ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS III. FASCICULI 1—2.

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1952

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 kötete'.

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), külföld számára pedig a „Kultúra" Könyv- és Hirlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin-út 2. Bankszámla 45-790-057-50-032 sz.) vagy külföldi képviseleteinél 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* 2. Текущий счет № 45-790-057-50-032), или его заграничные представительства и уполномоченные.

ANOTHER PROOF OF THE MARKOV—POST THEOREM*

By

LÁSZLÓ KALMÁR (Szeged), corresponding member of the Academy

The word problem for associative systems (i. e. semigroups without postulating the cancellation law) has been proved to be unsolvable by any (finite, general recursive) algorithm independently and simultaneously by POST[1] and MARKOV.[2] MARKOV's proof is simpler than that of POST's;[3] moreover, MARKOV obtained by means of his method some further results about the non-existence of some algorithms in the theory of associative systems.[4]

In the present paper, I shall give another proof for the non-existence of an algorithm to the effect of solving the word problem for associative systems. My proof seems to show some advantages over both POST's and

* Inaugural address, delivered 30 January 1950.

[1] E. L. POST, Recursive unsolvability of a problem of Thue, *The Journal of Symbolic Logic,* **12** (1947), pp. 1—11.

[2] A. Марков, Невозможность некоторых алгорифмов в теории ассоциативных систем, Доклады Академии Наук СССР, **55** (1947), pp. 587—590.

[3] Indeed, POST relies besides (1) on a theorem of CHURCH's stating the non-existence of a decision algorithm for λ-convertability (see A. CHURCH, An unsolvable problem of elementary number theory, *American Journal of Math.,* **58** (1936), pp. 345—363), (2) on ROSSER's combinatorial equivalent of the calculus of λ-conversion (see J. B. ROSSER, A mathematical logic without variables, *Annals of Math.,* (2) **36** (1935), pp. 127—150 and *Duke Math. Journal,* **1** (1935), pp. 328—355), (3) on POST's transformation method of logical systems in canonical form to those in normal form (see E. L. POST, Formal reductions of the general combinatorial decision problem, *American Journal of Math.,* **65** (1943), pp. 197—215), which have been used also by MARKOV, (4) on POST's (very easy) transformation method of logical systems in normal form to those in normal form and containing but two primitive letters (see E. L. POST, Recursively enumerable sets of positive integers and their decision problems, *Bulletin of the American Math. Society,* **50** (1944), pp. 284—316, especially footnote [5]), and (5) on TURING's theory of computing machines (see A. M. TURING, On computable numbers, with an application to the Entscheidungsproblem, *Proceedings of the London Math. Society,* (2) **42** (1937), pp. 230—265 and **43** (1937), pp. 544—546).

[4] See, besides loc. cit. [2], A. Марков, Невозможность некоторых алгорифмов в теории ассоциативных систем II, Доклады Академии Наук СССР, **58** (1947), pp. 353—356, as well as A. Марков, Невозможность некоторых алгоритмов в теории ассоциативных систем, Доклады Академии Наук СССР, **77** (1951), pp. 19—20.

MARKOV's proofs especially for readers who are not acquainted with CHURCH's theory of λ-convertability[5] nor with POST's reduction theory of logical systems.[6]

1. An *associative system* is a set in which an associative binary operation, called *multiplication*, has been defined. We denote the result of this operation, applied to the elements a and b, by ab; we call it the *product* of a and b. On account of the associative law $(ab)c=a(bc)$, a product $a_1a_2\dots a_l$ of several factors has an obvious sense (a_1 for $l=1$). We confine ourselves to associative systems containing a unit, i. e. an element e for which we have $ae=ea=a$ for every element a of the system. For $l=0$, we define the "empty product" $a_1a_2\dots a_l$ to denote e.

An equation of the form

$$a_{i_1}a_{i_2}\dots a_{i_r}=a_{j_1}a_{j_2}\dots a_{j_s} \tag{1}$$

is called a *relation;* moreover, a relation in $a_1, a_2, \dots, a_l$ if each of $i_1, i_2, \dots, i_r, j_1, j_2, \dots, j_s$ is one of $1, 2, \dots, l$. If $a_{i_1}, a_{i_2}, \dots, a_{i_r}, a_{j_1}, a_{j_2}, \dots, a_{j_s}$ denote some elements of an associative system $\mathfrak{A}$, we say, the relation (1) holds in $\mathfrak{A}$, or $\mathfrak{A}$ is satisfying (1), if $a_{i_1}a_{i_2}\dots a_{i_r}$ is the same element of $\mathfrak{A}$ as $a_{j_1}a_{j_2}\dots a_{j_s}$.

