bild  $A'B'C'$ . Da  $\tau$  auf die Fundamentelebene  $\gamma$  senkrecht steht, so liegt das von  $A'$  auf  $\gamma$  gefällte Lot  $A'\Omega$  in der Ebene  $\tau$  und das im Fußpunkt  $\Omega$  auf  $\tau$  errichtete Lot  $\Omega\Sigma = A'\Omega$  in der Ebene  $\gamma$ . Die Orthogonalhalbkugel mit dem Durchmesser  $\Omega\Sigma$  wird von  $\tau$  im Punkte  $\Omega$  berührt, da doch nach der Konstruktion  $\tau$  in  $\Omega$  auf  $\Omega\Sigma$  senkrecht steht. Wird demnach das Pseudodreieck  $A'B'C'$  von  $\Sigma$  aus auf diese Halbkugel projiziert, so geht es auf Grund des Hilfssatzes 3 in ein ihm pseudokon-



gruentes  $A''B''C''$  über, und  $A''$  ist wegen  $\Omega\Sigma = A'\Omega$  offenbar der Gipfel-  
punkt dieser Halbkugel (Fig. 6). Ist die Pseudoebene von  $ABC$  eine Ortho-  
gonalhalbebene, so bleibt die erste Projektion aus, und  $A'B'C'$  ist nur eine  
andere Bezeichnung für dieses Pseudodreieck. Jedenfalls kann also  $ABC$  in  
ein ihm pseudokongruentes Pseudodreieck  $A''B''C''$  übergeführt werden der-  
art, daß letzteres auf einer Orthogonalhalbkugel mit dem Gipfel-  
punkt  $A''$  liegt. Aus  $A_1B_1C_1$  entsteht in ähnlicher Weise das ihm pseudokongruente Pseudo-  
dreieck  $A_1'B_1'C_1''$ , und infolge der Voraussetzung (1) bestehen für die so her-  
gestellten Pseudodreiecke  $A''B''C''$  und  $A_1'B_1'C_1''$  (Fig. 7) die Pseudokon-  
gruenzen

$$\overline{A''B''} \equiv \overline{A_1'B_1''}, \quad \overline{A''C''} \equiv \overline{A_1'C_1''}, \quad \sphericalangle B''A''C'' \equiv \sphericalangle B_1'A_1''C_1''.$$

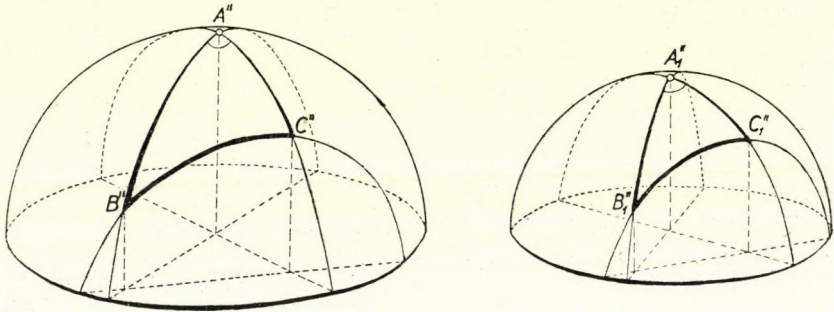


Fig. 7.

Da aber die Halbkugeln einander ähnlich, um ihre Gipfel-  
punkte in sich drehbar und in Bezug auf einen Hauptkreis durch den Gipfel-  
punkt symmetrisch sind, so folgt aus diesen Pseudokongruenzen offenbar die Pseudo-  
kongruenz

$$(6) \quad \sphericalangle A''B''C'' \equiv \sphericalangle A_1'B_1''C_1''.$$

Andererseits bestehen nach der Konstruktion

$$\sphericalangle ABC \equiv \sphericalangle A''B''C'', \quad \sphericalangle A_1B_1C_1 \equiv \sphericalangle A_1'B_1''C_1''.$$

Aus (6) folgt demnach die Pseudokongruenz (2), was zu beweisen war.

Solchergestalt ist auf elementargeometrischem Wege bewiesen, daß der  
Poincarésche Halbraum ein Modell des hyperbolischen Raumes darstellt, und  
somit ein elementargeometrischer Beweis für die Widerspruchsfreiheit der  
hyperbolischen Raumgeometrie erbracht.

(Eingegangen am 29. April 1954.)



## ЭЛЕМЕНТАРНО-ГЕОМЕТРИЧЕСКОЕ ДОКАЗАТЕЛЬСТВО НЕПРОТИВОРЕЧИВОСТИ ГИПЕРБОЛИЧЕСКОЙ ПРОСТРАНСТВЕННОЙ ГЕОМЕТРИИ, ПОСРЕДСТВОМ ПОЛУПРОСТРАНСТВА ПУАНКАРЕ

П. САС (Будапешт)

(Резюме)

Чтобы дать реализацию гиперболической пространственной геометрии, Г. Пуанкаре<sup>1</sup> создал известную модель полупространства, который мы здесь назовем псевдогеометрией.

В этой псевдогеометрии (которую нетрудно описать также с помощью элементарно-геометрических понятий) выполняются аксиомы связи, аксиомы порядка и аксиомы конгруэнтности, заданные Д. Гильбертом,<sup>2</sup> противоположная эвклидовой аксиоме параллельных, аксиома Архимеда и канторова аксиома непрерывности, т. е. полная система аксиом гиперболической геометрии. В настоящей работе это доказывается чисто элементарно-геометрическим путём. Можно при этом ограничиться доказательством аксиомы Ш<sub>5</sub> конгруэнтности,<sup>3</sup> ибо справедливость других аксиом непосредственно видна. Итак, нам достаточно доказать, что в этой псевдогеометрии имеет место следующая теорема:

Ш<sub>5</sub>. Если для псевдотреугольников  $ABC$  и  $A_1B_1C_1$  выполняются псевдоконгруенции

$$\overline{AB} \equiv \overline{A_1B_1}, \quad \overline{AC} \equiv \overline{A_1C_1}, \quad \sphericalangle BAC \equiv \sphericalangle B_1A_1C_1,$$

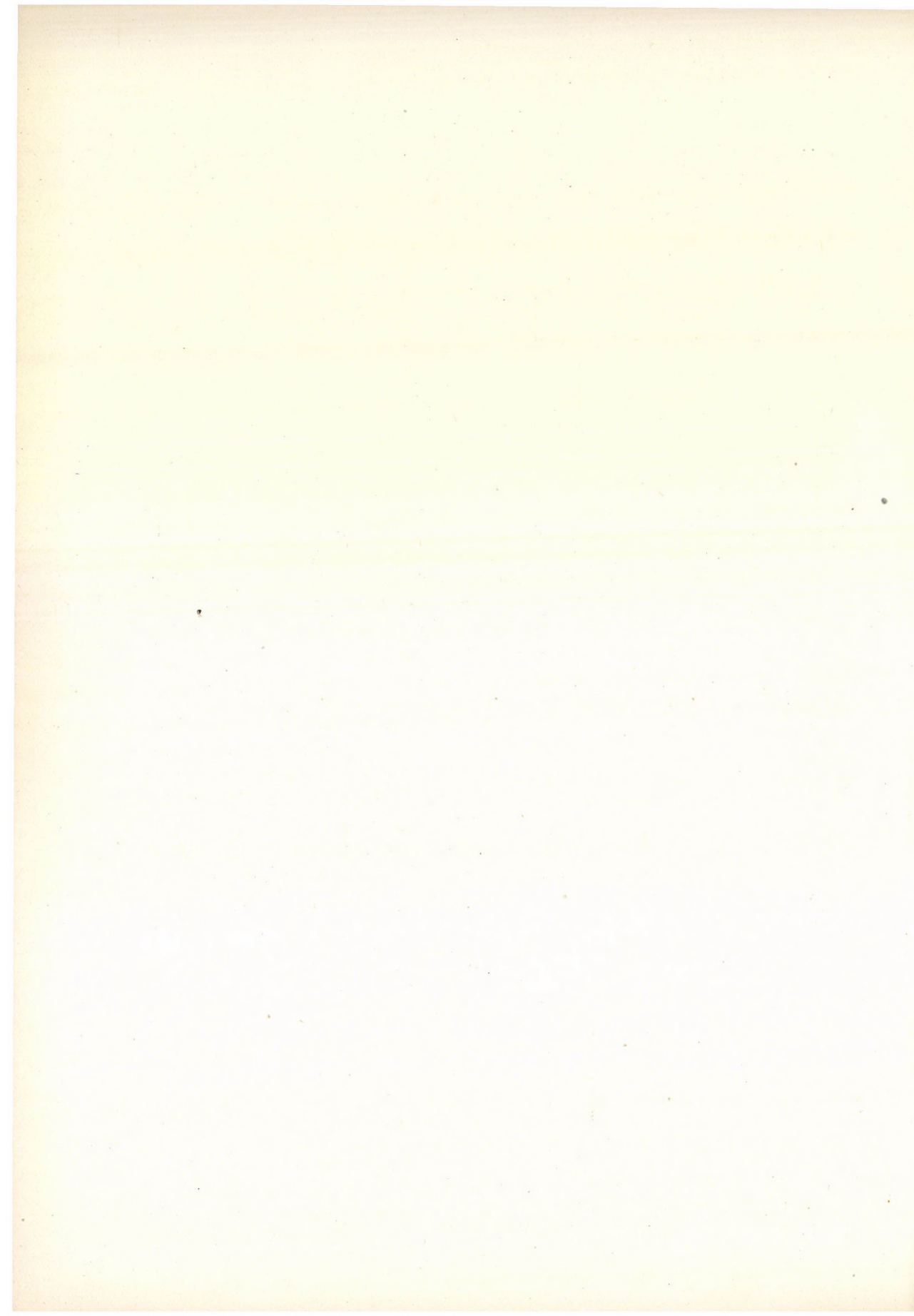
то выполняется и псевдоконгруенция

$$\sphericalangle ABC \equiv \sphericalangle A_1B_1C_1.$$

Доказательство использует элементарно-геометрическую теорему о том, что при стереографической проекции сферы всякая окружность переходит в окружность (считая при этом прямую вырожденной окружностью), углы же между окружностями сохраняются.<sup>4</sup>

Таким образом мы получаем элементарно-геометрическое доказательство того факта, что полупространство Пуанкаре может служить моделью гиперболического пространства, т. е. добьемся элементарно-геометрического доказательства непротиворечивости гиперболической пространственной геометрии.







# ARRANGEMENTS OF EQUAL SPHERES IN NON-EUCLIDEAN SPACES

By

H. S. M. COXETER (Toronto)

(Presented by G. HAJÓS)

## 1. Introduction

We consider two kinds of arrangements of equal spheres: *packings*, where no point is inside more than one sphere, and *coverings*, where no point is outside every sphere. Each arrangement determines a honeycomb whose cells are called *Dirichlet regions*. The Dirichlet region for any particular sphere is a polytope whose interior consists of all points that are nearer to the centre of that sphere than to the centre of any other [13, pp. 216—219]. If all the Dirichlet regions are of the same size, the ratio of the content of the sphere to the content of the Dirichlet region is called the *density* of the arrangement. If they are of various sizes, we use a suitable average. We are interested in cases where the density is as near to 1 as possible: close packings and thin coverings. In such cases the spheres are respectively inscribed in, and circumscribed about, the Dirichlet regions.

The two-dimensional results were obtained by FEJES TÓTH, who proved that the elliptic plane cannot be packed as closely, nor covered as thinly, as the Euclidean plane, and that no arrangements of finite circles in the hyperbolic plane are as "good" as the best arrangement of horocycles. Elliptic 3-space presents a different state of affairs: sixty spheres of radius  $\pi/10$ , each touching twelve others, form a packing of density 0,774, and somewhat larger spheres (having the same centres) form a covering of density 1,439, whereas in Euclidean space the maximum packing-density and minimum covering density are almost certainly 0,740 and 1,464 [14, pp. 171, 174].

In hyperbolic 3-space, no superior arrangements of finite spheres are known, but horospheres provide a packing of density 0,853 and a covering of density 1,280.

Finally, hyperbolic 4-space admits remarkable arrangements of finite spheres: a packing of density 0,691 and a covering of density 1,558, whereas the best densities known for Euclidean 4-space are 0,617 and 1,766.

In the course of these investigations we use some results on hyperbolic geometry (§§ 4 and 5) that may also be of some intrinsic interest.



## 2. Two-dimensional arrangements

Most of the results in this section have already been given in a different notation by SCHLEGEL [20, p. 360] and FEJES TÓTH [15; 16], but are included here for the sake of completeness.

Let  $\{p, q\}$  denote the regular tessellation of  $p$ -gons,  $q$  at each vertex, filling a Euclidean or non-Euclidean plane. Consider the characteristic triangle  $P_0P_1P_2$  formed by a vertex, the mid-point of one of the edges meeting at this vertex, and the centre of either of the faces meeting at this edge. Since the angles at  $P_0, P_1, P_2$  are

$$\pi/q, \quad \pi/2, \quad \pi/p$$

the plane must be elliptic, Euclidean, or hyperbolic according as the number

$$2p + 2q - pq = 4 - (p-2)(q-2)$$

is positive, zero, or negative.

Let  $\psi$  and  $\chi$  denote the in-radius and circum-radius of a face, so that

$$\psi = P_1P_2, \quad \chi = P_0P_2.$$

The formulae for a right-angled triangle yield

$$\cos \psi = \operatorname{cosec} \frac{\pi}{p} \cos \frac{\pi}{q}, \quad \cos \chi = \operatorname{ctg} \frac{\pi}{p} \operatorname{ctg} \frac{\pi}{q}$$

in the elliptic case [9, pp. 21, 24], and the same expressions for  $\operatorname{ch} \psi$ ,  $\operatorname{ch} \chi$  in the hyperbolic case.

It is known [14, p. 58] that the in- and circum-circles of the faces of the Euclidean tessellation  $\{6, 3\}$  form the closest packing and thinnest covering of equal circles in the Euclidean plane. Similarly, the in- and circum-circles of the faces of  $\{p, 3\}$  ( $p < 6$  or  $p > 6$ ), in the elliptic or hyperbolic plane, form a closer packing and a thinner covering than any other arrangement of circles of respective radii  $\psi$  and  $\chi$ , where

$$\cos \psi \text{ or } \operatorname{ch} \psi = \frac{1}{2} \operatorname{cosec} \frac{\pi}{p}, \quad \cos \chi \text{ or } \operatorname{ch} \chi = 3^{-\frac{1}{2}} \operatorname{ctg} \frac{\pi}{p}.$$

Clearly, the face of  $\{p, 3\}$  is a  $p$ -gon of angle  $2\pi/3$ . Comparing its area

$$\frac{1}{2}(p-6)\pi/3$$

with the area

$$2\pi \left| \frac{\cos \psi - 1}{\operatorname{ch} \psi - 1} \right| = \pi \left| \operatorname{cosec} \frac{\pi}{p} - 2 \right|$$

of its in-circle and with the area

$$2\pi \left| \frac{\cos \chi - 1}{\operatorname{ch} \chi - 1} \right| = \frac{2\pi}{\sqrt{3}} \left| \operatorname{ctg} \frac{\pi}{p} - \sqrt{3} \right|$$

of its circum-circle, we obtain the densities

$$d(p) = \frac{3}{p-6} \left( \operatorname{cosec} \frac{\pi}{p} - 2 \right) \quad \text{and} \quad D(p) = \frac{2\sqrt{3}}{p-6} \left( \operatorname{ctg} \frac{\pi}{p} - \sqrt{3} \right)$$



for the packing and covering, respectively. Taking the limits as  $p$  tends to 6, we obtain the densities

$$d(6) = \frac{\pi}{2\sqrt{3}} = 0,9069\dots \quad \text{and} \quad D(6) = \frac{2\pi}{3\sqrt{3}} = 1,2092\dots$$

for the closest packing and thinnest covering of the Euclidean plane. Taking instead the limits as  $p$  tends to infinity, we obtain the densities

$$d(\infty) = \frac{3}{\pi} = 0,9549\dots \quad \text{and} \quad D(\infty) = \frac{2\sqrt{3}}{\pi} = 1,1026\dots$$

for the packing and covering of the hyperbolic plane by the in- and circumscribed horocycles of the faces of the limiting tessellation  $\{\infty, 3\}$ , whose rotation group is the modular group  $S^3 = T^2 = 1$  [12, p. 425; 15, p. 107; 16, p. 113].

Since  $d(p)$  is an increasing function, every packing of equal circles in the elliptic plane (or of equal small circles on a sphere) has density less than  $d(6)$ , and every packing of equal circles in the hyperbolic plane [15, p. 106] has density less than  $d(\infty)$ .

Since  $D(p)$  is a decreasing function, every covering of the elliptic plane by equal circles (or of a sphere by equal small circles) has density greater than  $D(6)$ , and every covering of the hyperbolic plane by equal circles [16, p. 112] has density greater than  $D(\infty)$ .

### 3. Three-dimensional arrangements

The symbol  $\{p, q\}$ , which we have used for a spherical tessellation when  $(p-2)(q-2) < 4$ , is naturally used also for the corresponding regular polyhedron (or Platonic solid) in Euclidean or non-Euclidean space. Repetitions of this solid, of such a size that  $r$  of them can be fitted round a common edge, form a regular honeycomb  $\{p, q, r\}$  [9, p. 138].

Consider the *characteristic tetrahedron*  $P_0P_1P_2P_3$  formed by a vertex, the mid-point of an incident edge, the centre of an incident plane face, and the centre of an incident solid cell. Since the dihedral angles at the edges  $P_2P_3, P_0P_3, P_0P_1$  are

$$\pi/p, \quad \pi/q, \quad \pi/r,$$

while the remaining three are right angles, the space must be elliptic, Euclidean, or hyperbolic according as

$$\sin \frac{\pi}{p} \sin \frac{\pi}{r} - \cos \frac{\pi}{q}$$

is positive, zero, or negative [9, p. 135].

The in-radius and circum-radius of a cell are evidently

$$\psi = P_2P_3, \quad \chi = P_0P_3$$



[9, p. 139]. When the space is hyperbolic, we might allow the vertices to lie on the absolute quadric, so that  $\chi$  is infinite; e. g., the tetrahedron, octahedron and hexahedron described in an earlier paper [6, p. 29] are cells of  $\{3, 3, 6\}$ ,  $\{3, 4, 4\}$  and  $\{4, 3, 6\}$ , respectively.

The problems of close packing and thin covering in Euclidean space have not yet been completely solved. The difficulty may be attributed to the fact that, since the equation

$$\sin \frac{\pi}{p} \sin \frac{\pi}{3} - \cos \frac{\pi}{3} = 0$$

yields a value of  $p$  between 5 and 6, there is no Euclidean honeycomb of type  $\{p, 3, 3\}$ . (There is an obviously closest arrangement for four solid spheres, but this cannot be regularly continued.)

This difficulty disappears in the non-Euclidean spaces, where we have the elliptic (or spherical)  $\{5, 3, 3\}$  and the hyperbolic  $\{6, 3, 3\}$  [12, pp. 426, 428]. The former consists of 60 regular dodecahedra  $\{5, 3\}$  filling the elliptic space (or 120 filling the spherical space) [9, pp. 138, 176, 273; 10, p. 116; 11, p. 471; 12, p. 425]. The in- and circum-spheres of these dodecahedra provide a packing of spheres of radius

$$\psi = \pi/10$$

[9, p. 293] and a covering by spheres of radius

$$\chi = \frac{1}{3} \pi - \eta$$

where

$$\eta = \frac{1}{2} \arccos \frac{1}{4}.$$

Dividing the volumes

$$\pi(2\psi - \sin 2\psi) \quad \text{and} \quad \pi(2\chi - \sin 2\chi)$$

of these spheres by the volume

$$\pi^2/60$$

of the dodecahedral cell, we obtain the density

$$\frac{60}{\pi} (2\psi - \sin 2\psi) = \frac{60}{\pi} \left( \frac{\pi}{5} - \sin \frac{\pi}{5} \right) = 12 \left( 1 - \frac{5}{\pi} \sin \frac{\pi}{5} \right) = 0,774 \dots$$

for the packing, and

$$\frac{60}{\pi} (2\chi - \sin 2\chi) = \frac{60}{\pi} \left( \frac{2\pi}{3} - 2\eta - \frac{\sqrt{3}}{4} \tau \right) = 40 - \frac{15}{\pi} (8\eta + \sqrt{3}\tau) = 1,439 \dots$$

(where  $\tau = \frac{\sqrt{5}+1}{2}$ ) for the covering.