As an instance of an associative system, form the finite sequences or *"words"*, each element or *"letter"* of which is a member of a given finite set or *"alphabet"*. Two words are identical, by definition, if and only if they contain the same letters with the same multiplicity and in the same order. The product of two words is defined as the word formed of them by juxtaposition. In this associative system, called the *free associative system* on the given alphabet, the relation (1) holds if and only if it is "trivial", i. e. if $a_{i_1}a_{i_2}\dots a_{i_r}$ and $a_{j_1}a_{j_2}\dots a_{j_s}$ are identical words. However, there are associative systems in which some non-trivial relations hold; e. g. the associative systems $\mathfrak{A}_S$ defined below.

Given a finite system S of relations in $a_1, a_2, \dots, a_l$, a further relation in $a_1, a_2, \dots, a_l$ is called a *consequence* of S if and only if it holds in every associative system containing $a_1, a_2, \dots, a_l$ and satisfying each of the relations

[5] See, A. CHURCH, loc. cit. [3]; A set of postulates for the foundation of logic, *Annals of Math.*, (2) **33** (1932), pp. 346—366 and **34** (1933), pp. 839—864; A proof of freedom of contradiction, *Proceedings of the National Academy of Sciences* (Washington), **21** (1935), pp. 275–281; *Mathematical logic* (Princeton, N. J., 1936); *The calculi of lambda-conversion* (Princeton, N. J., 1941); A. CHURCH and J. B. ROSSER, Some properties of conversion, *Transactions of the American Math. Society,* **39** (1936), pp. 472—482; S. C. KLEENE, Proof by cases in formal logic, *Annals of Math.,* (2) **35** (1934), pp. 529–544; A theory of positive integers in formal logic, *American Journal of Math.,* **57** (1935), pp. 153—173 and 219—244; λ-definiability and recursiveness, *Duke Math. Journal,* **2** (1936), pp. 340—353; J. B. ROSSER, loc. cit. [3].

[6] See E. L. POST, loc. cit. [3]; A variant of a recursively unsolvable problem, *Bulletin of the American Math. Society,* **52** (1946), pp. 264—268; and loc. cit. [1].

of S. The *word problem for associative systems, relative to a given system* S *of relations*, consists in asking for an algorithm by means of which, given any relation in $a_1, a_2, \ldots, a_l$, we could decide if it is a consequence of S.

Obviously, a relation in $a_1, a_2, \ldots, a_l$ is certainly a consequence of a system S of relations in $a_1, a_2, \ldots, a_l$ if it can be obtained from the relation $e = e$ and from the relations of S by means of a finite number of multiplications by one of $a_1, a_2, \ldots, a_l$ on the left or on the right as well as applications of the transitivity law; i. e. if it is a theorem of the following formal system Φ_S. *Symbols* of Φ_S are $a_1, a_2, \ldots, a_l, =$. *Terms* of Φ_S are the words formed of the letters $a_1, a_2, \ldots, a_l$. *Formulae* of Φ_S are the relations between such words, i. e. $\mathbf{t} = \mathbf{u}$, where $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_S. *Axioms* of Φ_S are $e = e$ (e denoting the empty word) as well as the relations of S. *Theorems* of Φ_S are (i) the axioms of Φ_S; (ii) $a_1\mathbf{t} = a_1\mathbf{u}$, $a_2\mathbf{t} = a_2\mathbf{u}, \ldots, a_l\mathbf{t} = a_l\mathbf{u}$, $\mathbf{t}a_1 = \mathbf{u}a_1$, $\mathbf{t}a_2 = \mathbf{u}a_2, \ldots, \mathbf{t}a_l = \mathbf{u}a_l$ if $\mathbf{t} = \mathbf{u}$ is a theorem of Φ_S; (iii) $\mathbf{u} = \mathbf{v}$ if $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$ are theorems of Φ_S; (iv) nothing else.[7] Conversely, each consequence of the system S is a theorem of the formal system Φ_S. Indeed, the predicate[8] "$\mathbf{t} = \mathbf{u}$ is a theorem of Φ_S" between the terms (of Φ_S) is obviously an equivalence predicate, and, as easily seen, the equivalence classes of the free associative system on the alphabet $\{a_1, a_2, \ldots, a_l\}$ form an associative system[9] $\mathfrak{A}_S$ such that a relation in $a_1, a_2, \ldots, a_l$ holds in $\mathfrak{A}_S$ if and only if it is a theorem of Φ_S. Hence, the relations of the system S hold in $\mathfrak{A}_S$ and a relation which is not a theorem of Φ_S, not holding in $\mathfrak{A}_S$, is no consequence of S. Thus, the word problem relative to a system S of relations can be formulated alternatively as asking for an algorithm by means of which, given two terms $\mathbf{t}$ and $\mathbf{u}$ of Φ_S, we could decide if $\mathbf{t} = \mathbf{u}$ is a theorem of Φ_S. This formulation of the word problem is more appropriate for researches of a logical character.[10]

[7] I. e. the set of theorems of Φ_S is the smallest set (the intersection of all sets) having the axioms of Φ_S, further, together with $\mathbf{t} = \mathbf{u}$, the formulae $a_1\mathbf{t} = a_1\mathbf{u}$, $a_2\mathbf{t} = a_2\mathbf{u}, \ldots, a_l\mathbf{t} = a_l\mathbf{u}$, $\mathbf{t}a_1 = \mathbf{u}a_1$, $\mathbf{t}a_2 = \mathbf{u}a_2, \ldots, \mathbf{t}a_l = \mathbf{u}a_l$, finally, together with $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$, the formula $\mathbf{u} = \mathbf{v}$ as elements. Hence, we can prove a property of the theorems of Φ_S by proving it for the axioms of Φ_S; then, supposing that $\mathbf{t} = \mathbf{u}$ has the property in question, proving that the same holds for $a_1\mathbf{t} = a_1\mathbf{u}$, $a_2\mathbf{t} = a_2\mathbf{u}, \ldots, a_l\mathbf{t} = a_l\mathbf{u}$, $\mathbf{t}a_1 = \mathbf{u}a_1$, $\mathbf{t}a_2 = \mathbf{u}a_2, \ldots, \mathbf{t}a_l = \mathbf{u}a_l$ too, and, supposing that $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$ have the property in question, proving that the same holds for $\mathbf{u} = \mathbf{v}$ too. A similar remark applies for the set of the theorems of other formal systems (viz. $\Psi_E, \Phi_1, \Phi_2, \ldots, \Phi_5$) as well as for the set of the terms of some of them (viz. Ψ_E, Φ_1, Φ_2 and Φ_3) too.

[8] We use the expression "predicate" instead of "relation" for the latter is used in this paper in a particular sense.

[9] $\mathfrak{A}_S$ is called *the associative system generated by the system* S *of relations*. It can be characterized (irrespective of isomorphisms) by the following properties. (i) $\mathfrak{A}_S$ is an associative system each element of which can be written as a product each factor of which is one of some elements $a_1, a_2, \ldots, a_l$ of $\mathfrak{A}_S$; (ii) a relation in $a_1, a_2, \ldots, a_l$ holds in $\mathfrak{A}_S$ if and only if it is a consequence of the system S.

[10] A third formulation of the word problem relative to a system S of relations in $a_1, a_2, \ldots, a_l$, perhaps the most interesting one for the algebraist, is to ask for an algo-

2. My proof for the Markov—Post theorem is based on a theorem due to KLEENE[11] according to which a general recursive function $R(x, y)$ of two variables can be given for which there is no algorithm by means of which, given any non-negative integer k, we could decide if the equation $R(k, y)=0$ in y has a non-negative integer solution. Here a general recursive function[12] is an arithmetical function (i. e., a function with non-negative integers as arguments and values) for which, together with the successor function $F(x)=x+1$ and some additional arithmetical functions (called, together with F and R, the functions *"needed to the definition of R"*) a system E of equations (called a *"defining system of equations for R"*) can be given with the following properties. Each equation of E has to be a formula in the formal system Ψ_{E} defined below. For each function G needed to the definition of R, and for every sequence $k_1, k_2, \ldots, k_s$ of non-negative integers the number s of which coincides with the number of arguments of G, the equation $G(F^{k_1}(0), F^{k_2}(0), \ldots, F^{k_s}(0))=F^{k}(0)$ has to be a theorem of the formal system Ψ_{E} for one and only one non-negative integer k which we denote by $k=G(k_1, k_2, \ldots, k_s)$. Here $F^{k}(0)$ is an abbreviation for $F(F(\ldots F(0)\ldots))$ with k symbols F and the formal system Ψ_{E} is defined as follows. *Symbols* of Ψ_{E} are 0; an enumerable infinite set of *"variables"* $x, y, \ldots$; a finite set of symbols $F, \ldots, R$, called *"functors"*, for the functions needed to the definition of R to each of which a positive integer (for F, the integer 1) is attached as the *"number of its arguments"*;[13] parentheses (and) ; comma , ; equality symbol $=$. *Terms* of Ψ_{E} are (i) 0 and the variables; (ii) for any functor G, the sequence of symbols $G(\mathbf{t}_1 . \mathbf{t}_2, \ldots, \mathbf{t}_s)$ where s is the number of arguments of the functor G and $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ are terms of Ψ_{E}; (iii) nothing else. *Formulae* of Ψ_{E} are the sequences of symbols of the form $\mathbf{t}=\mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are terms of Ψ_{E}. *Axioms* of Ψ_{E} are the equations of E.

rithm by means of which, given any relation in $a_1, a_2, \ldots, a_l$, we could decide if it holds in the associative system $\mathfrak{A}_{\mathrm{S}}$ generated by S.