Since the cells are dodecahedra, each sphere in the packing touches twelve others [14, p. 177, Fig. 121]: the same number as in the familiar



close-packings of Euclidean space [14, pp. 171—172, Figs. 116, 118], whose density

$$\frac{\pi}{3\sqrt{2}} = 0,74048 \dots$$

is probably the greatest that can be maintained throughout the whole space, although greater densities are possible in a tubular region [2, pp. 308—310]. FEJES TÓTH [14, pp. 175—176] has shown that the density of the closest possible packing is almost certainly less than

$$\frac{\pi}{14} 5^{-1/4} \tau^{7/2} = 0,7545 \dots,$$

an amount still far short of the above elliptic density, 0,774....

Similarly, the elliptic covering density, 1,439..., compares favourably with

$$\frac{5\sqrt{5}\pi}{24} = 1,464 \dots,$$

which is the density of the covering of Euclidean space by the circum-spheres of the cells (truncated octahedra) of the honeycomb  $t_{1,2}\delta_4$  [7, p. 403; 14, p. 174]. BAMBAH [1, p. 26] has proved that this covering, in which each sphere intersects fourteen others, is the thinnest covering by spheres whose centres form a lattice (in this case, the body-centred cubic lattice); but the possibility of thinner irregular coverings remains open.

#### 4. A digression on the hyperbolic sine and tangent

Before investigating  $\{6, 3, 3\}$ , we need two lemmas about hyperbolic geometry. The first, which was proved quite differently by SOMMERVILLE [21, pp. 62, 77], may be regarded as the hyperbolic counterpart of the Euclidean

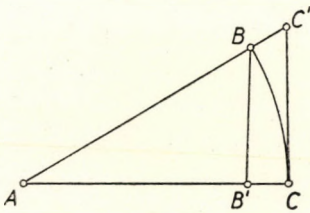


Fig. 1

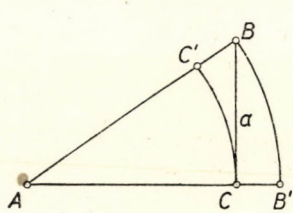


Fig. 2

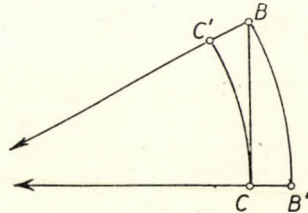


Fig. 3

definition for the sine and tangent of an arc of a unit circle: if lines  $BB'$  and  $CC'$  are drawn through the ends of an arc  $BC$ , perpendicular to the radius  $AC$ , so that  $B'$  lies on  $AC$ , and  $C'$  on  $AB$  produced (as in Fig. 1), then

$$BB' = \sin BC, \quad CC' = \operatorname{tg} BC.$$



In the hyperbolic plane, let arcs  $BB'$  and  $CC'$  of two circles with centre  $A$  be drawn through the vertices  $B$  and  $C$  of a right triangle  $ABC$ , so that  $B'$  lies on  $AC$  produced, and  $C'$  on  $AB$  (as in Fig. 2). The sides  $a = BC$ ,  $b = CA$ ,  $c = AB$  are connected by the classical formulae [8, p. 238]

$$\operatorname{sh} a = \sin A \operatorname{sh} c, \quad \operatorname{th} a = \operatorname{tg} A \operatorname{sh} b,$$

while the two arcs (of radii  $c$  and  $b$ ) are

$$(4.1) \quad BB' = A \operatorname{sh} c, \quad CC' = A \operatorname{sh} b.$$

Hence

$$BB' = \frac{A}{\sin A} \operatorname{sh} a, \quad CC' = \frac{A}{\operatorname{tg} A} \operatorname{th} a.$$

Taking the limit as  $A$  tends to zero, so that the triangle becomes singly-asymptotic (as in Fig. 3), we obtain the desired lemma:

*If  $BB'$  and  $C'C$  are corresponding arcs of two concentric horocycles, so situated that the tangent at  $C$  passes through  $B$ , then*

$$BB' = \operatorname{sh} BC, \quad CC' = \operatorname{th} BC.$$

### 5. The volume of a sector of a horosphere

Our second lemma is the solid analogue of another theorem of plane hyperbolic geometry: the area of a horocyclic sector is equal to its arc [8, p. 250].

Here it is to be understood that we are measuring length and area in terms of the natural units of LOBATSCHESKY: the unit arc of a horocycle is such that the tangent at one end is parallel to the diameter through the other end (produced outwards), and the unit area is bounded by this arc and the (parallel) diameters at its two ends.

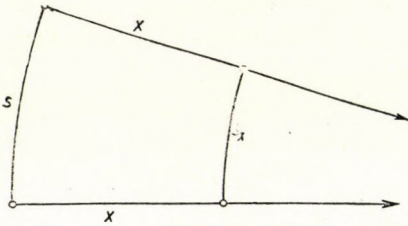


Fig. 4

Consider the "rectangle" bounded by two corresponding arcs of concentric horocycles and the intercepted segments of the diameters through their ends (as in Fig. 4). If the straight sides are both  $x$ , while the curved sides are  $s$  and  $s_r$  ( $s > s_r$ ), we have from (4.1)

$$(5.1) \quad \frac{s_r}{s} = \lim_{c \rightarrow \infty} \frac{\operatorname{sh}(c-x)}{\operatorname{sh} c} = e^{-x}$$

[3, p. 95]. It follows that the area of the horocyclic sector is

$$\int_0^{\infty} s_r dx = s \int_0^{\infty} e^{-x} dx = s.$$

Passing now to hyperbolic space, we observe that the surface of a horosphere is developable on the Euclidean plane [4, p. 205; 8, pp. 220, 251].



Accordingly, the natural unit of area on a horosphere is the "square" formed by four unit horocyclic arcs. Consider the "solid sector" bounded by this "square" and portions of the diametral planes through its sides. The section of this "solid sector" by a concentric horosphere, distant  $x$  from the bounding horosphere, is a "square" of side  $e^{-x}$  (by (5.1) with  $s=1$ ). Hence the volume of the "solid sector" is

$$\int_0^{\infty} (e^{-x})^2 dx = \int_0^{\infty} e^{-2x} dx = \frac{1}{2}.$$

Since any area can be expressed as a numerical multiple of the unit square, we deduce the desired lemma:

*The volume of any "solid sector" of a horosphere is equal to one-half the area of its horospherical boundary.*

This result can evidently be generalized to  $n$  dimensions, with  $1/(n-1)$  in place of  $\frac{1}{2}$ .

## 6. Three-dimensional arrangements, continued

Since  $\{6, 3\}$  is a Euclidean tessellation, the cell of the hyperbolic honeycomb  $\{6, 3, 3\}$  is an infinite polyhedron having an inscribed horosphere and a concentric circumscribed horosphere. Analogy with the above discussion of  $\{\infty, 3\}$  would suggest that every packing of equal spheres is less dense than the arrangement of horospheres inscribed in the cells of  $\{6, 3, 3\}$ , and that every covering of hyperbolic space by equal spheres is denser than the arrangement of horospheres circumscribed about these same cells.

Since the cells are infinite, the densities are most conveniently computed by comparing the singly-asymptotic characteristic tetrahedron  $P_0P_1P_2P_3$  ( $P_3$  at infinity) with the corresponding solid sectors of the respective horospheres.

The inscribed horosphere, touching a hexagonal face of the cell at its centre  $P_2$ , cuts out from the tetrahedron  $P_0P_1P_2P_3$  a horospherical triangle  $P_0P_1P_2$ , whose angles are  $\pi/3$ ,  $\pi/2$ ,  $\pi/6$ . The plane triangle  $P_0P_1P_2$  has likewise a right angle at  $P_1$  and  $\pi/6$  at  $P_2$ , while its side  $P_0P_1$  is

$$\varphi = \frac{1}{2} \log 2$$

[12, p. 427], so that  $\text{th } \varphi = \frac{1}{3}$ . Setting  $a = P_0P_1$ ,  $b = P_1P_2$ ,  $A = \pi/6$  in the formula

$$\text{sh } b = \frac{\text{th } a}{\text{tg } A}$$



for a right-angled triangle  $ABC$ , we obtain

$$\operatorname{sh} P_1 P_2 = \sqrt{\frac{1}{3}}.$$

Using the lemma in § 4, we see that this side  $P_1 P_2$  is the tangent corresponding to the horocyclic arc

$$P'_1 P_2 = \operatorname{th} P_1 P_2 = \frac{1}{2}.$$

By Euclidean mensuration, the area of the horocyclic triangle  $P'_0 P'_1 P_2$  is

$$\frac{1}{8\sqrt{3}}.$$

By the lemma in § 5, the volume of the corresponding solid sector is

$$\frac{1}{16\sqrt{3}}.$$

Similarly we have, on the circumscribed horosphere through  $P_0$ , a horospherical triangle  $P_0 P'_1 P''_2$  whose angles are again  $\pi/3, \pi/2, \pi/6$ . The side  $P_0 P_1 = \varphi$  of the corresponding plane triangle is the "sine" of the horocyclic arc

$$P_0 P'_1 = \operatorname{sh} \varphi = \frac{1}{2\sqrt{2}}.$$

Hence the area of  $P_0 P'_1 P''_2$  is  $\sqrt{3}/16$ , and the volume of the corresponding solid sector is

$$\frac{\sqrt{3}}{32}.$$

Using an observation of SCHLÄFLI [19, p. 162; see also 6, pp. 23, 27, 29], we obtain the volume of the singly-asymptotic tetrahedron  $P_0 P_1 P_2 P_3$  in the form

$$\begin{aligned} \frac{i}{4} S\left(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{6}\right) &= \frac{i}{16} S\left(\frac{\pi}{6}, \frac{\pi}{6}, \frac{\pi}{6}\right) = \frac{i}{24} S\left(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3}\right) = \\ &= \frac{\sqrt{3}}{32} \left(1 - \frac{1}{2^2} + \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} - \frac{1}{8^2} + \dots\right) = \\ &= \frac{1}{16\sqrt{3}} \left(1 + \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} + \frac{1}{8^2} - \dots\right). \end{aligned}$$

Hence the density of the packing is

$$\left(1 + \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} + \frac{1}{8^2} - \dots\right)^{-1} = 0,853\dots,$$

and the density of the covering is

$$\left(1 - \frac{1}{2^2} + \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} - \frac{1}{8^2} + \dots\right)^{-1} = 1,280\dots$$



When P. SCHERK saw the ratio of these two series, he proved a generalization in which the squares are replaced by sth powers. Setting

$$f(s) = 1 - \frac{1}{2^s} + \frac{1}{4^s} - \frac{1}{5^s} + \frac{1}{7^s} - \frac{1}{8^s} + \dots,$$

$$g(s) = 1 + \frac{1}{2^s} - \frac{1}{4^s} - \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{8^s} - \dots,$$

he observed that, if  $s > 1$ ,

$$g(s) - f(s) = 2 \left( \frac{1}{2^s} - \frac{1}{4^s} + \frac{1}{8^s} - \frac{1}{10^s} + \frac{1}{14^s} - \frac{1}{16^s} + \dots \right) = 2^{1-s} f(s),$$

whence

$$g(s) = (1 + 2^{1-s}) f(s)$$

and, in particular,

$$g(2) = \frac{3}{2} f(2).$$

## 7. Arrangements in hyperbolic 4-space

We saw in § 3 that, when

$$\sin \frac{\pi}{p} \sin \frac{\pi}{r} > \cos \frac{\pi}{q},$$

the honeycomb  $\{p, q, r\}$  fills elliptic or spherical 3-space. Taking the latter interpretation, we use the symbol  $g_{p,q,r}$  to denote the number of repetitions of the characteristic tetrahedron  $P_0P_1P_2P_3$  that will fill the whole spherical space, so that the volume of each is  $2\pi^2/g_{p,q,r}$ . This is a rather complicated function of  $p, q, r$ . For our present purpose it will suffice to note that

$$g_{3,3,3} = 120, \quad g_{5,3,3} = 120^2 = 14400$$

[9, pp. 133, 153].

The "Schläfli symbol"  $\{p, q, r\}$  is naturally used also for the corresponding regular polytope in Euclidean or non-Euclidean 4-space. Repetitions of this polytope, of such a size that  $s$  of them can be fitted round a common plane face, form a regular honeycomb  $\{p, q, r, s\}$ , whose *characteristic simplex*  $P_0P_1P_2P_3P_4$  is formed by a vertex, the mid-point of an incident edge, and the centres of other incident elements of dimension 2, 3, 4 [9, p. 136]. Since the condition for the space to be hyperbolic is

$$\frac{\cos^2 \pi/q}{\sin^2 \pi/p} + \frac{\cos^2 \pi/r}{\sin^2 \pi/s} > 1,$$

the only instance having finite cells, with  $s=3$ , is

$$\{5, 3, 3, 3\}.$$

Accordingly it is a natural conjecture that the closest packing and thinnest



covering of hyperbolic 4-space are provided by the in- and circum-spheres of the cells of this honeycomb. Since  $\{5, 3, 3\}$  is bounded by 120 dodecahedra, each sphere has 120 neighbours.

It was shown by SCHLÄFLI [19, p. 283; cf. 9, p. 233] that the content of the characteristic simplex for  $\{p, q, r, s\}$  is

$$\frac{\pi^2}{6} \left( \frac{8}{g_{p,q,r}} + \frac{8}{g_{q,r,s}} + \frac{2}{ps} - \frac{1}{p} - \frac{1}{q} - \frac{1}{r} - \frac{1}{s} + 1 \right).$$

In the present case this becomes

$$\frac{\pi^2}{6} \left( \frac{8}{14400} + \frac{8}{120} + \frac{2}{15} - \frac{1}{5} \right) = \frac{\pi^2}{10800}.$$

Multiplying by  $g_{5,3,3} = 14400$ , we deduce that the content of a cell  $\{5, 3, 3\}$  of the honeycomb  $\{5, 3, 3, 3\}$  is

$$\frac{4}{3} \pi^2.$$

This is to be compared with the contents of spheres (or rather, 4-dimensional hyperspheres), whose radii,

$$\psi = P_3 P_4 \quad \text{and} \quad \chi = P_0 P_4,$$

are given by the formulae

$$\operatorname{sech}^2 \psi = (3, 5)(-1, 4)/(-1, 3), \quad \operatorname{sech}^2 \chi = (0, 5)(-1, 4)$$

[9, pp. 161, 162], where, in the case of  $\{5, 3, 3, 3\}$ ,

$$(j, k) = k - j \quad (0 \leq j \leq k)$$

and

$$(-1, k) = 1 - k\tau^{-3} \quad \left( \tau = \frac{\sqrt{5}+1}{2}, \tau^{-3} = \sqrt{5}-2 \right),$$

so that

$$(3, 5) = 2, \quad (-1, 4) = \tau^{-6}, \quad (-1, 3) = 2\tau^{-4}, \quad (0, 5) = 5,$$

$$\operatorname{ch} \psi = \tau, \quad \operatorname{ch} \chi = 5^{\frac{1}{2}} \tau^3.$$

A sphere of radius  $\varrho$  in spherical 4-space has the same 3-dimensional content as one of radius  $\sin \varrho$  in Euclidean 4-space, namely  $2\pi^2 \sin^3 \varrho$ . Therefore its 4-dimensional content is

$$\int_0^{\varrho} 2\pi^2 \sin^3 \theta d\theta = \frac{2}{3} \pi^2 (2 + \cos \varrho) (1 - \cos \varrho)^2.$$

It follows that the 4-dimensional content of a sphere of radius  $\varrho$  in hyperbolic 4-space is

$$\frac{2}{3} \pi^2 (2 + \operatorname{ch} \varrho) (1 - \operatorname{ch} \varrho)^2.$$



Setting  $\varrho = \psi$  and  $\chi$  in turn, we obtain

$$\frac{2}{3} \pi^2 (2 + \tau) (1 - \tau)^2 = \frac{2}{3} \pi^2 \sqrt{5} \tau^{-1}$$

and

$$\frac{2}{3} \pi^2 \left(2 + 5^{-\frac{1}{2}} \tau^3\right) \left(1 - 5^{-\frac{1}{2}} \tau^3\right)^2 = \frac{8}{3} \pi^2 \frac{2 + 3\sqrt{5}}{5\sqrt{5}}$$

for the contents of the in- and circum-spheres.

Dividing by  $\frac{4}{3} \pi^2$ , we obtain the packing density

$$\frac{1}{2} \sqrt{5} \tau^{-1} = \frac{5 - \sqrt{5}}{4} = 0,69098 \dots$$

and the covering density

$$\frac{2(2 + 3\sqrt{5})}{5\sqrt{5}} = 1,55777 \dots$$

It is interesting to compare these with the corresponding results in Euclidean 4-space, where the greatest known packing density

$$\frac{\pi^2}{16} = 0,61685 \dots$$

is attained by the in-spheres of the cells  $\{3, 4, 3\}$  of the regular honeycomb  $\{3, 4, 3, 3\}$  [9, pp. 153—154, 156, 296; 18, p. 42], while the smallest known covering density

$$\frac{2\pi^2}{5\sqrt{5}} = 1,76553 \dots$$

is attained by the circum-spheres of the cells  $t_{0,1,2,3}a_4$  of HINTON's honeycomb [5, p. 334].

## 8. Postscript

After the above sections were written, Professor FEJES TÓTH kindly sent me the galley proof of his own work on the same subject [17]. This shows that he anticipated my use of the in-spheres of the cells of  $\{5, 3, 3\}$  for a packing in elliptic 3-space, and of the inscribed horospheres of the cells of  $\{6, 3, 3\}$  for a packing in hyperbolic 3-space. He and I obtained the same expressions for the density in both cases.