[11] S. C. KLEENE, General recursive functions of natural numbers, *Math. Annalen*, **112** (1936), pp. 727—742, especially theorem XV, p. 741. We do not make use of the fact that $R(x, y)$ is actually a *primitive recursive* function. The same result follows, with a *2-recursive* function R, from an unpublished proof of R. PÉTER for CHURCH's theorem on the existence of decision problems which are not solvable by any algorithm, by means of GÖDEL's theorem on the existence of truth problems unsolvable in a given postulate system, and, with an *elementary* function R, from an unpublished proof of mine for CHURCH's above theorem as a particular case of GÖDEL's above theorem (see also L. KALMÁR, On unsolvable mathematical problems, *Proceedings of the tenth International Congress of Philosophy* (Amsterdam, August 11—18, 1948), pp. 756—758).

[12] See KLEENE, loc. cit.[11], Definition $2b$, p. 731.

[13] We suppose that we never use the same functor for functions with a different number of variables (as $F(x)$ and $F(x, y)$, e. g.).

Theorems of Ψ_{E} are (i) the axioms of Ψ_{E}; (ii) the result of substitution of 0 or of $F(x)$ for a variable[14] x throughout a theorem $\mathbf{T}$ of Ψ_{E}; (iii) the result of replacement of a particular occurrence of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$, where $\mathbf{T}$ and $\mathbf{t}=\mathbf{u}$ are theorems of Ψ_{E}; (iv) nothing else.[15]

Instances for general recursive functions are given by $Q(x)=x!$ for which[16]

$$\begin{aligned} S(x,0)&=x,\\ S(x,F(y))&=F(S(x,y)),\\ P(x,0)&=0,\\ P(x,F(y))&=S(x,P(x,y)),\\ Q(0)&=F(0),\\ Q(F(x))&=P(Q(x),F(x)),\end{aligned}$$

further, by the Ackermann function[17] $A(x,y,z)$ for which

$$\begin{aligned} A(x,y,0)&=S(x,y),\\ A(x,0,F(0))&=0,\\ A(x,0,F^2(0))&=F(0),\\ A(x,0,F^3(z))&=x,\\ A(x,F(y),F(z))&=A(x,A(x,y,F(z)),z)\end{aligned}$$

is a defining system of equations. (Here, of course, $F^3(z)$ stands for $F(F(F(z)))$.)

3. In the sequel, let E denote the particular defining system of equations for KLEENE's general recursive function $R(x,y)$ referred to; Φ_1 denote the

[14] "For *a variable* x" (and not "for *the variable* x") means that x can be replaced by any other variable too. In a similar sense we use the clause "for a functor G". — Note that by substitution of $F(x)$ for x k times in succession and then, by substitution of 0 for x, we can substitute $F^k(0)$ for x. — Also remark that by substitution of a term for a variable throughout a term or a formula of Ψ_{E} or more generally, by replacement of one or more occurrences of some terms Ψ_{E} by other terms of Ψ_{E} in a term or a formula of Ψ_{E}, we get a term or a formula, respectively, of Ψ_{E} again. A similar remark holds for the formal systems Φ_2, and Φ_3 (and, trivially, for Φ_4 and Φ_5) too.

[15] As easily shown (by means of the proof method stated in footnote [7]), a formula $\mathbf{T}$ of Ψ_{E} is a theorem of Ψ_{E} if and only if there is a finite sequence $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_l$ of formulae of Ψ_{E} such that $\mathbf{T}_l$ is $\mathbf{T}$, and, for each $i=1,2,\ldots,l$, $\mathbf{T}_i$ is either an axiom of Ψ_{E}, or the result of substitution of 0 or of $F(x)$ for a variable x in one of $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$, or the result of replacement of a particular occurrence of $\mathbf{t}$ in one of $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$ by $\mathbf{u}$, where $\mathbf{t}=\mathbf{u}$ is one $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$. Such a sequence $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$ is called a *proof of* $\mathbf{T}$ *in the formal system* Ψ_{E}. Of course, the same holds for the formal systems $\Phi_1, \Phi_2, \ldots, \Phi_5$ too.

[16] On reasons which will become clear later on, we insist on consequent functional notation; e. g., we write $F(x), S(x,y), P(x,y), Q(x)$ rather than $x+1, x+y, xy, x!$.