(Received 18 June 1954)



## References

- [1] R. P. BAMBAH, On lattice coverings by spheres, *Proc. Nat. Inst. Sci. India*, **20** (1954), pp. 25—52.
- [2] A. H. BOERDIJK, Some remarks concerning close-packing of equal spheres, *Philips Research Reports*, **7** (1952), pp. 303—313.
- [3] H. S. CARSLAW, *The elements of non-Euclidean plane geometry and trigonometry* (London, 1916).
- [4] J. L. COOLIDGE, *The elements of non-Euclidean geometry* (Oxford, 1909).
- [5] H. S. M. COXETER, Wythoff's construction for uniform polytopes, *Proc. London Math. Soc.* (2), **38** (1935), pp. 327—339.
- [6] H. S. M. COXETER, The functions of Schläfli and Lobatschewsky, *Quart. J. Math.*, **6** (1935), pp. 13—29.
- [7] H. S. M. COXETER, Regular and semi-regular polytopes I, *Math. Z.*, **46** (1940), pp. 380—407.
- [8] H. S. M. COXETER, *Non-Euclidean geometry* (Toronto, 1947).
- [9] H. S. M. COXETER, *Regular polytopes* (London, 1948).
- [10] H. S. M. COXETER, Interlocked rings of spheres, *Scripta Mathematica*, **18** (1952), pp. 113—121.
- [11] H. S. M. COXETER, Regular honeycombs in elliptic space, *Proc. London Math. Soc.* (3), **4** (1954), pp. 471—501.
- [12] H. S. M. COXETER and G. J. WHITROW, World structure and non-Euclidean honeycombs, *Proc. Roy. Soc. A*, **201** (1950), pp. 417—437.
- [13] G. L. DIRICHLET, Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen, *J. reine angew. Math.*, **40** (1850), pp. 209—227.
- [14] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin, 1953).
- [15] L. FEJES TÓTH, Kreisausfüllungen der hyperbolischen Ebene, *Acta Math. Acad. Sci. Hung.*, **4** (1953), pp. 103—110.
- [16] L. FEJES TÓTH, Kreisüberdeckungen der hyperbolischen Ebene, *Acta Math. Acad. Sci. Hung.*, **4** (1953), pp. 111—114.
- [17] L. FEJES TÓTH, On close-packings of spheres in spaces of constant curvature, *Publicationes Mathematicae (Debrecen)*, **3** (1953), pp. 158—167.
- [18] D. HILBERT and S. COHN-VOSSEN, *Anschauliche Geometrie* (Berlin, 1932).
- [19] L. SCHLÄFLI, *Gesammelte mathematische Abhandlungen*, **2** (Basel, 1953).
- [20] V. SCHLEGEL, Theorie der homogenen zusammengesetzten Raumgebilde, *Verh. K. Leopold.-Carol. D. Akad. Naturf.*, **44** (1883), pp. 343—459.
- [21] D. M. Y. SOMMERVILLE, *The elements of non-Euclidean geometry* (London, 1914).

## РАЗМЕЩЕНИЕ КОНГРУЕНТНЫХ СФЕР В НЕЕВКЛИДОВЫХ ПРОСТРАНСТВАХ

X. C. M. КОКСЕТЕР (Торонто)

(Резюме)

Настоящая работа содержит описание некоторых способов расположения сфер, посредством которых получается сплошное заполнение и экономичное покрытие эллиптического и гиперболического пространств трех измерений, так же как и гиперболического пространства четырех измерений. Можно заполнять или покрывать эллиптическое пространство трех измерений с помощью 60 конгруэнтных сфер, с плотностью 0,774 и 1,439 соответственно. Гиперболическое пространство трех измерений допускает заполнение или покрытие с помощью горосфер, плотности 0,853 и 1,280 соответственно. Гиперболическое пространство четырех измерений допускает заполнение или покрытие с помощью конечных конгруэнтных сфер, плотности 0,691 и 1,558 соответственно.



## RESTGLIED EINES TAUBERSCHEN SATZES. III

Von

G. FREUD (Budapest)  
(Vorgelegt von A. RÉNYI)

### Einleitung

Eine Folge  $\{s_n\}$  sei Abel-limitierbar und  $A\text{-}\lim s_n = S$ . Bekanntlich ist dann  $\{s_n\}$  auch  $(C, 1)$ -summierbar, falls nur die Folge  $\{s_n\}$  einseitig beschränkt ist. In den ersten beiden Mitteilungen [1], [2] dieser Arbeit wurde folgende Verschärfung dieses Satzes gezeigt:

Es sei für  $s \rightarrow +0$

$$(1) \quad \sum_{n=0}^{\infty} s_n e^{-ns} = \frac{S}{s} [1 + O(R(s))]$$

und die Folge  $\{s_n\}$  sei einseitig beschränkt:

$$(2) \quad s_n > -K.$$

$R(s)$  bedeutet in Formel (1), und auch an jeder späteren Stelle eine monoton nicht abnehmende Funktion mit  $R(+0) = 0$ , die der Bedingung

$$(3) \quad R(ks) < e^{c_1 k} R(s), \quad k = 2, 3, \dots$$

genügt.<sup>1</sup>

Dann genügen die Cesàroschen Summen  $m$ -ter Ordnung  $\sigma_n^{(m)}$  der Folge  $\{s_n\}$  der Abschätzung

$$(4) \quad \sigma_n^{(m)} = S + O \left\{ \left( \log \frac{1}{R(1/n)} \right)^{-m} \right\}.$$

Genauer, wurde folgender, auf die Laplace—Stieltjes-Transformation verallgemeinerte Satz bewiesen.

<sup>1</sup> Infolge (3) besteht die Ungleichung

$$R(2^n s) < e^{2^{c_1 n}} R(s).$$

Es sei nun  $\frac{1}{2^{n+1}} \leq s < \frac{1}{2^n}$ ,  $n > 1$ ; dann erhält man

$$R(s) \geq R\left(\frac{1}{2^{n+1}}\right) > R(1)e^{-2(n+1)c_1} > R(1)e^{-4nc_1} \geq R(1) \exp\left(-\frac{4c_1}{\log 2} \log \frac{1}{s}\right) = R(1) s^{\frac{4c_1}{\log 2}}$$

Diese Ungleichung wird in unseren Überlegungen eine gewisse Bedeutung haben.



SATZ I. Es sei  $\tau(t)$  eine monoton nicht abnehmende Funktion, deren Laplace—Stieltjes—Transformierte

$$(5) \quad F(s) = \int_0^{\infty} e^{-st} d\tau(t)$$

für  $s > 0$  existiert.<sup>2</sup>

Es sei ferner für  $s \rightarrow +0$

$$(6) \quad F(s) = S \frac{\Gamma(\alpha+1)}{s^\alpha} [1 + O\{R(s)\}],$$

dann ist

$$(7) \quad \int_0^x (x-t)^m d\tau(t) = \frac{m! S x^{\alpha+m}}{(\alpha+1)(\alpha+2)\cdots(\alpha+m)} \left[ 1 + O\left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$

(In meiner Arbeit [2] wird etwas allgemeiner der Ausdruck

$$\int_0^x t^\beta (x-t)^m d\tau(t)$$

abgeschätzt.)

Als Ergänzung dieses Satzes sei bemerkt, daß (7) gültig bleibt, falls statt  $\tau(t)$  eine Funktion  $\tau^*(t)$  gewählt wird, die in jedem endlichen Teilintervall von  $(0, \infty)$  eine beschränkte Variation besitzt, und folgende Bedingung befriedigt:<sup>3</sup>

Es gibt eine für  $t \geq 0$  definierte monoton nicht abnehmende Funktion  $\beta_1(t)$  und eine positive Konstante  $L$ , so daß, für  $s \rightarrow +0$

$$\int_0^{\infty} e^{-st} d\beta_1(t) = \frac{\Gamma(\alpha+1)}{s^\alpha} [1 + O\{R(s)\}],$$

wobei das links stehende Integral für  $s > 0$  konvergiert, und

$$\gamma_1(t) = L\beta_1(t) + \tau^*(t)$$

eine monoton nicht abnehmende Funktion ist. Diese Verallgemeinerung ergibt sich einfach, indem man Satz I erst auf  $\beta_1(t)$ , dann auf  $\gamma_1(t)$  anwendet.

Falls man statt (2) voraussetzt, daß

$$(8) \quad a_n = s_n - s_{n-1} > -\frac{K}{n},$$

dann folgt aus dem bekannten Satze von HARDY und LITTLEWOOD sogar die Konvergenz von  $\{s_n\}$ . In dieser Mitteilung beschäftigen wir uns mit der Verschärfung dieses Hardy—Littlewoodschen Satzes, u. zw. in zwei verschiedenen Richtungen.

<sup>2</sup> Im folgenden sei unter  $\tau(t)$  immer eine Funktion mit diesen Eigenschaften verstanden. Ohne es in jedem einzelnen Falle hervorzuheben, sei jede auftretende Laplace—Stieltjes—Transformierte, bzw. Dirichletsche oder Potenzreihe für  $s > 0$  konvergent.

<sup>3</sup> Mit  $\tau^*(t)$  wird in weiteren immer eine Funktion mit dieser Eigenschaft bezeichnet.



Die erste Frage, die wir behandeln, ist folgende:

Wie kann (4) verschärft werden, falls statt (2) die Voraussetzung (8) besteht? Es wird gezeigt, daß das Restglied in (4) zu

$$O \left\{ \left( \log \frac{1}{R(1/n)} \right)^{-m-1} \right\}$$

verschärft werden kann. Es ist dann sogar für  $m=0$  gültig, und in diesem Falle liefert es eine Abschätzung für  $s_n - S$ . Unter der stärkeren Voraussetzung

$$a_n = O \left( \frac{1}{n} \right)$$

wurde dieser Satz auch von J. KOREVAAR [5] bewiesen. In einer unlängs erschienenen Arbeit bewies J. KOREVAAR [6] unter Anwendung meiner schon veröffentlichten Resultate einen Satz, der den Fall  $m=0$  meines Satzes für die Potenzreihe enthält. Im folgenden werden wir den Satz sogleich auf die Laplace—Stieltjes-Transformation verallgemeinerte Form beweisen. Der Satz von HARDY und LITTLEWOOD lautet dann<sup>4</sup> folgendermaßen: Die Laplace—Stieltjes-Transformierte  $F^*(s)$  von  $\tau^*(t)$  existiere für  $s > 0$  und es sei  $F^*(+0) = A$ . Es gebe eine Funktion  $\beta_1(t) \sim t$  mit  $\beta_1(0) = 0$  und eine positive Konstante  $L$ , so daß

$$\gamma_1(t) = L\beta_1(t) + \int_0^t u d\tau^*(u)$$

eine monoton nicht abnehmende Funktion von  $t$  ist, dann gilt

$$\lim_{t \rightarrow \infty} \tau(t) = A.$$

Wir beweisen, daß dieser Satz folgenderweise mit einem Restglied ergänzt werden kann:

SATZ II.

a) Für  $s > 0$  sei

$$(9) \quad F^*(s) = \int_0^{\infty} e^{-st} d\tau^*(t) = A + O\{R(s)\}.$$

b) Es gebe eine monoton nicht abnehmende Funktion  $\beta_2(t)$  und eine positive Konstante  $L$ , so daß für  $s \rightarrow +0$

$$(10) \quad \int_0^{\infty} e^{-st} d\beta_2(t) = \frac{1}{s} [1 + O\{R(s)\}]$$

und

$$(11) \quad \gamma_2(t) = L\beta_2(t) + \int_0^t u d\tau^*(u)$$

<sup>4</sup> Vgl. G. H. HARDY [3], Theorem 102, S. 160.



eine monoton nicht abnehmende Funktion von  $t$  ist. Unter diesen Bedingungen besteht für jede nichtnegative ganze Zahl  $m$

$$(12) \quad \int_0^{\infty} (x-t)^m d\tau(t) = A \frac{x^m}{m!} \left[ 1 + O \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$

Die zweite Frage, mit der wir uns beschäftigen werden, ist folgende. Wie kann die Taubersche Bedingung b) in Satz II geschwächt werden, falls wir aus (9) nur folgern wollen, daß  $\tau(x) \rightarrow A$  ist, ohne daß sich auch eine Abschätzung von  $|\tau(x) - A|$  ergebe. Der zu beweisende Satz lautet folgendermaßen:

SATZ III.

a) Für  $s \rightarrow +0$  sei für ein positives  $\delta$

$$(13) \quad F^*(s) = \int_0^{\infty} e^{-st} d\tau^*(t) = A + O(s^\delta).$$

b) Es gebe eine monoton nicht abnehmende Funktion  $\beta_3(t)$ , so daß für  $t \rightarrow \infty$

$$(14) \quad \beta_3(t) = t \log t + Bt + O(t^{1-\delta})$$

gilt, und für ein beliebiges positives  $\varepsilon$  gebe es ein  $t_0(\varepsilon)$ , so daß für  $t > t_0(\varepsilon)$

$$(15) \quad \gamma_3(t) = \varepsilon \beta_3(t) + \int_0^t u d\tau^*(u)$$

eine monoton nicht abnehmende Funktion von  $t$  ist;  
dann gilt

$$\lim_{t \rightarrow \infty} \tau^*(t) = A.$$

Auf die Potenzreihe angewendet, erhalten wir hieraus folgenden Satz:

Es sei für ein positives  $\delta$

$$\sum_{n=0}^{\infty} a_n e^{-ns} = A + O(s^\delta)$$

und

$$(15a) \quad \liminf \frac{na_n}{\log n} \geq 0,$$

dann ist die Reihe  $\sum a_n$  konvergent und ihre Summe ist gleich  $A$ .

Aus einem Gegenbeispiel von J. KOREVAAR [4], [5] folgt, daß die Größenordnung des Restgliedes in (12) für kein einziges  $m$  verbessert werden kann. Dasselbe Beispiel zeigt auch, daß die Taubersche Bedingung (15a) nicht durch

$$a_n = O\left(\frac{\log n}{n}\right)$$



ersetzt werden kann. Zwischen den Sätzen II und III kann man eine ganze Reihe von Sätzen einschalten, in denen die Forderung b) des Satzes III verschärft wird, und somit eine Abschätzung von  $|x^*(t) - A|$  erhalten wird. Die schon erwähnte, jüngst erschienene Arbeit von J. KOREVAAR [6] ist hauptsächlich diesem Problem im Falle der Potenzreihe gewidmet.

### 1. Beweis des Satzes II

Wir erinnern an folgenden Satz von J. E. LITTLEWOOD [7], der in Anwendungen auf Taubersche Sätze eine hervorragende Rolle spielt:

Es sei  $f(x)$  eine in  $(0, a)$  definierte zweimal differenzierbare Funktion, es existiere

$$\lim_{x \rightarrow +0} f(x) = A$$

und es sei

$$f''(x) > -Kx^{-2},$$

dann ist

$$f'(x) = o(x^{-1}).$$

Wir benötigen folgende Verallgemeinerung dieses Satzes:

HILFSSATZ I.  $G(s)$  sei eine Funktion, die in einer rechtseitigen Umgebung der Stelle  $s=0$  zweimal differenzierbar ist, und es besteht

$$(16) \quad |G(s) - A| < r(s),$$

wobei  $r(s)$  eine monoton zunehmende Funktion ist, ferner sei  $r(+0) = 0$  und

$$(17) \quad r(2s) < c_3 r(s).$$

Andererseits sei

$$(18) \quad G''(s) > -\varrho(s),$$

wobei  $\varrho(s)$  eine monoton abnehmende positive Funktion ist, für welche  $\varrho(+0) = \infty$  und

$$(19) \quad \varrho\left(\frac{s}{2}\right) < c_4 \varrho(s)$$

besteht. Endlich sei, für genügend kleine  $s$ ,

$$(20) \quad r(s) < \frac{1}{4} s^2 \varrho(s).$$

Unter diesen Bedingungen gilt für  $s \rightarrow +0$

$$(21) \quad G'(s) = O(\sqrt{R(s)\varrho(s)}).$$

BEWEIS. Es gilt

$$G(s \pm h) = G(s) \pm h G'(s) + \frac{h^2}{2} G''(s \pm \vartheta h)$$

mit  $0 \leq \vartheta \leq 1$ .



Es sei

$$h = \sqrt{\frac{r(s)}{\varrho(s)}},$$

dann ist infolge (20)  $h < \frac{1}{2}s$ , also ist

$$\begin{aligned} G'(s) &= \frac{G(s+h) - G(s)}{h} - \frac{h}{2} G''(s + \vartheta h) < \\ &< (c_3 + 1) \frac{r(s)}{h} + \frac{h}{2} \varrho(s) = \left(c_3 + \frac{3}{2}\right) \sqrt{R(s)\varrho(s)} \end{aligned}$$

und ähnlicherweise

$$\begin{aligned} G'(s) &= \frac{G(s) - G(s-h)}{h} + \frac{h}{2} G''(s + \vartheta h) > \\ &> -\frac{2r(s)}{h} - \frac{c_4}{2} h \varrho(s) > -\left(2 + \frac{c_4}{2}\right) \sqrt{r(s)\varrho(s)}, \end{aligned}$$

und damit ist unser Hilfssatz bewiesen.

Wir gehen zum Beweis des Satzes II über. Aus (9) und (11) ergibt sich

$$(22) \quad F^{**}(s) = \int_0^{\infty} e^{-st^2} d\tau^*(t) = \int_0^{\infty} e^{-st} d\gamma_2(t) - L \int_0^{\infty} e^{-st} d\beta_2(t).$$

Infolge (10) ist  $\beta_2(t) \sim t$ , also

$$(23) \quad \int_0^{\infty} e^{-st} d\beta_2(t) \sim \frac{1}{s^2}.$$

Da  $\gamma_2(t)$  monoton nicht abnehmend ist, folgt aus (22) und (23)

$$(24) \quad F^{**}(s) > -\frac{c_5}{s^2}.$$

Aus dem angeführten Hilfssatze folgt also, indem man (9) und (24) beachtet

$$(25) \quad F^{**}(s) = \int_0^{\infty} e^{-st} t d\tau(t) = O\left(\frac{\sqrt{R(s)}}{s}\right).$$

Aus (11), (10) und (25) ergibt sich

$$(26) \quad \int_0^{\infty} e^{-st} d\gamma_2(t) = \frac{L}{s} [1 + O(R(s)) + O(\sqrt{sR(s)})].$$

Nach Fußnote <sup>3</sup> besteht

$$\log \frac{1}{\sqrt{sR(s)}} = O\left\{\log \frac{1}{R(s)}\right\}$$



und somit folgt aus Satz I wegen (26)

$$(27) \quad \int_0^x (x-t)^m d\gamma_2(t) = L \frac{x^{m+1}}{(m+1)!} \left[ 1 + O \left\{ \log \left( \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$

Aus (10) folgt wegen Satz I

$$(28) \quad \int_0^x (x-t)^m d\beta_2(t) = \frac{x^{m+1}}{(m+1)!} \left[ 1 + O \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$

Aus (11), (27) und (28) erhält man

$$(29) \quad \int_0^x (x-t)^m t d\tau^*(t) = x^{m+1} O \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\}.$$

Für  $m=0$  folgt hieraus

$$\lim_{x=\infty} \frac{1}{x} \int_0^x t d\tau^*(t) = 0.$$

Aus der wohlbekanntem Verallgemeinerung des Satzes von TAUBER folgert man hieraus

$$\lim_{x=\infty} \tau^*(x) = A$$

und somit ist  $\tau^*(t)$  bestimmt beschränkt:  $|\tau^*(t)| < M$ . Mit einer partiellen Integration ergibt sich aus (9)

$$\int_0^\infty e^{-st} \tau^*(t) dt = \frac{A}{s} [1 + O\{R(s)\}].$$

Hier steht links die Laplace—Stieltjes-Transformierte von  $\int_0^t \tau^*(u) du$ . Infolge der Beschränktheit von  $\tau^*(t)$  ist

$$Mt + \int_0^t \tau^*(u) du$$

eine monoton wachsende Funktion, infolgedessen gilt wegen der Verallgemeinerung des Satzes I

$$(30) \quad \int_0^x (x-t)^m \tau^*(t) dt = A \frac{x^{m+1}}{(m+1)!} \left[ 1 + O \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$



Beachten wir jetzt die Identität

$$\begin{aligned} 0 &= [(x-t)^{m+1} \tau^*(t)]_0^x = -(m+1) \int_0^x (x-t)^m \tau^*(t) dt + \int_0^x (x-t)^{m+1} d\tau^*(t) = \\ &= -(m+1) \int_0^x (x-t)^m \tau^*(t) dt + x \int_0^x (x-t)^m d\tau^*(t) - \int_0^x (x-t)^m t d\tau^*(t). \end{aligned}$$

Hieraus erhält man infolge (29) und

$$\begin{aligned} (31) \quad \int_0^x (x-t)^m d\tau^*(t) &= \frac{1}{x} \left[ (m+1) \int_0^x (x-t)^m \tau^*(t) dt - \right. \\ &\left. - \int_0^x (x-t)^m t d\tau^*(t) \right] = A \frac{x^m}{m!} \left[ 1 + \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right]. \end{aligned}$$

## 2. Anwendung auf Dirichletsche Reihen und Potenzreihen

Unser Ausgangspunkt ist folgender Satz von G. H. HARDY und J. E. LITTLEWOOD:

Es sei  $\{\lambda_n\}$  eine monoton gegen  $+\infty$  zunehmende Zahlenfolge, für welche

$$(32) \quad \lim_{n \rightarrow \infty} \frac{\lambda_{n+1}}{\lambda_n} = 1$$

besteht; es sei ferner

$$(33) \quad \lim_{s \rightarrow +0} \sum_{n=1}^{\infty} a_n e^{-\lambda_n s} = A$$

und

$$(34) \quad a_n > -K \frac{\lambda_n - \lambda_{n-1}}{\lambda_n},$$

dann konvergiert die Reihe  $\sum a_n$  und ihre Summe ist gleich A. ANANDA RAU<sup>5</sup> zeigte, daß aus den Bedingungen dieses Satzes (32) nicht gestrichen werden kann. (Fordert man aber statt (34)

$$a_n = O\left(\frac{\lambda_n - \lambda_{n-1}}{\lambda_n}\right),$$

dann ist die Bedingung (32) unnötig.)