[17] See W. ACKERMANN, Zum Hilbertschen Aufbau der reellen Zahlen, *Math. Annalen*, 99 (1928), pp. 118—133. $A(x,y,z)$ is an instance for a general recursive function which is not primitive recursive, whereas $Q(x)$ is a primitive recursive function.

corresponding formal system Ψ_{E}. I shall transform the question "has the equation $R(k, y)=0$ in y a non-negative integer soluton?" into the question "is the relation $\mathfrak{r}_k$ a consequence of the system S of relations?", $\mathfrak{r}_0, \mathfrak{r}_1, \ldots$ being a particular sequence of relations in the letters of some alphabet and S a particular system of relations in the same letters. Then, *the word problem for associative systems, relative to the system* S *of relations, cannot be solvable by any algorithm,* for in the opposite case, there would be an algorithm by means of which, given any non-negative integer k, we could decide if $\mathfrak{r}_k$ is a consequence of S, i. e. if the equation $R(k, y)=0$ has a non-negative integer solution y, in contradiction to KLEENE's theorem referred to.

Our first step towards this transformation is to transform the above question to an equation depending on k. To this effect, let us introduce the arithmetical function [18]

$$(2) \qquad U(x, y)=\begin{cases} 0, & \text{if there is an integer } m \geqq y \text{ for which } R(x, m)=0, \\ 1 & \text{otherwise.} \end{cases}$$

Then, the equation $R(k, y)=0$ has a non-negative integer solution if and only if we have $U(k, 0)=0$.

Now U is no general recursive function.[19] However, with the aid of the auxiliary (general recursive) function

$$(3) \qquad V(x, y)=\begin{cases} 0 & \text{if } y=0, \\ x & \text{if } y \neq 0 \end{cases}$$

it can be partially [20] defined by the system E' of equations arising from E by subjoining the equations [21]

$$(4) \qquad U(x, y)=V(U(x, F(y)), R(x, y)),$$

$$(5) \qquad V(U(x, F(y)), 0)=0,$$

$$(6) \qquad V(U(x, F(y)), F(z))=U(x, F(y)).$$

[18] We suppose, U (and V) is a functor different from those for the functions $F, \ldots, R$ needed to the definition of R.

[19] Indeed, in the opposite case, the computation algorithm of $U(k, 0)$ by means of a defining system of equations for U would furnish an algorithm for deciding if we have $U(k, 0)=0$, i. e. ifthere is a non-negative integer solution of the equation $R(k, y)=0$ in y.

[20] In a sense similar to the notion of a partial recursive function; see S. C. KLEENE, On notation for ordinal numbers, *The Journal of Symbolic Logic,* 3 (1938), pp. 150—155, especially pp. 151—152. However, the function $U(x, y)$ as defined by (2), is not partial recursive (for it is defined for all values of its arguments and it is not general recursive); E' is rather a defining system of equations for the partial recursive function

$$U(x, y)=\begin{cases} 0 & \text{if there is an integer } m \geqq y \text{ for which } R(x, m)=0, \\ & \text{undefined otherwise.} \end{cases}$$

[21] The idea of introducing these equations is due, in another form, to KLEENE; see S. C. KLEENE, A theory of positive integers in formal logic, loc. cit. [5], pp. 230—231.

We define a formal system Φ_2, analogous to Φ_1 (except a slight modification) but based on the system of equations E' instead of E as follows. *Symbols* of Φ_2 are 0, the variables $x, y, \ldots$, the functors $F, \ldots, R$ as well as U and V; parentheses (and) ; comma , ; equality symbol $=$. *Terms* of Φ_2 are (i) 0 and the variables; (ii) for any functor G, the sequence of symbols $G(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s)$ where s is the number of arguments of G and $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ are terms of Φ_2; (iii) nothing else. *Formulae* of Φ_2 are the sequences of symbols of the form $\mathbf{t} = \mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_2. *Axioms* of Φ_2 are (i) the equations of E'; (ii) the formulae of the form $\mathbf{t} = \mathbf{t}$ where $\mathbf{t}$ is a term of Φ_2. (The purpose of the modification (ii) will be clear in section **6**, lemma 9). *Theorems* of Φ_2 are (i) the axioms of Φ_2; (ii) the result of substitution of 0 or of $F(x)$ for a variable x throughout a theorem of Φ_2; (iii) the result of replacement of a particular occurrence of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$, where $\mathbf{T}$ and $\mathbf{t} = \mathbf{u}$ are theorems of Φ_2: (iv) nothing else. Then we have

LEMMA 1. *Let k denote a non-negative integer. The formula*

$$U(F^k(0), 0) = 0 \tag{7}$$

of Φ_2 is a theorem of Φ_2 if and only if the equation $R(k, y) = 0$ has a non-negative integer solution in y. Hence, there is no algorithm by means of which, given a non-negative integer k, we could decide if (7) *is a theorem of Φ_2.*

Indeed, suppose first, the equation $R(k, y) = 0$ has a solution in non-negative integers y; let l denote its least solution. Then, we have on the one hand $R(k, i) \neq 0$, i. e. $R(k, i) = r_i + 1$ for some non-negative integers r_i $(i = 0, 1, \ldots, l-1)$, on the other hand $R(k, l) = 0$. Hence, the formulae $R(F^k(0), F^i(0)) = F^{r_i+1}(0)$, i. e.