Im folgenden wird dieser Satz unter Anwendung des Satzes II mit einer Abschätzung des Restgliedes erweitert.

<sup>5</sup> *Journal of the London Math. Soc.*, 3 (1928), und *Proc. of the London Math. Soc.* (2), 30 (1930).



Es sei aber ausdrücklich hervorgehoben, daß in dieser Weise die Möglichkeiten der Anwendung des Satzes II auf Dirichletsche Reihen nicht im mindesten ausgeschöpft werden.

SATZ IV. Für die monoton gegen  $+\infty$  strebende Zahlenfolge  $\{\lambda_n\}$  gelte

$$(35) \quad \lambda_{n+1} - \lambda_n = O(\lambda_n^{1-\delta}) \quad \delta > 0,$$

und es sei für  $s \rightarrow +0$

$$(36) \quad \sum_{n=1}^{\infty} a_n e^{-\lambda_n s} = A + O\{R(s)\},$$

wobei

$$(37) \quad a_n > -K \frac{\lambda_n - \lambda_{n-1}}{\lambda_n}$$

besteht; hieraus folgt mit jeder festen ganzen nichtnegativen Zahl  $m$

$$\sum_{\lambda_n \leq x} \left(1 - \frac{\lambda_n}{x}\right)^m a_n = A + O\left\{\left[\log \frac{1}{R(1/x)}\right]^{-m-1}\right\}.$$

BEWEIS. Es sei

$$\beta(t) = \lambda_{n(t)},$$

wobei

$$\lambda_{n(t)} \leq t < \lambda_{n(t)+1}.$$

Dann ist infolge (35)

$$\beta(t) = t + O(t^{1-\delta}),$$

also gilt

$$\int_0^{\infty} e^{-st} d\beta(t) = \frac{1}{s} [1 + O(s^\delta)]$$

und wegen (35) ist

$$K\beta(t) + \sum_{\lambda_n \leq t} \lambda_n a_n$$

eine monoton wachsende Funktion von  $t$ . Somit sind alle Bedingungen des Satzes II erfüllt, falls nur

$$s^\delta = O\{R(s)\}$$

besteht. Infolge Fußnote<sup>1</sup> bedeutet das aber keine Einschränkung.

Für den Fall  $\lambda_n = n$  (Potenzreihe) ist (35) jedenfalls erfüllt. Also folgt aus dem angeführten Satze, daß falls für  $s \rightarrow +0$

$$(38) \quad \sum_{n=1}^{\infty} a_n e^{-ns} = A + O\{R(s)\}$$

gilt und

$$(39) \quad a_n > -\frac{K}{n}$$



besteht, dann gilt

$$\sum_{n \equiv x} \left(1 - \frac{n}{x}\right)^m \alpha_n = A + O \left\{ \left[ \log \frac{1}{R(1/x)} \right]^{-m-1} \right\}.$$

Hieraus ergibt sich mit derselben elementaren Überlegung, wie im Teile V der zweiten Mitteilung, folgender Satz:

SATZ V. Ist (38) und (39) befriedigt, und bezeichnet  $\sigma_n^{(m)}$  das  $n$ -te Cesàrosche Mittel  $m$ -ter Ordnung, dann besteht

$$(40) \quad \sigma_n^{(m)} = A + O \left\{ \left[ \log \frac{1}{R(1/x)} \right]^{-m-1} \right\}.$$

### 3. Beweis des Satzes III

HILFSSATZ II. Es sei  $\tau(t)$  eine monoton nicht abnehmende Funktion,  $\tau(0) = 0$ , und es sei für ein festes positives  $\delta$

$$(41) \quad \left| \int_0^\infty e^{-st} d\tau(t) - A\Gamma(\alpha+1) \frac{\log \frac{1}{s}}{s^\alpha} - A' \frac{1}{s^\alpha} \right| < \frac{C}{s^{\alpha-\delta}},$$

dann gilt für  $x > 2$

$$(42) \quad |\tau(x) - Ax^\alpha \log x| < c_6 \left[ (A+A')x^\alpha + Cx^{\alpha-\frac{\delta}{2}} \right],$$

wobei  $c_6$  eine nur von  $\alpha$  und  $\delta$  abhängige positive Konstante bedeutet.

Der Beweis dieses Hilfssatzes gestaltet sich ebenso, wie der Beweis des Hauptsatzes der ersten Mitteilung. Darum soll es nur in großen Zügen angedeutet werden.

In (41) ersetzen wir  $s$  durch  $(k+1)s$ :

$$(43) \quad \left| \int_0^\infty e^{-st} (e^{-st})^k d\tau(t) - A\Gamma(\alpha+1) \left( \frac{\log \frac{1}{s}}{s^\alpha} \cdot \frac{1}{(k+1)^\alpha} - \frac{1}{s^\alpha} \frac{\log(k+1)}{(k+1)^\alpha} \right) - A' \frac{1}{s^\alpha} \frac{1}{(k+1)^\alpha} \right| < \frac{C}{(k+1)^{\alpha-\delta}} \frac{1}{s^{\alpha-\delta}}.$$

Die von  $k$  abhängigen Faktoren an der rechten Seite wollen wir als Momente  $k$ -ter Ordnung darstellen:

$$\frac{1}{(k+1)^\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^1 \left( \log \frac{1}{x} \right)^{\alpha-1} x^k dx$$



und

$$\frac{\log(k+1)}{(k+1)^\alpha} = \frac{\Gamma'(\alpha)}{[\Gamma(\alpha)]^2} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} x^k dx - \\ - \frac{1}{\Gamma(\alpha)} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} \log \log \frac{1}{x} x^k dx.$$

Ist also

$$H_N(x) = \sum_{k=0}^N \beta_k x^k$$

ein beliebiges Polynom höchstens  $N$ -ten Grades, dann erhalten wir aus (43):

$$(44) \quad \left| \int_0^\infty e^{-st} H_N(e^{-st}) d\tau(t) - A \frac{\log \frac{1}{s}}{s^\alpha} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} H_N(x) dx - \right. \\ \left. - \frac{A\alpha\Gamma'(\alpha)}{\Gamma(\alpha)s^\alpha} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} H_N(x) dx + \frac{A\alpha}{s^\alpha} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} \log \log \frac{1}{x} H_N(x) dx - \right. \\ \left. - \frac{A'}{\Gamma(\alpha)s^\alpha} \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} H_N(x) dx \right| < \frac{C}{s^{\alpha-\delta}} \sum_{k=0}^N \frac{|\beta_k|}{(k+1)^{\alpha-\delta}}.$$

Es sei

$$(45) \quad g(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq e^{-1}, \\ \frac{1}{x} & \text{für } e^{-1} < x \leq 1, \end{cases}$$

und wir konstruieren Polynome  $\varphi_N(x)$  und  $\Phi_N(x)$  vom Grade  $N$ , für welche

$$(46) \quad \varphi_N(x) \leq g(x) \leq \Phi_N(x)$$

und

$$(47) \quad \int_0^1 \left(\log \frac{1}{x}\right)^{\alpha-1} \left[1 + \left|\log \log \frac{1}{x}\right|\right] \left[\Phi_N(x) - \varphi_N(x)\right] dx < \frac{c_7}{N}.$$

Die Konstruktion dieser Polynome erfolgt in derselben Weise, wie die Konstruktion der entsprechenden Polynome im Teil II der ersten Mitteilung. Es wurde ebenfalls in der ersten Mitteilung gezeigt, daß die Koeffizienten  $b_k$  bzw.  $B_k$  der Polynome  $\varphi_N(x)$  bzw.  $\Phi_N(x)$  die Ungleichungen

$$(48) \quad \sum_{k=0}^N |b_k| < c_8 e^{c_9 N} \quad \text{bzw.} \quad \sum_{k=0}^N |B_k| < c_8 e^{c_9 N}$$

befriedigen. (Die Durchführung einiger Modifizierungen im Beweise sei dem



Leser überlassen.) Infolge (45) und (47) ist

$$(49) \quad \left| \int_0^1 \left( \log \frac{1}{x} \right)^{\alpha-1} \Phi_N(x) dx - \frac{1}{\alpha} \right| < \frac{c_7}{N},$$

ferner

$$\left| \int_0^1 \left( \log \frac{1}{x} \right)^{\alpha-1} \log \log \frac{1}{x} \Phi_N(x) dx \right| < c_{10}.$$

Somit folgt aus (44)

$$(50) \quad \left| \int_0^\infty e^{-st} \Phi_N(e^{-st}) d\tau(t) - A \frac{\log \frac{1}{s}}{s} \right| < \\ < \frac{Ac_7}{N} \frac{\log \frac{1}{s}}{s} + (A + A')c_{11} \frac{1}{s^\alpha} + \frac{1}{s^{\alpha-\delta}} \sum_{k=0}^N \frac{|B_k|}{(k+1)^{\alpha-\delta}}.$$

Wegen (48) besteht

$$(51) \quad \sum_{k=0}^N \frac{|B_k|}{(k+1)^{\alpha-\delta}} < c_{12} e^{c_{13}N},$$

also folgt aus (50) und (51), falls man

$$x = \frac{1}{s} \quad N = \left[ \frac{\delta}{2c_{13}} \log x \right] + 1$$

setzt,

$$(52) \quad \left| \int_0^\infty e^{-\frac{t}{x}} \Phi_N(e^{-\frac{t}{x}}) d\tau(t) - Ax^\alpha \log x \right| < (A + A')c_{14}x^\alpha + c_{15}Bx^{\alpha-\frac{\delta}{2}}.$$

Ähnlicherweise erhält man

$$(53) \quad \left| \int_0^\infty e^{-\frac{t}{x}} \varphi_N(e^{-\frac{t}{x}}) d\tau(t) - Ax^\alpha \log x \right| < c_{16}(A + A')x^\alpha + c_{17}Bx^{\alpha-\frac{\delta}{2}},$$

aus (45) folgt

$$(54) \quad \int_0^\infty e^{-\frac{t}{x}} \varphi_N(e^{-\frac{t}{x}}) d\tau(t) \leq \tau(x) \leq \int_0^\infty e^{-\frac{t}{x}} \Phi_N(e^{-\frac{t}{x}}) d\tau(t),$$

also ist wegen (52), (53) und (54) die Abschätzung (42) befriedigt. Q. e. d.

Wir kommen jetzt zum Beweise des Satzes III. Aus

$$\beta_3(t) = t \log t + Bt + O(t^{1-\delta})$$

folgt

$$(55) \quad \int_0^\infty e^{-st} d\beta_3(t) = \frac{\log \frac{1}{s}}{s} + \frac{B'}{s} + O(s^{-1+\delta})$$



und

$$(56) \quad \int_0^{\infty} e^{-st} t d\beta_3(t) = O\left(\frac{\log \frac{1}{s}}{s^2}\right).$$

Infolge der Voraussetzung b) ist, mit einer passend gewählten positiven Zahl  $t_1$ ,

$$\gamma(t) = \beta_3(t) + \int_0^t u d\tau^*(u)$$

eine für  $t > t_1$  monoton nicht abnehmende Funktion von  $t$ . Somit erhalten wir aus (56)

$$(57) \quad \begin{aligned} F^{**'}(s) &= \int_0^{\infty} e^{-st} t^2 d\tau^*(t) = \int_0^{\infty} e^{-st} t d\gamma(t) - \int_0^{\infty} e^{-st} t d\beta_3(t) > \\ &> \int_0^{t_1} e^{-st} t d\gamma(t) - c_{14} \frac{\log \frac{1}{s}}{s^2} > -c_{15} \frac{\log \frac{1}{s}}{s^2}. \end{aligned}$$

Hieraus folgt wegen Hilfssatz II

$$(58) \quad F^{**'}(s) = \int_0^{\infty} e^{-st} t d\tau^*(t) = O(s^{-1+\frac{\delta}{3}}).$$

Es sei nun  $\varepsilon$  eine beliebig kleine feste positive Zahl, und  $t_0 = t_0(\varepsilon)$  die in Voraussetzung b) definierte Zahl. Infolge (15), (55) und (58) gilt

$$\begin{aligned} \int_0^{\infty} e^{-st} d\gamma_3(t) &= \varepsilon \int_0^{\infty} e^{-st} d\beta_3(t) + \int_0^{\infty} e^{-st} t d\tau^*(t) = \\ &= \varepsilon \left( \frac{\log \frac{1}{s}}{s} + \frac{B'}{s} \right) + O(s^{-1+\frac{\delta}{3}}), \end{aligned}$$

also besteht

$$(59) \quad \int_{t_0}^{\infty} e^{-st} d\gamma_3(t) = \varepsilon \left( \frac{\log \frac{1}{s}}{s} + \frac{B'}{s} \right) + O(s^{-1+\frac{\delta}{3}}).$$

Allerdings ist hier die  $O$ -Abschätzung an der rechten Seite nicht mehr in  $\varepsilon$  gleichmäßig. Aus (59) erhält man infolge des Hilfssatzes II

$$(60) \quad |\gamma_3(x) - \gamma_3(t_0) - Ax \log x| \leq c_6(1+B')\varepsilon x + K_1(\varepsilon)x^{1-\frac{\delta}{6}}.$$

Wegen (14) und (15) folgt hieraus, daß für genügend großes  $x$

$$\left| \int_0^x u d\tau^*(u) \right| \leq c_{15}\varepsilon x$$



besteht, wobei  $c_{15} = 2c_6(1 + B')$  sowohl von  $\varepsilon$ , als von  $x$  unabhängig ist. Das bedeutet aber

$$(61) \quad \lim_{x \rightarrow \infty} \frac{1}{x} \int_0^x u d\tau^*(u) = 0.$$

Aus (61) folgt nach einem bekannten Satze

$$\lim_{x \rightarrow \infty} \tau^*(x) = A,$$

was zu beweisen war.

(Eingegangen am 3. August 1954.)

### Literaturverzeichnis

- [1] G. FREUD, Restglied eines Tauberschen Satzes. I, *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 299—308.
- [2] G. FREUD, Restglied eines Tauberschen Satzes. II, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 299—307.
- [3] G. H. HARDY, *Divergent series* (Oxford, 1949).
- [4] J. KOREVAAR, An estimate of the error in tauberian theorems for power series, *Duke Math. Journal*, **18** (1951), S. 723—734.
- [5] J. KOREVAAR, Best  $L_1$ -approximation and the remainder in Littlewood's theorem, *Indagationes Math.*, **15** (1953), S. 281—293.
- [6] J. KOREVAAR, A very general form of Littlewood's theorem, *Indagationes Math.*, **16** (1954), S. 36—45.
- [7] J. E. LITTLEWOOD, The converse of Abel's theorem on power series, *Proc. of the London Math. Soc.* (2), **9** (1911), 434—448.



## ОБ ОСТАТОЧНОМ ЧЛЕНЕ НЕКОТОРОЙ ТЕОРЕМЫ ТИПА ТАУБЕРА. III

Г. ФРАЙД (Будапешт)

(Резюме)

Теорема 1. Пусть  $z^*(t)$  есть функция, определенная на  $(0, \infty)$ , ограниченной вариации на всяком отрезке конечной длины и удовлетворяющая следующим условиям:

$$F^*(s) = \int_0^{\infty} e^{-st} dz^*(t)$$

(преобразованная Лапласа—Стилтьеса) существует для  $s > 0$ . При  $s \rightarrow +0$

$$F^*(s) = A + O\{R(s)\},$$

где  $R(s)$  монотонная неубывающая функция, определенная для  $s > 0$ ,  $R(+0) = 0$  и

$$R(kx) < e^{c_k} R(x) \quad (k = 2, 3, \dots).$$

Существуют положительная постоянная  $L$  и монотонная неубывающая функция  $\beta_2(t)$ , для которых при  $s \rightarrow +0$

$$\int_0^{\infty} e^{-st} d\beta_2(t) = \frac{1}{s} [1 + O\{R(s)\}],$$

и

$$\gamma_2(t) = L\beta_2(t) + \int_0^t u dz^*(u)$$

есть монотонная неубывающая функция  $t$ .

При таких условиях имеет место соотношение

$$\int_0^x (x-t)^m dz^*(t) = A \frac{x^m}{m!} \left[ 1 + O \left\{ \left( \log \frac{1}{R(1/x)} \right)^{-m-1} \right\} \right].$$

Теорема 2. Пусть выполняются следующие условия:

$$F^*(s) = A + O(s^\delta) \quad (\delta > 0).$$

Для монотонной неубывающей функции  $\beta_3(t)$  имеет место

$$\beta_3(t) = t \log t + Bt + O(t^{1-\delta}).$$

Для любого  $\varepsilon > 0$  можно найти такое число  $t_0(\varepsilon)$ , что при  $t > t_0(\varepsilon)$

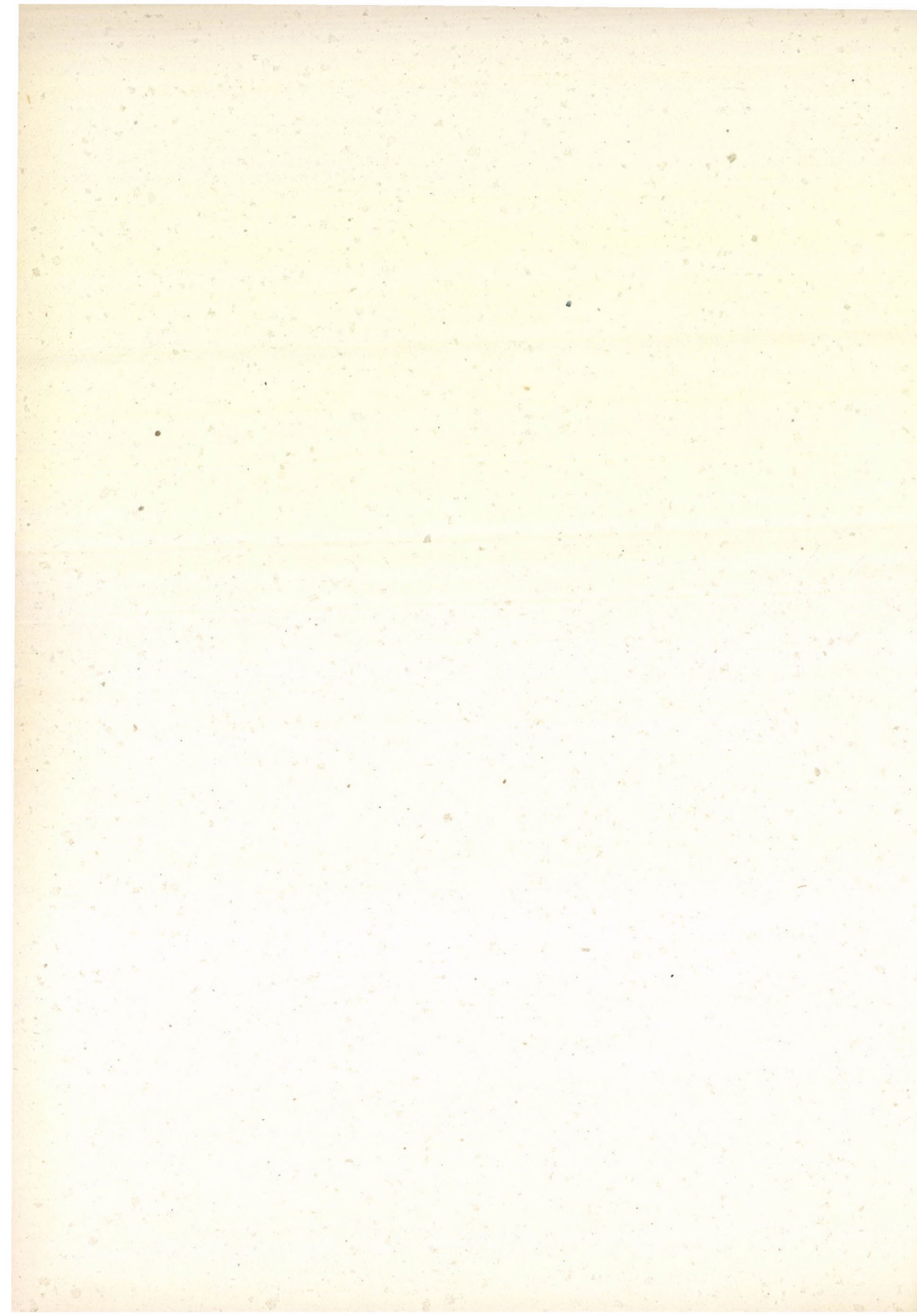
$$\gamma_3(t) = \varepsilon \beta_3(t) + \int_0^t u dz^*(u)$$

является монотонной неубывающей функцией  $t$ .

При таких условиях

$$\lim_{t \rightarrow \infty} z^*(t) = A.$$







# ÜBER ORTHOGONALE POLYNOME

Von

G. FREUD (Budapest)

(Vorgelegt von P. TURÁN)

Es sei  $w(x)$  eine in  $[-1, +1]$  definierte, nichtnegative  $L$ -integrierbare Gewichtsfunktion,

$$\int_{-1}^{+1} w(x) dx > 0.$$

Wir bezeichnen die Folge der normierten Orthogonalpolynome, die zu dieser Gewichtsfunktion im Orthogonalitätsintervall  $[-1, +1]$  gehören, mit  $\{p_n(x)\}$ . Verfasser untersuchte die Entwicklungen, die nach Orthogonalpolynomen fortschreiten, falls diese der Bedingung

$$(1) \quad \sum_{\nu=0}^{n-1} [p_\nu(x)]^2 = O(n)$$

unterworfen sind ([2], [3]). In [2] wurde gezeigt, daß jede mit der Gewichtsfunktion  $w(x)$  quadratisch integrierbare Funktion in fast allen Punkten des Orthogonalitätsintervalles, wo (1) erfüllt ist, stark  $(C, 1)$ -summierbar ist. In [3] wurden die Sätze von S. N. BERNSTEIN über die absolute Konvergenz von Fourierreihen auf Entwicklungen nach Orthogonalpolynomen verallgemeinert. Verfasser zeigte [2], daß falls in einem Teilintervall  $(a, b)$  von  $[-1, +1]$   $w(x) \geq m > 0$  besteht, (1) in jedem inneren Teilintervall von  $(a, b)$  gleichmäßig erfüllt ist. Die Bedeutung dieser Sätze liegt darin, daß man auf die absolute Konvergenz bzw.  $(C, 1)$ -Summierbarkeit der Orthogonalpolynomreihe aus dem Verlauf der Gewichtsfunktion schließen kann, ohne etwas über die Beschränktheit der Orthogonalpolynome voraussetzen zu müssen.

Das Problem, ob (1) nicht auch an fast allen Stellen besteht, wo  $w(x) > 0$  ist, scheint sehr tief zu sein. Der Verfasser ist nicht in der Lage es zu beantworten. Doch wird es gezeigt, daß dies der Fall ist, wenn die Gewichtsfunktion einer allgemeinen strukturellen Bedingung genügt.

Es sei

$$(2) \quad W(\theta) = \begin{cases} w(\cos \theta) \sin \theta & 0 \leq \theta \leq \pi, \\ 0 & \theta \in [0, \pi]. \end{cases}$$

SATZ I. Für genügend kleine Werte von  $h$  sei

$$\frac{W(\theta+h) - W(\theta)}{W(\theta)}$$



$L$ -integrierbar, ferner sei

$$(3) \quad \int_0^\pi \frac{|W(\theta+h) - W(\theta)|}{W(\theta)} d\theta = O\left(\log^{-\alpha} \frac{1}{|h|}\right)$$

mit  $\alpha > 1$ ; dann ist (1) an fast jeder Stelle  $x$  erfüllt.

Andererseits wird es gezeigt, daß das Verhalten des Ausdruckes (1) auf der Punktmenge  $N$ , die aus den Nullstellen von  $w(x)$  besteht, ohne jede Nebenbedingung beherrscht werden kann:

SATZ II. Es gilt fast überall auf  $N$

$$(4) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(x) = \infty.$$

Es sei auch die bemerkenswerteste Folgerung des Satzes I hervorgehoben:

SATZ III. Besteht die Voraussetzung (3) des Satzes I, dann ist die Entwicklung nach Orthogonalpolynomen jeder Funktion  $f \in L^2(w)$  fast überall stark  $(C, 1)$  summierbar.

Die Untersuchungen des Verfassers folgten eine Arbeit von K. TANDORI [5]. Er bewies die starke  $(C, 1)$ -Summierbarkeit der Orthogonalpolynomreihe einer Funktion  $f \in L^2(w)$  unter der stärkeren Bedingung, daß die normierten Orthogonalpolynome in dem betrachteten Teilintervall gleichmäßig beschränkt sind. Weiter bemerkte er, daß aus seinem Satze folgende Behauptung folgt:

Verschwindet  $w(x)$  auf der Punktmenge  $N$  vom positiven Maße, dann können die normierten Orthogonalpolynome auf dieser Punktmenge nicht beschränkt sein. Satz II kann als eine Präzisierung dieser Behauptung betrachtet werden.

### 1. Approximation von $\frac{1}{n} \sum_{k=0}^{n-1} p_k^2(x)$ im $L$ -Raume

In den folgenden Betrachtungen, wie auch in einigen früheren Arbeiten des Verfassers, wird folgender Hilfssatz eine hervorragende Rolle spielen:

HILFSSATZ I. Es seien alle Polynome  $\Phi_{n-1}(t)$  höchstens  $n-1$ -ten Grades betrachtet, welche die Ungleichung

$$(5) \quad \int_{-1}^{+1} \Phi_{n-1}^2(t) w(t) dt \leq 1$$

befriedigen. Dann besteht

$$(6) \quad \Phi_{n-1}^2(x) \leq \sum_{k=0}^{n-1} p_k^2(x),$$



und der Höchstwert an der rechten Seite wird für

$$\Phi_{n-1}^*(t) = \frac{\sum_{k=0}^{n-1} p_k(x) p_k(t)}{\left\{ \sum_{k=0}^{n-1} p_k^2(x) \right\}^{1/2}}$$

erreicht. (Vgl. G. SZEGŐ [4], Theorem 3.1.3, S. 38.)

Nach einem Gedanken von P. ERDŐS und P. TURÁN ([1], S. 539—540) setzen wir als Konkurrenzpolynom

$$(7) \quad \Phi_{n-1}(t) = \frac{\sum_{k=0}^{n-1} T_k(x) T_k(t)}{\left\{ \int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right] w(u) du \right\}^{1/2}},$$

wobei  $T_0(t) \equiv \frac{1}{\sqrt{2}}$  ist und  $T_n(x)$  das Tschebyscheffsche Polynom  $n$ -ten Grades bedeutet. Infolge unseres Hilfssatzes folgt also mit der Bezeichnung  $\theta = \arccos x$

$$(8) \quad \sum_{k=0}^{n-1} p_k^2(x) \geq \frac{\left[ \sum_{k=0}^{n-1} T_k^2(x) \right]^2}{\int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right]^2 w(u) du} = \left( \frac{n - \frac{1}{2} + \frac{1}{2} \frac{\sin(2n-1)\theta}{\sin \theta}}{2} \right)^2 \frac{1}{\int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right]^2 w(u) du}.$$

Das Integral im Nenner wollen wir mittels (2) umformen:

$$\begin{aligned} & \int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right]^2 w(u) du = \\ & = \int_0^\pi \left[ \frac{1}{2} + \sum_{k=1}^{n-1} \cos k\theta \cos k\varphi \right]^2 W(\varphi) d\varphi = \\ & = \frac{\pi}{4} \left( n - \frac{1}{2} + \frac{1}{2} \frac{\sin(2n-1)\theta}{\sin \theta} \right) W(\theta) + \\ & + \int_0^\pi \left( \frac{1}{2} + \sum_{k=1}^{n-1} \cos k\theta \cos k\varphi \right)^2 [W(\varphi) - W(\theta)] d\varphi. \end{aligned}$$



Aus einer elementaren, aus der Theorie der Fourierreihen bekannten Abschätzung erhält man

$$\left(\frac{1}{2} + \sum_{k=1}^{n-1} \cos k\theta \cos k\varphi\right)^2 \leq C_1 \operatorname{Min}\left(n^2, \frac{1}{|\theta - \varphi|^2}\right)$$

für  $0 \leq \theta, \varphi \leq \pi$ . Somit besteht

$$(9) \quad \int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right]^2 w(u) du \leq \\ \leq \frac{\pi}{4} \left( n - \frac{1}{2} + \frac{1}{2} \frac{\sin(2n-1)\theta}{\sin \theta} \right) W(\theta) + \\ + C_2 \int_0^\pi \operatorname{Min}\left(n, \frac{1}{n|\theta - \varphi|^2}\right) |W(\theta) - W(\varphi)| d\varphi \left\}.$$

Aus (8) und (9) erhalten wir mit Hilfe der elementaren Ungleichung

$$\frac{1}{a+b} \geq \frac{a-b}{a^2} \\ (10) \quad \sum_{k=0}^{n-1} p_k^2(\cos \theta) \geq \left( n - \frac{1}{2} + \frac{1}{2} \frac{\sin(2n-1)\theta}{\sin \theta} \right) \cdot \\ \frac{W(\theta) - C_2 \int_0^\pi \operatorname{Min}\left(n, \frac{1}{n|\theta - \varphi|^2}\right) |W(\theta) - W(\varphi)| d\varphi}{\pi W^2(\theta)}.$$

Mittels einer einfachen Umformung folgt hieraus

$$\frac{1}{\pi} - \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(\cos \theta) W(\theta) \leq C_3 \int_0^\pi \operatorname{Min}\left(n, \frac{1}{n|\theta - \varphi|^2}\right) \frac{|W(\theta) - W(\varphi)|}{W(\theta)} d\varphi + \\ (11) \quad + \frac{1}{2n} \left( 1 + \left| \frac{\sin(2n-1)\theta}{\sin \theta} \right| \right).$$

Es sei für den „positiven Teil von  $a$ “ die Bezeichnung

$$\{a\}^+ = \begin{cases} a & \text{für } a > 0, \\ 0 & \text{für } a \leq 0 \end{cases}$$

eingeführt. Da die rechte Seite von (11) immer positiv ist, bleibt diese Ungleichung richtig, wenn man die linke Seite durch ihren positiven Teil ersetzt. Nach  $\theta$  integriert, erhält man somit



$$\begin{aligned}
 & \int_0^\pi \left\{ \frac{1}{\pi} - \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(\cos \theta) W(\theta) \right\}^+ d\theta \leq \\
 (12) \quad & \leq C_3 \int_0^\pi \int_0^\pi \text{Min} \left( n, \frac{1}{n|\theta - \varphi|^2} \right) \frac{|W(\theta) - W(\varphi)|}{W(\theta)} d\theta d\varphi + \\
 & + \frac{1}{2n} \int_0^\pi \left( 1 + \left| \frac{\sin(2n-1)\theta}{\sin \theta} \right| \right) d\theta.
 \end{aligned}$$

Hier ist das letzte Glied von der Größenordnung  $O\left(\frac{\log n}{n}\right)$ . Im ersten Gliede auf der rechten Seite soll unter dem Integralzeichen die neue Veränderliche  $h = \theta - \varphi$  eingeführt werden. Indem wir (2) beachten, bekommen wir

$$\begin{aligned}
 & \int_0^\pi \int_0^\pi \text{Min} \left( n, \frac{1}{n|\theta - \varphi|^2} \right) \frac{|W(\theta) - W(\varphi)|}{W(\theta)} d\theta d\varphi = \\
 & = \int_{-\pi}^\pi \text{Min} \left( n, \frac{1}{nh^2} \right) \left\{ \int_0^\pi \frac{|W(\theta + h) - W(\theta)|}{W(\theta)} d\theta \right\} dh = \\
 & = n \int_{-\frac{1}{n}}^{+\frac{1}{n}} \left\{ \int_0^\pi \frac{|W(\theta + h) - W(\theta)|}{W(\theta)} d\theta \right\} dh + \\
 & + \frac{1}{n} \int_{\frac{1}{n}}^\pi \left\{ \int_0^\pi \frac{|W(\theta + h) - W(\theta)|}{W(\theta)} d\theta \right\} dh + \int_0^\pi \frac{|W(\theta - h) - W(\theta)|}{W(\theta)} d\theta \left\{ \frac{dh}{h^2} \right.
 \end{aligned}$$

und somit unter Anwendung der Voraussetzung (3)

$$\int_0^\pi \int_0^\pi \text{Min} \left( n, \frac{1}{n|\theta - \varphi|^2} \right) \frac{|W(\theta) - W(\varphi)|}{W(\theta)} d\theta d\varphi = O(\log^\alpha n),$$

und hieraus ergibt sich wegen (12)

$$(13) \quad \int_0^\pi \left\{ \frac{1}{\pi} - \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(\cos \theta) W(\theta) \right\}^+ d\theta = O(\log^\alpha n).$$

Infolge der Normierung der Polynome  $p_k(x)$  besteht

$$\int_0^\pi \left\{ \frac{1}{\pi} - \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(\cos \theta) W(\theta) \right\} d\theta = 0,$$



und hieraus schließt man wegen (13)

$$(14) \quad I_n = \int_0^\pi \left| \pi - \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(\cos \theta) W(\theta) \right| d\theta = O(\log^{-\alpha} n).$$

Diese Abschätzung bildet den Kern unseres Beweises für Satz I.

## 2. Abschätzung von $\sum_{k=0}^{n-1} p_k^2(x)$

Aus der Ungleichung (14) folgt

$$\sum_{\nu=1}^{\infty} I_{2^\nu} < \infty,$$

also gilt für fast alle  $\theta$

$$(15) \quad \lim_{\nu \rightarrow \infty} \frac{1}{2^\nu} \sum_{k=0}^{2^\nu-1} p_k^2(\cos \theta) = \frac{\pi}{W(\theta)}.$$

Betrachten wir jetzt ein  $\theta$ , für welches (15) erfüllt ist. Es gibt eine von  $\nu$  unabhängige positive Zahl  $K(\theta)$ , so daß

$$\sum_{k=0}^{2^\nu-1} p_k^2(\cos \theta) < K(\theta) 2^\nu \quad (\nu = 0, 1, 2, \dots)$$

besteht. Es sei  $n$  eine beliebige natürliche Zahl, und

$$2^{m-1} \leq n < 2^m,$$

dann ist

$$\sum_{k=0}^{n-1} p_k^2(\cos \theta) \leq \sum_{k=0}^{2^m-1} p_k^2(\cos \theta) \leq K(\theta) 2^m \leq 2K(\theta)n.$$

Da die Ausnahmemenge in  $\theta$  nach der Abbildung  $x = \cos \theta$  wieder in eine Nullmenge übergeht, sind wir mit unserem Beweise fertig.

## 3. Beweis des Satzes II

Fast alle Punkte der Menge der ins Innere von  $(-1, +1)$  fallenden Nullstellen von  $w(x)$  sind Lebesguesche Punkte, d. h.

$$(15) \quad \int_{x-h}^{x+h} w(t) dt = o(h).$$

Wir zeigen, daß in jedem solchen Punkte (4) befriedigt ist. Es wird wieder Hilfssatz I angewandt, und wieder wird als Konkurrenzpolynom der Ausdruck (7) gewählt. Mit diesem Polynom ist (5) befriedigt, sogar mit dem Gleichheitszeichen. Aus (15) erhält man — mit derselben schon klassisch gewordenen



Schlußweise, wie man den Fejér—Lebesgueschen Satz über Fourierreihen beweist — daß

$$(17) \quad \int_{-1}^{+1} \left[ \sum_{k=0}^{n-1} T_k(x) T_k(u) \right]^2 w(u) du = o(n)$$

gültig ist.

Aus (7) erhält man somit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \Phi_{n-1}^2(x) = \infty,$$

und endlich wegen (6)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} p_k^2(x) = \infty,$$

was zu beweisen war.

(Eingegangen am 9. August 1954.)

### Literaturverzeichnis

- [1] P. ERDŐS and P. TURÁN, On interpolation III, *Annals of Math.*, **41** (1940), S. 510—553.
- [2] G. FREUD, Über die starke (C, 1)-Summierbarkeit von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 83—88.
- [3] G. FREUD, Über die absolute Konvergenz von orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **4** (1953), S. 127—136.
- [4] G. SZEGŐ, *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ., Vol. XXIII (New York, 1939).
- [5] K. TANDORI, Über die Cesàrosche Summierbarkeit der orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 73—82.



## ОБ ОРТОГОНАЛЬНЫХ ПОЛИНОМАХ

Г. ФРАЙД (Будапешт)

(Резюме)

Пусть  $p_0(x), p_1(x), \dots, p_n(x)$  полиномы, ортогональные и нормированные с весом  $w(x) \geq 0$  на интервале  $[-1, +1]$ , где

$$W(\theta) = \begin{cases} w(\cos \theta) \sin \theta & 0 \leq \theta \leq \pi, \\ 0 & \theta \in [0, \pi]. \end{cases}$$

Пусть при  $h \rightarrow 0$

$$\int_0^\pi \frac{|W(\theta+h) - W(\theta)|}{W(\theta)} d\theta = O\left(\log^{-\alpha} \frac{1}{|h|}\right), \quad \alpha > 1;$$

тогда для почти всех значений  $x \in [-1, +1]$

$$\sum_{\nu=0}^{n-1} [p_\nu(x)]^2 = O(n).$$

С другой стороны, пусть  $w(x) \in L$  есть некоторая, определенная на  $[-1, +1]$  неотрицательная весовая функция. Обозначим через  $N$  множество нулей функции  $w(x)$ . При таких условиях для почти всех точек  $N$  выполняется

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=0}^{n-1} [p_\nu(x)]^2 = \infty.$$



# A LATTICE-THEORETIC DISCUSSION OF SOME PROBLEMS IN ADDITIVE IDEAL THEORY

By

L. FUCHS (Budapest)  
(Presented by L. RÉDEI)

To Professor L. FEJÉR on his 75th birthday

## § 1. Introduction

The present paper has for its origin the results of a previous paper [6].<sup>1</sup> In that paper we have made use of the known idea of considering, instead of the totality of the ideals in a commutative ring, the elements of a lattice  $L$  which is at the same time a semigroup, having the fundamental properties of the ideal lattice of a commutative ring;<sup>2</sup> our purpose was to obtain a far-reaching generalization of the theory of *primary* and *quasiprimary* ideals.<sup>3</sup> Our main tool in the discussions was the notion of a "closure operator  $\Phi$ " which associated with each element  $x$  of the lattice  $L$  an element  $\Phi(x)$  in  $L$  such that  $x \subseteq \Phi(x)$ ,  $\Phi(\Phi(x)) = \Phi(x)$  and  $\Phi(x \cap y) = \Phi(x) \cap \Phi(y)$  (i. e.  $\Phi(x)$  obeys the same rules as the radical of  $x$ ). It turned out that the theorems got previously for primary and quasiprimary ideals in a commutative ring substantially retain their validity even in this abstract form for the so-called  $\Phi$ -primary and  $\Phi$ -prime elements of the lattice.<sup>4</sup>

In another paper [5] we have introduced another type of ideal, called *primal* ideal, and have shown that, in commutative rings, for primal ideals a theory may be worked out almost entirely similar to those for primary and quasiprimary ideals. These results were generalized to the non-commutative case by CH. W. CURTIS [2]. However, the results on primal ideals can not be generalized immediately to lattices in the same manner as in [6], for the

<sup>1</sup> The numbers in brackets refer to the Bibliography given at the end of the paper.