$$R(F^k(0), F^i(0)) = F(F^{r_i}(0)) \qquad (i = 0, 1, \ldots, l-1) \tag{8}$$

and

$$R(F^k(0), F^l(0)) = 0 \tag{9}$$

are theorems of Φ_1 and thus of Φ_2 too. On the other hand, by (4) (see footnote [14], second sentence),

$$U(F^k(0), F^i(0)) = V(U(F^k(0), F^{i+1}(0)), R(F^k(0), F^i(0)) \qquad (i = 0, 1, \ldots, l),$$

hence, by (8) and (9),

$$U(F^k(0), F^i(0)) = V(U(F^k(0), F^{i+1}(0)), F(F^{r_i}(0))) \qquad (i = 0, 1, \ldots, l-1), \tag{10}$$

and

$$U(F^k(0), F^l(0)) = V(U(F^k(0), F^{l+1}(0)), 0) \tag{11}$$

are theorems of Φ_2. The same holds, by (6) and (5), for

$$V(U(F^k(0), F^{i+1}(0)), F(F^{r_i}(0))) = U(F^k(0), F^{i+1}(0)) \qquad (i = 0, 1, \ldots, l-1) \tag{12}$$

and

$$V(U(F^k(0), F^{l+1}(0)), 0) = 0 \tag{13}$$

too. By (10) and (12) on the one hand, (11) and (13) on the other hand, we see that

$$\begin{aligned} &U(F^k(0), 0) = U(F^k(0), F(0)), \\ &U(F^k(0), F(0)) = U(F^k(0), F^2(0)), \\ &\quad \cdot \quad \cdot \quad \cdot \qquad\qquad \cdot \quad \cdot \quad \cdot \\ &U(F^k(0), F^{l-1}(0)) = U(F^k(0), F^l(0)), \\ &U(F^k(0), F^l(0)) = 0 \end{aligned}$$

are theorems of Φ_2; thus $U(F^k(0), 0) = 0$, i. e. (7) too.

Suppose now, (7) is a theorem of Φ_2. By $F(x) = x+1$ and the equations of E, the arithmetical functions needed to the definition of R have been defined so as to render each equation of E verifiable, i. e. true for each replacement of its variables by non-negative integers. Defining the additional arithmetical functions U and V by (2) and (3), the same holds for the equations (4), (5) and (6) too. This is obvious for (5) and (6), for we have $F(m) = m+1 \neq 0$ for each non-negative integer m. As to (4), we have for any non-negative integers k and l, in case $R(k, l) = 0$

$$U(k, l) = 0 = V(U(k, F(l)), 0) = V(U(k, F(l), R(k, l)),$$

whereas in case $R(k, l) \neq 0$ we have

$$U(k, l) = U(k, l+1) = U(k, F(l)) = V(U(k, F(l)), R(k, l)),$$

for either we have $R(k, m) = 0$ for some integer $m \geqq l+1$, thus $U(k, l) = U(k, l+1) = 0$, or $R(k, m) \neq 0$ for every integer $m \geqq l$ (for we have $R(k, l) \neq 0$), thus $U(k, l) = U(k, l+1) = 1$. Also, obviously, the equations of the form $\mathbf{t} = \mathbf{t}$ are verifiable.

Obviously, substitution of 0 or of $F(x)$ for a variable x in a verifiable formula of Φ_2 yields a verifiable formula again; also, if $\mathbf{T}$ and $\mathbf{t} = \mathbf{u}$ are verifiable formulae of Φ_2, then so is the formula obtained by replacement of a particular occurrence of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$. Hence, each theorem of Φ_2 is verifiable, thus, by hypothesis, (7) too; i. e. there is an integer $n \geqq 0$ such that $R(k, n) = 0$ as stated.

4. The second part of the proof of lemma 1 (from "suppose now" on), besides being somewhat sketchy, does not fulfil the requirements of proof theory for it does not provide a calculation method of a non-negative integer n such that $R(k, n) = 0$ by means of a given proof of the formula $U(F^k(0), 0) = 0$ in the formal system Φ_2. Indeed, the notion of verifiability as defined above is not a constructive one, owing to the non-constructive definition (2) of $U(x, y)$. For those readers who are interested in constructiveness distinctions, I give a proof-theoretical alternative for the second part of the proof. The other readers may continue reading with section **5.**

Let us call the terms $F^n(0)$ of Φ_2 $(n = 0, 1, \ldots)$ *numerals*. A term or a formula obtained from a term $\mathbf{t}$ or a formula $\mathbf{T}$ of Φ_2 by substituting a numeral for each of its variables is called a *numerical instance* of $\mathbf{t}$ or of $\mathbf{T}$, respectively. A term or a formula containing no variables is called a numerical term or formula;[22] it is the only numerical instance of itself. E. g., (7) is a numerical formula. A finite sequence $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_l$ of (numerical) formulae such that, for $i = 1, 2, \ldots, l$, $\mathbf{T}_i$ is either a numerical instance of an axiom of Φ_2, or the result of replacement of a particular occurrence of $\mathbf{t}$ in one of $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$ by $\mathbf{u}$, where $\mathbf{t} = \mathbf{u}$ is one of $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{i-1}$, is called a *numerical proof of* $\mathbf{T}_l$. Our arguments are based on

LEMMA 2. *If* $\mathbf{T}$ *is a theorem of* Φ_2, *and* $\overline{\mathbf{T}}$ *a numerical instance of* $\mathbf{T}$, *then there is a numerical proof of* $\overline{\mathbf{T}}$.

Indeed,[23] this holds if $\mathbf{T}$ is an axiom of Φ_2, for then, $\overline{\mathbf{T}}$ alone forms a numerical proof of $\overline{\mathbf{T}}$. Suppose, lemma 2 holds for a theorem $\mathbf{T}$ of Φ_2 and let $\mathbf{T}'$ and $\mathbf{T}''$ be the results of substitution of 0 and of $F(x)$, respectively, for a variable x throughout $\mathbf{T}$. Then, lemma 2 holds for $\mathbf{T}'$ and $\mathbf{T}''$ too. Indeed, any numerical instance of $\mathbf{T}'$ or $\mathbf{T}''$ is a numerical instance of $\mathbf{T}$ too. For if $x, y, \ldots, v$ are the variables of $\mathbf{T}$, then substitution of the numerals $F^n(0), \ldots, F^q(0)$ for the variables $y, \ldots, v$, respectively, in $\mathbf{T}'$ amounts to substitution of $0, F^n(0), \ldots, F^q(0)$ for the variables $x, y, \ldots, v$, respectively, in $\mathbf{T}$, whereas substitution of the numerals $F^m(0), F^n(0), \ldots, F^q(0)$ for the variables $x, y, \ldots, v$, respectively, in $\mathbf{T}''$ amounts to substitution of $F^{m+1}(0)$, $F^n(0), \ldots, F^q(0)$ for the variables $x, y, \ldots, v$, respectively, in $\mathbf{T}$. Again, suppose, lemma 2 holds for some theorems $\mathbf{T}$ and $\mathbf{T}'$ of Φ_2, and let $\mathbf{T}''$ be the result of replacement of a particular occurrence of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$, where $\mathbf{T}'$ is $\mathbf{t} = \mathbf{u}$. Then lemma 2 holds for $\mathbf{T}''$ too. Indeed, let $\overline{\mathbf{T}''}$ be any numerical instance of $\mathbf{T}''$. Substitute in $\mathbf{T}$ as well as in $\mathbf{T}'$ (i. e., in $\mathbf{t}$ and $\mathbf{u}$) the same numerals for the variables figuring in $\mathbf{T}''$, (in so far as they figure in $\mathbf{T}$ or in $\mathbf{T}'$) as when forming $\overline{\mathbf{T}''}$ out of $\mathbf{T}''$, whereas for the other variables figuring in $\mathbf{T}$ and $\mathbf{T}'$, substitute 0, say. (There are such variables if and only if there is a single occurrence of $\mathbf{t}$ in $\mathbf{T}$ and some variables figure in $\mathbf{t}$ but neither

[22] Obviously, the notion of a numerical term and of a numerical formula could be defined as follows. Numerical terms are (i) 0; (ii) for any functor G, the sequence of symbols $G(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s)$, where s is the number of arguments of G and $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ are numerical terms; (iii) nothing else. Numerical formulae are the sequences of symbols of the form $\mathbf{t} = \mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are numerical terms. Hence, a function can be defined over the class of the numerical terms by defining it for 0 and, supposing its value known for $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$, by defining it for $G(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s)$, where G is a functor having s as its number of arguments.

[23] The proof is a variant of HILBERT's method of removing back the substitutions (Rückverlegung der Einsetzungen), however, without dissolution into proof-files (Auflösung in Beweisfäden); see D. HILBERT and P. BERNAYS, *Grundlagen der Mathematik*, I (Berlin, 1934), pp. 221—228.

in $\mathbf{u}$, nor in $\mathbf{T}$ outside the occurrence of $\mathbf{t}$.) Thus, we get a numerical instance $\overline{\mathbf{T}}$ of the formula $\mathbf{T}$ as well as numerical instances $\bar{\mathbf{t}}$ and $\bar{\mathbf{u}}$ of the terms $\mathbf{t}$ and $\mathbf{u}$, respectively; and obviously, $\overline{\mathbf{T}''}$ is the result of replacement of some occurrence of $\bar{\mathbf{t}}$ in $\overline{\mathbf{T}}$ by $\bar{\mathbf{u}}$. By hypothesis, there are numerical proofs of $\overline{\mathbf{T}}$ as well as of the numerical instance $\bar{\mathbf{t}}=\bar{\mathbf{u}}$ of T′. By juxtaposing them and subjoining the formula $\overline{\mathbf{T}''}$, we get a numerical proof of $\overline{\mathbf{T}''}$. Hence, lemma 2 holds for every theorem $\mathbf{T}$ of Φ_2.