<sup>2</sup> For such a point of view of discussion see e. g. KRULL [7] in the commutative and [8] in the non-commutative case.

<sup>3</sup> The classical theory of primary ideals in abstract rings is known to be due to NOETHER [11]; see also KRULL [10] and VAN DER WAERDEN [12]. For the theory of quasiprimary ideals we refer to our paper [3].

<sup>4</sup> See [6].



operator which makes the adjoint prime ideal correspond to the primal ideal fails to satisfy the "distributive law"  $\Phi(x \cap y) = \Phi(x) \cap \Phi(y)$  (true in the case of forming the radical) which has played a distinguished role in [6]. Consequently, if we should like to subsume the results on primal ideals under a general theory like that in [6], then we are forced to drop the convenient and useful "distributive law" and consider more general operators  $\Psi$ . In accordance with this, in the present paper we deal with a quite general operator<sup>5</sup>  $\Psi$  which in general does not satisfy any condition, and shall see that — in spite of the great generality of this concept — the essential part of the results on primal ideals may be carried over to lattices  $L$  almost without change. Considering that the commutativity of multiplication has proved not to be an essential hypothesis, we shall discuss with the same trouble the non-commutative case. Moreover, we have succeeded in extending the results of [6] to the non-commutative case and to more general operators than those considered there, namely to operators  $\Phi$  subject to the single condition:  $x \leq \Phi(y)$  implies  $\Phi(x) \leq \Phi(y)$ .

The arrangement of the present paper is as follows. First of all we lay down the definition and main properties of the lattice whose elements will be dealt with. The next section is devoted to the definitions of the notions  $\Psi$ -prime,  $\Psi$ -primary,  $\Psi$ -primal and  $\Psi^*$ -primal elements, as well as to their simplest consequences. Although, strictly speaking, the notion of  $\Psi^*$ -primal element is a direct generalization of our old concept of primal ideal, the  $\Psi$ -primality will also prove an appropriate definition. After a detailed discussion of various examples which themselves will at once convince ourselves of the usefulness of our new notions, we turn our attention to examining the intersection of a finite number of elements possessing one of the properties defined previously. Certain difficulties arise from the fact that we do not know anything of the effect of  $\Psi$  on the meet of elements; we essentially overcome them by introducing the concepts "quasi- $\Psi$ -prime" etc. By making use of these notions we shall be able to prove certain unicity statements concerning representations as the meet of a finite number of  $\Phi$ -prime etc. elements.

## § 2. Preliminaries

We begin with recalling a few known definitions and facts on which the sequel depends.

Let  $L$  be a *lattice-ordered semigroup* which is *complete* as lattice,<sup>6</sup> in other words,  $L$  is a complete lattice in which a binary associative (but not

<sup>5</sup> Otherwise expressed,  $\Psi$  is an arbitrary function on the lattice  $L$  with values in  $L$ .

<sup>6</sup> See BIRKHOFF [1].



necessarily commutative) multiplication is defined satisfying the infinite distributive laws<sup>7</sup>

$$(1) \quad a(\bigcup_{\alpha} b_{\alpha}) = \bigcup_{\alpha} (ab_{\alpha}) \quad \text{and} \quad (\bigcup_{\alpha} b_{\alpha})a = \bigcup_{\alpha} (b_{\alpha}a).$$

Hence it results at once that multiplication is monotone:

$$(2) \quad a \leq b \text{ implies } ax \leq bx \text{ and } xa \leq xb \text{ for all } x \in L.$$

We also assume that  $L$  contains a zero  $0$  and a unity  $e$  such that

$$(3) \quad 0 \leq x \leq e, \quad 0x = x0 = 0, \quad ex = xe = x \text{ for all } x \in L.$$

Let us remark that (2), (3) imply

$$(4) \quad ab \leq a \text{ and } ba \leq a \text{ for all } a, b \in L,$$

or more generally,

$$(5) \quad a_1 \cdots a_n \leq a_1 \cap \cdots \cap a_n.$$

It follows easily the existence of a *right (left) residual*  $a : b$  ( $a :: b$ ) defined as the join of all  $x$  satisfying  $xb \leq a$  ( $bx \leq a$ ). By (1) we have  $(a : b) b \leq a$  [ $b(a :: b) \leq a$ ]. (4) shows that always  $a : b \geq a$  ( $a :: b \geq a$ ); if  $a : b = a$  ( $a :: b = a$ ), we call  $b$  *right (left) prime to  $a$* .

For the residuals one has the following important rules which we shall need throughout:

$$(6) \quad (\bigcap_{\alpha} a_{\alpha}) : b = \bigcap_{\alpha} (a_{\alpha} : b),$$

$$(7) \quad a : (\bigcup_{\beta} b_{\beta}) = \bigcap_{\beta} (a : b_{\beta}),$$

$$(8) \quad a : (bc) = (a : c) : b,$$

and symmetrically for the left residual.

We shall say that  $L$  satisfies the *ascending chain condition* (or *maximal condition*) if no infinite ascending chain  $a_1 < a_2 < \cdots$  with different terms  $a_i$  exists.

In our discussions a fundamental role is played by *operators*  $\Psi$  which associate with each element  $x$  of  $L$  again an element  $\Psi(x)$  of  $L$ .

In certain special cases we have to do with operators possessing some of the following properties:

$$(9) \quad x \leq \Psi(x);$$

$$(10) \quad \Psi(\Psi(x)) = \Psi(x);$$

$$(11) \quad x \leq \Psi(y) \text{ implies } \Psi(x) \leq \Psi(y);$$

$$(12) \quad x \leq y \text{ implies } \Psi(x) \leq \Psi(y);$$

$$(13) \quad \Psi(x \cap y) = \Psi(x) \cap \Psi(y).$$

Obviously, (9) and (11) imply (12), (10) and (12) imply (11), (13) implies (12), while (12) implies only the sign  $\leq$  in (13).

<sup>7</sup>  $\alpha$  and (later)  $\beta$  vary over an arbitrary finite or infinite set of indices.



We emphasize that in general we do not assume anything on the operator  $\Psi$  considered.

$x = a_1 \cap \dots \cap a_n$  is called a *representation* of  $x$  and  $a_i$  ( $i = 1, \dots, n$ ) the *components* of  $x$  in this representation. It is said to be *irredundant* if no component may be omitted, *reduced* if no  $a_i$  may be replaced by an element  $b_i$  with  $b_i > a_i$ . Finally, if the components  $a_i$  have certain property  $P$ , then we shall call the representation *shortest* if it is irredundant and no subset of the  $a_i$  has a meet again with property  $P$ .

### § 3. Definitions

The following concepts are fundamental in what follows.

**A.** An element  $p$  is called  $\Psi$ -prime if<sup>8</sup>  $x_1 \dots x_k \leq p$  ( $k$  arbitrary) implies  $x_i \leq \Psi(p)$  for some index  $i$ .

**B.** An element  $q$  is said to be  $\Psi$ -primary if<sup>8</sup>  $x_1 \dots x_k \leq q$  ( $k$  arbitrary) implies that for each  $i$  one has either  $x_i \leq q$  or  $x_j \leq \Psi(q)$  for some  $j = 1, \dots, i-1, i+1, \dots, k$ .

**C.** An element  $r$  will be called *right  $\Psi$ -primal* if  $\Psi(r)$  is the join of all  $x$  not right prime to  $r$ , i. e.  $r : x > r$ .  $r$  is  $\Psi$ -primal if it is both right and left  $\Psi$ -primal.

**D.** We shall call the element  $s$  *strongly right  $\Psi$ -primal*, or *right  $\Psi^*$ -primal*, if  $x \leq \Psi(s)$  is equivalent to  $s : x > s$ . It is obvious that a right  $\Psi$ -primal element  $r$  is right  $\Psi^*$ -primal if and only if  $r : \Psi(r) > r$  holds.<sup>9</sup>

Now let us consider the simplest consequences of these definitions.

First of all we remark that if  $y (\neq e)$  is a  $\Psi$ -prime,  $\Psi$ -primary, right  $\Psi$ -primal or right  $\Psi^*$ -primal element, then we have necessarily

$$(14) \quad y \leq \Psi(y).$$

In fact, in the first two cases apply the definition to  $ye \leq y$ ,<sup>10</sup> while in the last two cases take into account that  $y : y = e > y$ .

In order to make the connections between the four notions defined above more apparent, let us remark the following. At first, from the definition and (14) it is evident that each  $\Psi$ -primary element is at the same time  $\Psi$ -prime, further a right  $\Psi^*$ -primal element is necessarily right  $\Psi$ -primal. It is not so obvious, but it is not hard to verify that *if  $y$  is a  $\Psi$ -primal element then it is  $\Psi$ -primary too*.<sup>11</sup> For, let  $y$  be a  $\Psi$ -primal element and suppose

<sup>8</sup> It is to be noted that it would not be sufficient to state the definition for  $k = 2$ .

<sup>9</sup> It is uninteresting whether  $e$  has a  $\Psi$ -property or not.

<sup>10</sup> Take into account that  $e \leq \Psi(y)$  implies also  $y < e = \Psi(y)$ !

<sup>11</sup> This generalizes a statement in [5] according to which in a commutative ring an ideal is necessarily primary if it is both primal and quasiprimal such that its adjoint prime ideal and its radical coincide.



$x_1 \dots x_k \leq y$  where  $x_j \not\leq \Psi(y)$  for  $j = 1, \dots, i-1, i+1, \dots, k$ . By definition **C** we must have  $y : x_j = y$  and  $y :: x_j = y$  for these  $x_j$ , so that the hypothesis  $x_1 \dots x_k \leq y$  implies  $x_i \leq y$  establishing the  $\Psi$ -primary character of  $y$ .

What we have just said may be summed up into the following chain of implications:

$$(15) \quad \Psi^*\text{-primal} \Rightarrow \Psi\text{-primal} \Rightarrow \Psi\text{-primary} \Rightarrow \Psi\text{-prime.}$$

It is easy to illustrate by examples that in general none of the converse implications holds, not even in the commutative case and under assumption of the ascending chain condition.<sup>12</sup>

Next we show that if  $s$  is a right  $\Psi^*$ -primal element, then  $\Psi(s)$  is a prime element in the usual sense.<sup>13</sup> Indeed,  $xy \leq \Psi(s)$  implies  $s < s : (xy) = (s : y) : x$ . This shows that if  $s : y = s$ , then  $s : x > s$ , i. e. either  $y \leq \Psi(s)$  or  $x \leq \Psi(s)$ , as stated.

### § 4. Examples

From among the great variety of illustrations the following ones are perhaps the most interesting.

**I.** Define  $\Psi$  as the identity operator in  $L$ , that is to say,  $\Psi(x) = |x| = x$ . In this special case (for elements  $x \neq e$ ) all the four notions coincide: *the l-prime, l-primary, l-primal and l\*-primal elements are identical with those elements which are commonly called primes*. In order to verify this assertion, on account of (15) it suffices to show that l-primeness (=primeness) implies l\*-primality. But this is obvious, considering that for prime elements  $p$  the relation  $x \leq p = |p)$  is equivalent to  $p < p : x = e$  ( $p < p :: x = e$ ).

**II.** The case in which  $P(x)$  is defined as the join of all  $a$  with the property:  $a^k \leq x$  for some natural integer  $k$  seems to have certain special interest. ( $P(x)$  may be called the *radical* of  $x$ .<sup>14</sup>) If  $L$  satisfies the ascending chain condition, then we have also  $P(x)^k \leq x$  for some  $k$ . In fact,  $P(x)$  is, by the ascending chain condition, the join of a finite number of elements,  $P(x) = a_1 \cup \dots \cup a_n$  such that  $a_i^{k_i} \leq x$  ( $i = 1, \dots, n$ ). Owing to the distribution law (1) we have

$$P(x)^k = (a_1 \cup \dots \cup a_n)^k = \cup a_{i_1} \dots a_{i_k}$$

where  $i_1, \dots, i_k$  run over all variations of  $1, 2, \dots, n$  which are of class  $k$ , with repetitions. If we put  $k = k_1 + k_2 + \dots + k_n$  and take (4) into consideration, we obtain that each product  $a_{i_1} \dots a_{i_k}$  is  $\leq a_i^{k_i}$  for some  $i$ , thus also  $\leq x$ . Hence the statement.

<sup>12</sup> For illustrations we refer to [2], [3], [5].

<sup>13</sup>  $p$  is a prime element if  $ab \leq p$  implies either  $a \leq p$  or  $b \leq p$ . We conclude that the same is true for more factors:  $x_1 \dots x_n \leq p$  implies  $x_i \leq p$  for some  $i$ .

<sup>14</sup> KRULL [9], [10].



It is readily seen that  $p$  is  $P$ -prime if and only if  $P(p)$  is a prime element, provided in  $L$  the ascending chain condition holds. To prove this, suppose  $p$  is  $P$ -prime and  $x_1 x_2 \leq P(p)$ . Then, by the preceding observation,  $(x_1 x_2)^k \leq p$  for a certain  $k$ , whence, in view of the  $P$ -primeness of  $p$  we are led to  $x_1 \leq P(p)$  or  $x_2 \leq P(p)$ . Conversely, let  $P(p)$  be prime and  $x_1 \dots x_m \leq p$ . Then  $x_1 \dots x_m \leq P(p)$  whence by hypothesis we obtain  $x_i \leq P(p)$  for an  $i$ , indeed.

It is easy to see that for the operator  $P$  the relations (9), (12), (13) always hold; the validity of (10), (11) becomes obvious as soon as one assumes the ascending chain condition.

Our next aim is to show that the notions  $P$ -primary and  $P$ -primal coincide.<sup>11</sup> By (15) it is enough to verify that the  $P$ -primary character of an element  $q (\neq e)$  implies that it is  $P$ -primal too. Take into account that if  $a^k \leq q (\neq e)$  and  $k$  is the least exponent with this property, then  $a$  is not (right) prime to  $q$ , for  $a^{k-1} \leq q : a$  but  $a^{k-1} \not\leq q$ . Therefore  $P(q) \leq \cup a$  where  $a$  runs over all elements not (right) prime to  $q$ . But if  $q : a > q$  then  $(q : a)a \leq q$  implies, by the hypothesis on  $q$ , that  $a \leq P(q)$ . This establishes the statement.

If the ascending chain condition happens to hold in  $L$ , then it is also true that the notions  $P$ -primary,  $P$ - and  $P^*$ -primal are identical. The proof runs on the same lines as above with the sole modification that we can work with  $P(q)$  itself, for  $P(q)^k \leq q$ .

Calling a prime element  $p$  a minimal prime element associated with  $x$  if  $x \leq p$  but there is no prime  $p'$  with  $x \leq p' < p$ , we may define another operator  $P_1$  by putting  $P_1(x)$  equal to the meet of all minimal prime elements associated with  $x$ . (It is to be remarked that if there is no minimal prime element associated with  $x$ , then we set  $P_1(x) = e$ .) We may easily conclude that for each  $x$  we have  $P(x) \leq P_1(x)$ . As a matter of fact, if  $p$  is any minimal prime element associated with  $x$ , then for each  $a$  with  $a^k \leq x$  we have  $a^k \leq p$  whence  $a \leq p$ , and this implies the assertion.

Now we show that in case  $L$  satisfies the ascending chain condition, then the converse is also true:  $P_1(x) \leq P(x)$ . For, by the assumed condition,  $P(x)^k \leq x$ , so that from  $b \not\leq P(x)$  it follows that no power of  $b$  is  $\leq x$ , consequently, again by hypothesis, there is a maximal  $p$  such that  $x \leq p$  but  $b^l \not\leq p$  for no  $l$ . This  $p$  is a prime, for  $a_j \not\leq p$  implies  $b^{l_j} \leq a_j \cup p$  for some  $l_j$  and therefore, from  $a_1 a_2 \leq p$  we should obtain  $b^{l_1} b^{l_2} \leq (a_1 \cup p)(a_2 \cup p) \leq p$ , a contradiction to the choice of  $p$ . Hence  $P(x)$  is the meet of all primes  $p$  with  $x \leq p$ . What remains to be proved is that to each prime  $p$  with  $x \leq p$  there exists a minimal prime  $p'$  associated with  $x$  such that  $p' \leq p$ . The elements  $\not\leq p$  form, by the definition of primeness, a semigroup  $S$  in  $L$ . ZORN's lemma guarantees the existence of a maximal semigroup  $S^*$  containing  $S$  such that the elements in  $S^*$  are  $\not\leq x$ . If  $p'$  is an element maximal with respect to the property that the elements of  $S^*$  are  $\not\leq p'$ , then by making use of the same



argument as above in the proof of the primeness of  $p$  we conclude that this  $p'$  is a prime element and it is obvious that it must be a minimal one. This establishes our assertion.

Hence we arrive at the result: *if  $L$  satisfies the ascending chain condition, then  $P(x) = P_1(x)$  for all  $x \in L$ .*

III. Let  $A_r(x)$  be the join of all  $a$  with the property<sup>15</sup>

$$(16) \quad x:(a \cup b) = x \text{ (for some } b) \text{ implies } x:b = x.$$

We observe that  $A_r(x)$  may alternatively be defined as the set of all  $a$  such that  $x:(a \cup b) > x$  whenever  $x:b > x$ . Hence it is immediately seen that if  $x \neq e$  we must have  $x:a > x$  for  $a$  in (16). The operation  $A_r$  may analogously be defined by using left residual in place of right residual.

In the presence of the ascending chain condition,  $A_r(x)$  may be characterized by the property that it is the (only) maximal element  $a$  with property (16). Indeed,  $A_r(x)$  possesses this property, for by hypothesis  $A_r(x) = a_1 \cup \dots \cup a_n$  where the elements  $a_i$  have property (16), and thus  $x:(a_1 \cup a_2 \cup \dots \cup a_n \cup b) = x$  implies  $x:(a_1 \cup \dots \cup a_n \cup b) = x, \dots$ , finally,  $x:(a_n \cup b) = x$  implies  $x:b = x$ , establishing the statement.

Next we show that any  $A_r$ -primary element  $x$  is at the same time right  $A_r$ -primal. If  $x$  is  $A_r$ -primary and  $x:a > x$ , then  $(x:a)a \leq x$  implies  $a \leq A_r(x)$ . But  $A_r(x)$  is by definition the join of certain  $b$  with  $x:b > x$ . Therefore,  $A_r(x)$  is actually the join of all  $b$  with  $x:b > x$ , i. e.  $x$  is right  $A_r$ -primal. By our remark in the preceding paragraph it follows at once that *if in  $L$  the ascending chain condition holds, then any  $A_r$ -primary element is right  $A_r^*$ -primal.*

Each meet-irreducible element<sup>16</sup> is right  $A_r$ -primal. Let  $x$  be a meet-irreducible element and  $a$  an element such that  $x:a > x$ . If we have  $x:(a \cup b) = x$ , or what is the same,  $x:a \cap x:b = x$ , then the meet-irreducibility of  $x$  implies  $x:b = x$ . Consequently, all  $a$  with  $x:a > x$  have property (16), and therefore  $A_r(x)$  is the join of all  $a$  with  $x:a > x$ , i. e.  $x$  is right  $A_r$ -primal.

This is the right place to observe that *if  $x$  is a strongly meet-irreducible element and is right  $\Psi$ -primal, then it is right  $\Psi^*$ -primal too.* For, putting  $\Psi(x) = \bigcup_{\alpha} a_{\alpha}$  with  $x:a_{\alpha} > x$ , it is impossible to have  $x:\Psi(x) = x$ , for hence  $x = x:(\bigcup_{\alpha} a_{\alpha}) = \bigcap_{\alpha} (x:a_{\alpha})$ , in contradiction to the hypothesis on  $x$ .

The last two remarks imply: *each strongly meet-irreducible element is right  $A_r^*$ -primal.*

<sup>15</sup> Cf. CURTIS [2] and FUCHS [4].

<sup>16</sup> We call  $x$  *meet-irreducible* if  $x = a_1 \cap \dots \cap a_n$  implies  $a_i = x$  for at least one  $i$ ;  $x$  is *strongly meet-irreducible* if  $x = \bigcap_{\alpha} a_{\alpha}$  ( $\alpha$  runs over a finite or infinite set of indices) implies  $a_{\alpha} = x$  for a certain  $\alpha$ .



We shall call  $p$  a *right maximal element associated with  $x$*  if 1)  $p$  is the join of some  $a_\alpha$  such that for a finite number of the  $a_\alpha$  one has

$$(17) \quad x:(a_{\alpha_1} \cup a_{\alpha_2} \cup \dots \cup a_{\alpha_n}) > x$$

and 2) no  $p' > p$  exists with property 1).

In view of  $x:a = x:(a \cup x)$ , it results that for any right maximal element  $p$  associated with  $x$  one has  $x \leq p$ . By ZORN'S lemma (or the equivalent lemma of TUKEY) one gets at once that for each  $a$  with  $x:a > x$  there exists at least one right maximal element  $p$  associated with  $x$  so that  $a \leq p$ .

By familiar arguments we may infer that if  $L$  enjoys the ascending chain condition, then  $p$  may be defined as an element maximal with respect to the property  $x:p > x$ . In this case it may readily be checked that  $p$  is a prime element. For, assume  $a_1 a_2 \leq p$  whence  $(a_1 \cup p)(a_2 \cup p) \leq p$ . Therefore

$$x < x:p \leq x: [(a_1 \cup p)(a_2 \cup p)] = [x:(a_2 \cup p)]:(a_1 \cup p).$$

This ensures the existence of an element  $a_1$  with the property  $x:(a_1 \cup p) > x$ . This, together with the maximality of  $p$ , leads us to the conclusion  $a_1 \cup p \leq p$ , i. e.  $a_1 \leq p$ , as we wished to prove.

By making use of the last notion we may introduce an operator  $A'_r$  by putting the element  $A'_r(x)$  equal to the meet of all right maximal elements associated with  $x$ . Then we have  $A_r(x) \leq A'_r(x)$ . As a matter of fact, let  $p$  denote a right maximal element associated with  $x$  for which  $A_r(x) \leq p$  does not hold. Then there exists an element  $a$  for which (16) holds, although  $a \leq p$  is not true. By the definition of  $p$ , there are elements  $a_1, \dots, a_m \leq p$  satisfying both  $x:(a_1 \cup \dots \cup a_m) > x$  and  $x:(a \cup a_1 \cup \dots \cup a_m) = x$ . But this contradicts (16).

If the ascending chain condition is assumed to hold in  $L$ , then the converse is also true:  $A'_r(x) \leq A_r(x)$ . If  $b$  is not right prime to  $x$ ,  $x:b > x$ , then  $b \leq p$  for some right maximal element  $p$  associated with  $x$ . Thus  $A'_r(x) \cup b \leq p$ , and so  $x:(A'_r(x) \cup b) \geq x:p > x$ . This involves the statement.

Thus we have proved:<sup>15</sup> if  $L$  satisfies the ascending chain condition, then  $A_r(x) = A'_r(x)$  for all  $x \in L$ .

**IV.** Let  $p$  be an arbitrary but fixed prime ( $\neq e$ ) and let  $\Pi_p(x)$  be the join of all  $a$  having the property: there is some  $c \not\leq p$  such that  $ac \leq x$  (i. e.  $x::a \not\leq p$ ).<sup>17</sup> It is evident that  $x \not\leq p$  implies  $\Pi_p(x) = e$  (because  $ex \leq x$ ), but for no  $x$  with  $x \leq p$  can  $\Pi_p(x) = e$  hold, as in this case clearly  $\Pi_p(x) \leq p$ .

If in  $L$  the maximal condition holds, then  $\Pi_p(x)$  coincides with the (unique) maximal element  $a$  with the property  $ac \leq x$  for some  $c \not\leq p$ . In fact, by the ascending chain condition we have  $\Pi_p(x) = a_1 \cup \dots \cup a_n$  where  $a_i c_i \leq x$  for some  $c_i \not\leq p$ . Taking  $c = c_1 \dots c_n$ , we obtain  $c \not\leq p$  and

$$\begin{aligned} \Pi_p(x) \cdot c &= (a_1 \cup \dots \cup a_n) c_1 \dots c_n = (a_1 c_1 \dots c_n) \cup \dots \cup (a_n c_1 \dots c_n) \leq \\ &\leq a_1 c_1 \cup \dots \cup a_n c_n \leq x, \end{aligned} \qquad \text{q. e. d.}$$

<sup>17</sup> Cf. the principal component of KRULL [9].



The following general observation might be of some interest. The operator  $\Psi$  occurring in this assertion is completely arbitrary.

$\Pi_p(x) \leq \bigcap_{\alpha} r_{\alpha}$  holds where  $r_{\alpha}$  varies over all right  $\Psi$ -*primal* elements  $r_{\alpha}$  satisfying  $x \leq r_{\alpha}$  and  $\Psi(r_{\alpha}) \leq p$ .

Let  $a$  have the property:  $ac \leq x$  for some  $c \leq p$ . Then for  $r_{\alpha}$  defined in the theorem we have  $ac \leq r_{\alpha}$ , and by  $c \leq \Psi(r_{\alpha})$  and the definition of right  $\Psi$ -primality we conclude  $r_{\alpha} : c = r_{\alpha}$ . Consequently,  $a \leq r_{\alpha}$  and hence  $\Pi_p(x) \leq \bigcap_{\alpha} r_{\alpha}$  as stated.

### § 5. The intersection of elements with the same $\Psi$ -property

We proceed to the problem which consists in finding a necessary and sufficient condition under which the meet of a finite number of elements with a certain  $\Psi$ -property has again the same  $\Psi$ -property. In order to get satisfactory criteria, it seems to be necessary to suppose in the case of  $\Psi$ -prime and  $\Psi$ -primary elements that the operator satisfies (11). Since in the mentioned cases we shall be concerned exclusively with such operators, in order to avoid the continuous reference to property (11), we shall introduce a new notation for operators satisfying (11): in what follows  $\Phi$  will mean an operator such that

$$(18) \quad x \leq \Phi(y) \text{ implies } \Phi(x) \leq \Phi(y).$$

It is to be noted that we need (18) only in cases in which both  $x$  and  $y$  have the same  $\Phi$ -property.

We begin with the results on  $\Phi$ -primes.

**THEOREM 1.** *The meet of a finite number of  $\Phi$ -primes,  $p = p_1 \cap \dots \cap p_n$ , is  $\Phi$ -prime again if and only if [ $\Phi(p_i) \geq \Phi(p)$  for all  $i$  and]  $\Phi(p_i) = \Phi(p)$  for at least one  $i$ .*

Suppose that  $p_1, \dots, p_n$  are  $\Phi$ -primes and, for example,  $\Phi(p_1) = \Phi(p)$ . If  $x_1 \dots x_m \leq p$ , then also  $x_1 \dots x_m \leq p_1$  whence  $x_j \leq \Phi(p_1) = \Phi(p)$  for some  $j$ .

Conversely, if  $p_1, \dots, p_n, p$  are  $\Phi$ -primes, then  $p_1 \dots p_n \leq p_1 \cap \dots \cap p_n = p$  implies  $p_j \leq \Phi(p)$  for some  $j$ . By (18) we get  $\Phi(p_j) \leq \Phi(p)$ . Further, in view of (14) we have  $p \leq p_i \leq \Phi(p_i)$  whence, using (18) again, we obtain  $\Phi(p) \leq \Phi(p_i)$  for all  $i$ . This completes the proof.

**THEOREM 2.** *An irredundant meet of a finite number of  $\Phi$ -primary elements,  $q = q_1 \cap \dots \cap q_n$ , is again  $\Phi$ -primary if and only if  $\Phi(q_i) = \Phi(q)$  for every  $i = 1, \dots, n$ .*

Assume  $q = q_1 \cap \dots \cap q_n$  with  $\Phi$ -primary components satisfying  $\Phi(q_i) = \Phi(q)$  for  $i = 1, \dots, n$ . If  $x_1 \dots x_m \leq q$  and  $x_j \leq q$ , then for at least one index  $i$  we have  $x_j \leq q_i$  and  $x_1 \dots x_m \leq q_i$ . By the  $\Phi$ -primary property of



$q_i$  we conclude  $x_k \leq \Phi(q_i) = \Phi(q)$  for some  $k = 1, \dots, j-1, j+1, \dots, m$ , establishing that  $q$  is again  $\Phi$ -primary.

Let conversely the irredundant meet  $q = q_1 \cap \dots \cap q_n$  be again  $\Phi$ -primary where all of the components are  $\Phi$ -primary. For each  $i$  we have

$$(q_1 \cap \dots \cap q_{i-1} \cap q_{i+1} \cap \dots \cap q_n) q_i \leq q$$

where the first factor is not  $\leq q$ , as irredundancy was assumed. By the  $\Phi$ -primary character of  $q$  we obtain  $q_i \leq \Phi(q)$  whence by (18)  $\Phi(q_i) \leq \Phi(q)$ . On the other hand, (14) implies  $q \leq q_i \leq \Phi(q_i)$  whence  $\Phi(q) \leq \Phi(q_i)$ , q. e. d.

In the case of  $\Psi$ -primal and  $\Psi^*$ -primal elements we must restrict ourselves to *reduced* intersections.

**THEOREM 3.** *A reduced intersection of right  $\Psi$ -primal elements,  $r = r_1 \cap \dots \cap r_n$ , is right  $\Psi$ -primal again if and only if  $\Psi(r) = \Psi(r_1) \cup \dots \cup \Psi(r_n)$ .*

The proof will easily follow from the following observation. *a is right prime to r if and only if a is right prime at least to one of  $r_1, \dots, r_n$ .* That  $r:a > r$  implies  $r_i:a > r_i$  for some  $i$ , is by  $r:a = r_1:a \cap \dots \cap r_n:a$  evident. The converse is obtained from the reducedness of the representation of  $r$ .

Turning to the proof of Theorem 3, it is immediate that if  $r = r_1 \cap \dots \cap r_n$  is a reduced representation and  $\Psi(r) = \bigcup_i \Psi(r_i)$ , then  $\Psi(r)$  is the join of all  $a$  with  $r_i:a > r_i$  for a certain  $i$ , i. e., in view of the preceding observation, the join of all  $a$  with  $r:a > r$ . Hence  $r$  is right  $\Psi$ -primal. On the other hand, if  $r$  is right  $\Psi$ -primal, then  $\Psi(r)$  is the join of all  $a$  with  $r:a > r$ , i. e. the join of all  $\Psi(r_i)$  ( $i = 1, \dots, n$ ), in fact.

As a trivial consequence of Theorem 3 we mention:

**COROLLARY.** *A reduced intersection of right  $\Phi$ -primal elements,*

$$r = r_1 \cap \dots \cap r_n,$$

*is right  $\Phi$ -primal again if and only if  $\Phi(r) = \Phi(r_1) = \dots = \Phi(r_n)$ .*

As a matter of fact, in the case of operators  $\Phi$  we have  $\Phi(r) \leq \Phi(r_i)$  for all  $i$ . The rest of the proof is in view of the preceding theorem obvious.

**THEOREM 4.** *A reduced intersection of right  $\Psi^*$ -primal elements,  $s = s_1 \cap \dots \cap s_n$ , is right  $\Psi^*$ -primal if and only if among the  $\Psi(s_i)$  there is a maximal one,  $\Psi(s_j)$ , such that  $\Psi(s_j) \geq \Psi(s_i)$  for all  $i$  and besides  $\Psi(s) = \Psi(s_j)$ .*

If  $s_1, \dots, s_n, s$  are all right  $\Psi^*$ -primal, then it follows as above that we must have  $\Psi(s) = \bigcup_i \Psi(s_i)$ . But since  $s:\Psi(s) > s$ , there exists an  $s_j$  with  $s_j:\Psi(s) > s_j$ , that is to say,  $\Psi(s) \leq \Psi(s_j)$ . Consequently, we obtain  $\Psi(s) = \Psi(s_j)$  and  $\Psi(s_i) \leq \Psi(s_j)$  for all  $i$ .

To prove the converse, assume  $s_1, \dots, s_n$  are right  $\Psi^*$ -primal and  $\Psi(s_j)$  has the properties in question. Then  $s:x > s$  is equivalent to  $x \leq \Psi(s_j) = \Psi(s)$ , q. e. d.



We omit the formulation of the corollary got from Theorem 4 by specialization of the operator  $\Psi$  to operator  $\Phi$ , since we shall not need it in the sequel.

### § 6. Representations by means of elements with certain $\Psi$ -property

Our aim in this section is to examine those elements of  $L$  which can be represented as a finite number of elements having the same  $\Psi$ -property. The main problem is to get certain unicity statements for such representations. Since in the present general discussion we have not assumed anything on the  $\Psi$ -element associated with the meet of elements, we can not in general conclude from the  $\Psi$ -elements of the components to the  $\Psi$ -element of their intersection. This is the reason why we are inclined to introduce the following notions in terms of which we shall be able to state unicity statements analogous to those obtained for the ideals of commutative rings.

We shall call the element

$$(19) \quad x = x_1 \cap \cdots \cap x_n$$

**A. quasi- $\Phi$ -prime** if  $x_i$  ( $i=1, \dots, n$ ) are  $\Phi$ -primes and there is an  $x_j$  such that  $\Phi(x_j) \subseteq \Phi(x_i)$  for  $i=1, \dots, n$ ; we put  $\Phi[x] = \Phi(x_j)$  and call  $x_j$  a *dominating* element of  $x$  (viz. with respect to the representation (19));

**B. quasi- $\Phi$ -primary** if  $x_i$  are  $\Phi$ -primary and  $\Phi(x_1) = \cdots = \Phi(x_n)$ ; we set  $\Phi[x] = \Phi(x_i)$  and call each  $x_i$  a dominating element of  $x$ ;

**C. right quasi- $\Phi$ -primal** if the intersection (19) is reduced, every  $x_i$  is right  $\Phi$ -primal and  $\Phi(x_1) = \cdots = \Phi(x_n)$ ; we define  $\Phi[x] = \Phi(x_i)$  and each  $x_i$  is a dominating element of  $x$ ;

**D. right quasi- $\Psi^*$ -primal** if the representation (19) is reduced, every  $x_i$  is right  $\Psi^*$ -primal and there is an  $x_j$  such that  $\Psi(x_j) \supseteq \Psi(x_i)$  for  $i=1, \dots, n$ ; now let  $\Psi[x] = \Psi(x_j)$  and  $x_j$  a dominating element of  $x$ .

Clearly, the idea of these notions originates from Theorems 1, 2, 4 and Corollary to Theorem 3.

It is evident that if

$$(20) \quad a = p_1 \cap \cdots \cap p_m$$

is a representation of the element  $a$  as the meet of a finite number of  $\Phi$ -primes  $p_1, \dots, p_m$ , then from this representation we may get, by uniting some components in a suitable way, a *shortest quasi- $\Phi$ -prime representation*,

$$(21) \quad a = p_1^* \cap \cdots \cap p_n^*$$

where the components  $p_i^*$  are quasi- $\Phi$ -prime elements, no component  $p_i^*$  may be omitted and no subset of the  $p_i^*$  has a quasi- $\Phi$ -prime meet. In a shortest quasi- $\Phi$ -prime representation the elements  $\Phi[p_i^*]$  have the property that



$\Phi[p_i^*] \leq \Phi[p_j^*]$  implies  $i=j$  (cf. definition **A**). — All what we have now stated of  $\Phi$ -primes are also true for  $\Phi$ -primary, right  $\Phi$ - and  $\Psi^*$ -primal elements with the sole modification that in a shortest quasi- $\Phi$ -primary or right quasi- $\Phi$ -primal representation only  $\Phi[p_i^*] = \Phi[p_j^*]$  (and not  $\leq$ ) implies  $i=j$ .

Now we are ready to examine the unicity statements concerning representations by elements of certain  $\Psi$ -property.

**THEOREM 5.** *If the element  $a$  is representable as the meet of a finite number of  $\Phi$ -prime elements, then it has a shortest quasi- $\Phi$ -prime representation. In two such representations,*

$$a = p_1^* \cap \dots \cap p_n^* = x_1^* \cap \dots \cap x_m^*,$$

*the number of the components is the same:  $n=m$ , and by proper arrangement we have  $\Phi[p_i^*] = \Phi[x_i^*]$  ( $i=1, \dots, n$ ).*

Representing the elements  $p_i^*$  as the meet of  $\Phi$ -primes  $p_r$ , it is clear that the product of these  $\Phi$ -primes  $p_r$  is  $\leq a$ , and so, a fortiori, is  $\leq x_j$  where  $x_j$  is any dominating element of  $x_j^*$ . From the  $\Phi$ -prime character of  $x_j$  it follows that there is some  $r = r(j)$  with  $p_r \leq \Phi(x_j)$ . Hence  $\Phi(p_r) \leq \Phi(x_j)$ , and thus we conclude to the existence of a  $p_i^*$  with  $\Phi[p_i^*] \leq \Phi(p_r) \leq \Phi(x_j) = \Phi[x_j^*]$ . Using this  $i$ , the same reasoning leads to  $\Phi[x_i^*] \leq \Phi[p_i^*]$  for a suitable  $k = k(i)$ . Therefore  $\Phi[x_i^*] \leq \Phi[x_j^*]$ , implying, by virtue of definition **A**, that  $x_j^* \cap x_k^*$  is again quasi- $\Phi$ -prime. This contradiction vanishes if  $j=k$ . Consequently,  $\Phi[p_i^*] = \Phi[x_j^*]$  as we wished to prove.

**THEOREM 6.** *If  $a$  has a representation as the meet of a finite number of  $\Phi$ -primary elements, then it has shortest quasi- $\Phi$ -primary representations and for two such representations,*

$$a = q_1^* \cap \dots \cap q_n^* = x_1^* \cap \dots \cap x_m^*,$$

*we have  $n=m$  and the elements  $\Phi[q_i^*]$  are, up to order, equal to the elements  $\Phi[x_j^*]$ .*

Let  $\Phi[q_i^*]$  be a maximal one among  $\Phi[q_1^*], \dots, \Phi[q_n^*], \Phi[x_1^*], \dots, \Phi[x_m^*]$ , and assume there is no  $x_j^*$  with  $\Phi[q_i^*] \leq \Phi[x_j^*]$  (clearly, instead of  $\leq$  it would be sufficient to write  $=$ ). Then for each  $x_k^*$  we have  $x_k^* : q_i^* = x_k^*$ , for if we had  $>$ , then from  $q_i^* = q'_1 \cap \dots \cap q'_r$  ( $q'$  are  $\Phi$ -primary) we should get

$$(22) \quad x_k^* < x_k^* : q_i^* = x_k^* : (q'_1 \cap \dots \cap q'_r) \leq x_k^* : (q'_1 \cdot \dots \cdot q'_r) = [(x_k^* : q'_r) : q'_{r-1}] : \dots$$

and hence  $x_k^* : q'_\lambda > x_k^*$  for some  $\lambda$ . This implies  $q'_\lambda \leq \Phi[x_k^*]$  whence  $\Phi[q_i^*] = \Phi(q'_\lambda) \leq \Phi[x_k^*]$ , against hypothesis. Thus on the one hand we have

$$a : q_i^* = q_1^* : q_i^* \cap \dots \cap q_i^* : q_i^* \cap \dots \cap q_n^* : q_i^* > a$$

and on the other hand

$$a : q_i^* = x_1^* : q_i^* \cap \dots \cap x_m^* : q_i^* = x_1^* \cap \dots \cap x_m^* = a.$$

This contradiction shows that the same maximal  $\Phi[q_i^*]$  belong to the components of any quasi- $\Phi$ -primary representation of  $a$ .



Let now  $\Phi[q_1^*] = \Phi[x_1^*]$  be a maximal one in the set of all  $\Phi[q_i^*]$ ,  $\Phi[x_i^*]$ . Then  $q_1^* \cap x_1^*$  is again a quasi- $\Phi$ -primary element such that  $\Phi[q_1^* \cap x_1^*] = \Phi[q_1^*] = \Phi[x_1^*]$ . Therefore, the same argument as above leads us to  $q_i^* : (q_1^* \cap x_1^*) = q_i^*$  if  $i \neq 1$  and  $x_j^* : (q_1^* \cap x_1^*) = x_j^*$  if  $j \neq 1$ . Consequently, we have

$$a : (q_1^* \cap x_1^*) = q_1^* : (q_1^* \cap x_1^*) \cap \cdots \cap q_n^* : (q_1^* \cap x_1^*) = q_2^* \cap \cdots \cap q_n^*$$

and

$$a : (q_1^* \cap x_1^*) = x_1^* : (q_1^* \cap x_1^*) \cap \cdots \cap x_m^* : (q_1^* \cap x_1^*) = x_2^* \cap \cdots \cap x_m^*.$$

An easy induction completes the proof of our theorem.

In addition, if e. g.  $\Phi[q_n^*] = \Phi[x_n^*]$  is a minimal one in the set of all  $\Phi[q_i^*]$ ,  $\Phi[x_i^*]$ , then  $q_n^* = x_n^*$ . (This is the analogue of the known fact that the isolated primary components are uniquely determined by the ideal in a commutative ring.) In fact, let  $c = q_1^* \cap \cdots \cap q_{n-1}^* \cap x_1^* \cap \cdots \cap x_{n-1}^*$ . Using the same method as in (22), we may conclude that  $q_n^* : c > q_n^*$  would imply  $\Phi[q_n^*] \leq \Phi[q_n^*]$  (or  $\Phi[x_n^*] \leq \Phi[q_n^*]$ ) for some  $i \neq n$  ( $j \neq n$ ), contrary to hypothesis. Therefore  $q_n^* : c = q_n^*$  and, similarly,  $x_n^* : c = x_n^*$ . Thus we find that

$$a : c = q_1^* : c \cap \cdots \cap q_n^* : c = q_n^* \quad \text{and} \quad a : c = x_n^*$$

whence  $q_n^* = x_n^*$  as stated.

**THEOREM 7.** *If a is a reduced intersection of a finite number of right  $\Phi$ -primary elements, then it has shortest reduced right quasi- $\Phi$ -primary representations, and two such representations,*

$$a = r_1^* \cap \cdots \cap r_n^* = x_1^* \cap \cdots \cap x_m^*,$$

imply  $n = m$  as well as that the sets  $\Phi[r_1^*], \dots, \Phi[r_n^*]$  and  $\Phi[x_1^*], \dots, \Phi[x_m^*]$  differ merely in the order of their elements.

Since in the proof of the preceding theorem the  $\Phi$ -primary property was needed only in a form which holds also for right  $\Phi$ -primary elements, it is obvious that this proof may be regarded as a demonstration of Theorem 7 too.

**THEOREM 8.** *Let the element a have a reduced intersection of a finite number of right  $\Psi^*$ -primary elements. Then a has a shortest reduced right quasi- $\Psi^*$ -primary representation and this is unique in the sense that*

$$a = s_1^* \cap \cdots \cap s_n^* = x_1^* \cap \cdots \cap x_m^*$$

implies  $n = m$  and  $\Psi[s_i^*] = \Psi[x_i^*]$  ( $i = 1, \dots, n$ ) for appropriate ordering of the elements.

For the proof assume the elements  $s_i^*$  and  $x_j^*$  are represented as a reduced intersection of right  $\Psi^*$ -primary elements and let  $s_i(x_j)$  denote a dominating element of  $s_i^*(x_j^*)$ . Now  $\Psi(s_i) = \Psi[s_i^*]$  is not right prime to  $s_i$ . Owing to reducedness the same holds for  $s_i^*$  and hence for  $a$  in place of  $s_i$ . Consequently, there exists an element  $x_j^*$  and hence  $x_j$  to which  $\Psi[s_i^*]$  is not right



prime; thus  $\Psi[s_i^*] \leq \Psi(x_j) = \Psi[x_j^*]$ . Changing the roles of  $s$  and  $x$ , we can find an  $s_k^*$  such that  $\Psi[x_j^*] \leq \Psi[s_k^*]$ . By the definition of right quasi- $\Psi^*$ -primality and the hypothesis that the representations are shortest, it follows that  $\Psi[s_i^*] \leq \Psi[s_k^*]$  implies  $i = k$ . Hence  $\Psi[s_i^*] = \Psi[x_j^*]$  and the remaining part of the proof is obvious.

(Received 2 September 1954)

### Bibliography

- [1] G. BIRKHOFF, *Lattice theory*, Amer. Math. Soc. Coll. Publ., vol. 25, 2nd ed. (New York, 1948).
- [2] CH. W. CURTIS, On additive ideal theory in general rings, *Amer. J. Math.*, **74** (1952), pp. 687—700.
- [3] L. FUCHS, On quasi-primary ideals, *Acta Sci. Math. Szeged*, **11** (1947), pp. 174—183.
- [4] L. FUCHS, On a special property of the principal components of an ideal, *Kgl. Norske Vid. Selsk. Forh.*, **22** (1949), pp. 28—30.
- [5] L. FUCHS, On primal ideals, *Proc. Amer. Math. Soc.*, **1** (1950), pp. 1—6.
- [6] L. FUCHS, The meet-decomposition of elements in lattice-ordered semi-groups, *Acta Sci. Math. Szeged*, **12A** (1950), pp. 105—111.
- [7] W. KRULL, Axiomatische Begründung der allgemeinen Idealtheorie, *Sitzungsberichte d. phys.-med. Sozietät Erlangen*, **56** (1924), pp. 47—63.
- [8] W. KRULL, Zur Theorie der zweiseitigen Ideale in nichtkommutativen Bereichen, *Math. Zeitschrift*, **28** (1928), pp. 481—503.
- [9] W. KRULL, Idealtheorie in Ringen ohne Endlichkeitsbedingung, *Math. Annalen*, **101** (1929), pp. 729—744.
- [10] W. KRULL, *Idealtheorie*, Ergebnisse d. Math. u. Grenzgeb., IV. 3. (Berlin, 1935).
- [11] E. NOETHER, Idealtheorie in Ringbereichen, *Math. Annalen*, **83** (1921), pp. 24—66.
- [12] B. L. VAN DER WAERDEN, *Moderne Algebra*, vol. II, 2nd ed. (Berlin, 1940).



# ИССЛЕДОВАНИЕ НЕКОТОРЫХ ПРОБЛЕМ АДДИТИВНОЙ ТЕОРИИ ИДЕАЛОВ С ПОМОЩЬЮ ТЕОРИИ СТРУКТУР

Л. ФУКС (Будапешт)

## (Резюме)

Пусть  $L$  полная структура, элементы которой образуют полугруппу относительно некоторого (ассоциативного, но не обязательно коммутативного) умножения, так, что  $L$  обладает всеми свойствами, присущими структуре двусторонних идеалов кольца. Чтобы получить далеко идущие, очень удобные обобщения некоторых вопросов аддитивной теории идеалов коммутативных колец, предположим, что в  $L$  определена операция  $\Psi$ , ставящая в соответствие элементам  $x$  из  $L$  элементы  $\Psi(x)$ , принадлежащие также  $L$ . Вообще мы не предполагаем ничего специального насчет операции  $\Psi$ . На этой операции основаны следующие понятия, играющие важную роль в настоящей работе:

А. Элемент  $p$  называется  $\Psi$ -простым, если из  $a_1 \dots a_k \leq p$  для некоторого  $i$  следует  $a_i \leq \Psi(p)$ ;

В. Элемент  $q$  называется  $\Psi$ -примарным, если из  $a_1 \dots a_k \leq q$  для любого  $i$  следует или  $a_i \leq q$  или же  $a_j \leq \Psi(q)$ , где  $j$  означает одно из чисел  $1, \dots, i-1, i+1, \dots, k$ ;

С.  $r$  есть  $\Psi$ -примальный элемент справа, если  $\Psi(r)$  совпадает с соединением элементов, справа непростых относительно  $r$ ;

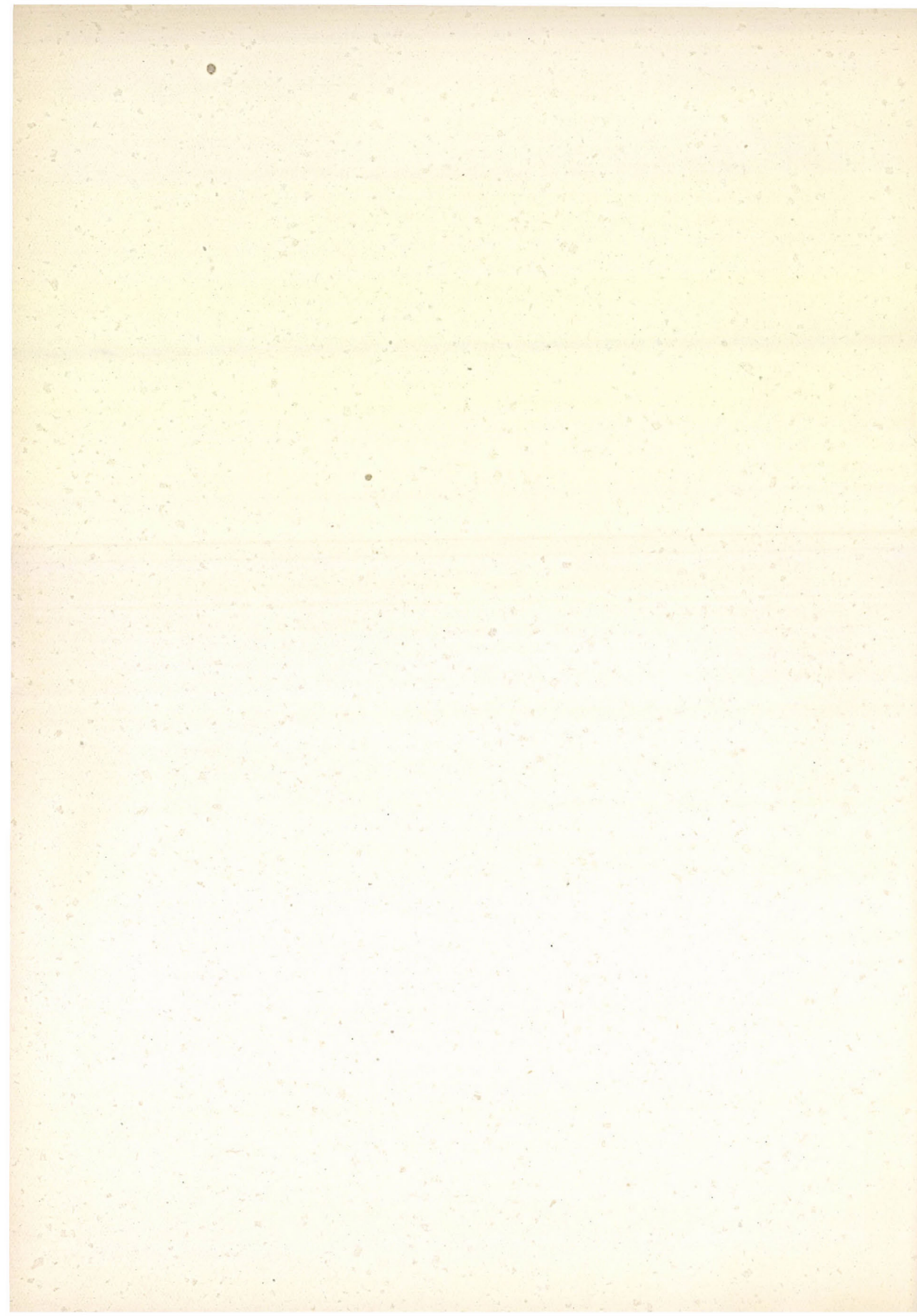
Д.  $s$  есть  $\Psi^*$ -примальный элемент справа, если  $x \leq \Psi(s)$  эквивалентно тому, что  $x$  является простым справа относительно  $s$ .

Специализируя операцию  $\Psi$ , мы приходим к различным важным частным случаям, часть которых была уже раньше подробно исследована.

В дальнейшем в работе исследуется необходимое и достаточное условие того, чтобы пересечение конечного числа элементов обладающих одним и тем же свойством  $\Psi$ , также обладало этим свойством. При проведении этих исследований в случаях А и В оказывается нужным предполагать, что  $x \leq \Psi(y)$  влечет за собой  $\Psi(x) \leq \Psi(y)$ .

В заключение автором доказывается, что для представления элементов, записывающихся в виде пересечения конечного числа элементов со свойством  $\Psi$ , справедливы теоремы единственности, очень похожие на классические результаты Нётера.







Les Acta Mathematica paraissent en russe, français, anglais et allemand et publient des mémoires du domaine des sciences mathématiques.

Les Acta Mathematica sont publiés sous forme de fascicules qui seront réunis en un volume.

On est prié d'envoyer les manuscrits destinés à la rédaction et écrits à la machine à l'adresse suivante :

*Acta Mathematica, Budapest 62, Postafiók 440.*

Toute correspondance doit être envoyée à cette même adresse.

Le prix de l'abonnement est de 110 forints (\$ 6.50) par volume.

On peut s'abonner à l'Entreprise pour le Commerce Extérieur des Livres et Journaux „Kultúra“ (Budapest, VI., Sztálin út 21. Compte-courant No. 45-790-057-50-032) ou à l'étranger chez tous les représentants ou dépositaires.

---

The Acta Mathematica publish papers on mathematics in Russian, French, English and German.

The Acta Mathematica appear in parts of various size, making up one volume.

Manuscripts should be typed and addressed to :

*Acta Mathematica, Budapest 62, Postafiók 440.*

Correspondence with the editors and publishers should be sent to the same address.

The rate of subscription to the Acta Mathematica is 110 forints (\$ 6.50) a volume. Orders may be placed with „Kultúra“ Foreign Trade Company for Books and Newspapers (Budapest, VI., Sztálin út 21. Account No. 45-790-057-50-032) or with representatives abroad.

---

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französischer, englischer und deutscher Sprache.

Die Acta Mathematica erscheinen in Heften wechselnden Umfanges. Mehrere Hefte bilden einen Band.

Die zur Veröffentlichung bestimmten Manuskripte sind, mit Maschine geschrieben, an folgende Adresse zu senden :

*Acta Mathematica, Budapest 62, Postafiók 440.*

An die gleiche Anschrift ist auch jede für die Redaktion und den Verlag bestimmte Korrespondenz zu richten.

Abonnementspreis pro Band 110 Forints (\$ 6.50). Bestellbar bei dem Buch- und Zeitungs-Außenhandels-Unternehmen „Kultúra“ (Budapest, VI., Sztálin út 21. Bankkonto Nr. 45-790-057-50-032) oder bei seinen Auslandsvertretungen und Kommissionären.



## INDEX

<i>Turán, P.</i> , On Lindelöf's conjecture. П. Туран, О гипотезе Линделефа . . . . .	145
<i>Sz.-Nagy, Gy.</i> , Konvexes Polyeder als geometrischer Ort. Д. С.-Надь, Выпуклый многогранник как геометрическое место . . . . .	165
<i>Rédei, L.</i> , Die Holomorphentheorie für Gruppen und Ringe. Л. Рэдэи, Теория голоморфов групп и колец . . . . .	169
<i>Szendrei, J.</i> , Eine neue Definition des Holomorphen der Gruppe und der Holomorphen des Ringes. Й. Сендрей, Новое определение голоморфа группы и голоморфов кольца	197
<i>Takács, L.</i> , On secondary processes generated by a Poisson process and their applications in physics. Л. Такач, О стохастических процессах, порождаемых процессом Пуассона, и об их применениях в физике . . . . .	203
<i>Tandori, K.</i> , Über die Cesàrosche Summierbarkeit der orthogonalen Polynomreihen. II. К. Тандори, О суммируемости методом Чезаро рядов ортогональных полиномов. II . . . . .	237
<i>Szász, P.</i> , Elementargeometrischer Beweis der Widerspruchsfreiheit der hyperbolischen Raumgeometrie mit Hilfe des Poincaréschen Halbraumes. П. Сас, Элементарно-геометрическое доказательство непротиворечивости гиперболической пространственной геометрии, посредством полупространства Пуанкаре . . . . .	255
<i>Coxeter, H. S. M.</i> , Arrangements of equal spheres in non-Euclidean spaces. Х. С. М. Коксетер, Размещение конгруэнтных сфер в неевклидовых пространствах . . . . .	263
<i>Freud, G.</i> , Restglied eines Tauberschen Satzes. III. Г. Фрайд, Об остаточном члене некоторой теоремы типа Таубера. III . . . . .	275
<i>Freud, G.</i> , Über orthogonale Polynome. Г. Фрайд, Об ортогональных полиномах . . . . .	291
<i>Fuchs, L.</i> , A lattice-theoretic discussion of some problems in additive ideal theory. Л. Фукс, Исследование некоторых проблем аддитивной теории идеалов с помощью теории структур . . . . .	299

Technikai szerkesztő: Molnár Ferenc

A kiadásért felelős: az Akadémiai Kiadó igazgatója

Műszaki felelős: Tóth Ferenc

A kézirat beérkezett: 1954. XI. 29. — Példányszám: 800 — Terjedelem: 15 (A/5) iv, 12 ábra.

Csongrádmegyei Nyomdaipari Vállalat, Szeged, 54-6792

Felelős vezető: Vincze György



# ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM  
HUNGARICAE

ADIUVANTIBUS

G. ALEXITS, E. EGERVÁRY, L. FEJÉR, CH. JORDAN,  
L. KALMÁR, L. RÉDEI, A. RÉNYI, F. RIESZ,  
B. SZ.-NAGY, P. TURÁN, O. VARGA

REDIGIT

G. HAJÓS

TOMUS V



MAGYAR TUDOMÁNYOS AKADÉMIA  
BUDAPEST, 1954