Given a non-negative integer k, we attach to each numerical term a non-negative integer, called its value, by means of the following definition. The value of 0 is 0 (zero). If G is one of the functors $F, \ldots, R$ for the functions needed to the definition of R, and s is the number of its arguments, then if $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ are numerical terms whose values are $k_1, k_2, \ldots, k_s$, respectively, we define the value of the numerical term $G(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s)$ as the only non-negative integer l for which $G(F^{k_1}(0), F^{k_2}(0), \ldots, F^{k_s}(0)) = F^l(0)$ is a theorem of Φ_1. If $\mathbf{t}$ and $\mathbf{u}$ are numerical terms the values of which are m and n, respectively, then we define the value of the numerical term $U(\mathbf{t}, \mathbf{u})$ to be 1 or 0 according as $m=k$ or $m \neq k$ and the value of the numerical term $V(\mathbf{t}, \mathbf{u})$ to be 0 or m according as $n=0$ or $n \neq 0$. Obviously, the value of a numerical term $\mathbf{t}$ does not change if an occurrence of a term in $\mathbf{t}$ is replaced by another numerical term with the same value. We call a numerical formula $\mathbf{t}=\mathbf{u}$ *true* or *false* according as the numerical terms $\mathbf{t}$ and $\mathbf{u}$ have the same value or not.

Suppose now, the numerical formula (7) is a theorem of Φ_2. Then, by lemma 2, there is a numerical proof of (7). Let $\mathbf{T}$ be the first formula of this numerical proof which is false (there is such a formula for (7) is false, the value of its left-hand side being 1 and that of its right-hand side 0). Then, $\mathbf{T}$ is a numerical instance of an axiom of Φ_2. For in the opposite case, there would be numerical formulae $\mathbf{T}'$ and $\mathbf{t}=\mathbf{u}$, belonging to, and prior to $\mathbf{T}$ in, the numerical proof in question, thus true, such that $\mathbf{T}$ is the result of replacement of a particular instance of $\mathbf{t}$ in $\mathbf{T}'$ by $\mathbf{u}$. However, this is impossible, for, $\mathbf{t}=\mathbf{u}$ being true, $\mathbf{t}$ and $\mathbf{u}$ have the same value; thus, the left and the right hand side of $\mathbf{T}$ have the same value as those of $\mathbf{T}'$; but, $\mathbf{T}'$ being true, its left and right hand side have the same value and thus, the same holds for $\mathbf{T}$ too; i. e. $\mathbf{T}$ would be true.

Now, the numerical instances of the equations of E are true in consequence of the unicity requirement in the definition of general recursive functions. Also, the numerical instances of the axioms (5) and (6) are true by the definition of the value of a numerical term of the form $V(\mathbf{t}, \mathbf{u})$. Plainly, also the numerical instances of an equation of the form $\mathbf{t}=\mathbf{t}$ are true. Hence, $\mathbf{T}$ must be a numerical instance of the axiom (4), i. e. a formula of the form $U(F^m(0), F^n(0)) = V(U(F^m(0), F^{n+1}(0)), R(F^m(0), F^n(0)))$. Here, we must have $m=k$, for in the opposite case, $U(F^m(0), F^n(0))$ and $U(F^m(0), F^{n+1}(0))$

would have the value 0, hence the same would hold for $V(U(F^m(0), F^{n+1}(0)), R(F^m(0), F^n(0)))$ independently of the value of $R(F^m(0), F^n(0))$ and thus, **T** would be true. Hence, **T** is of the form $U(F^k(0), F^n(0)) = V(U(F^k(0), F^{n+1}(0)), R(F^k(0), F^n(0)))$. Here, $U(F^k(0), F^n(0))$ and $U(F^k(0), F^{n+1}(0))$ have the value 1; hence, $R(F^k(0), F^n(0))$ has the value 0, for in the opposite case, $V(U(F^k(0), F^{n+1}(0)), R(F^k(0), F^n(0)))$ would have the value 1 and **T** would be true again. I. e., we have $R(k, n) = 0$, as stated.

5. Our second step towards the transformation of the question "has the equation $R(k, y) = 0$ in y a non-negative integer solution?" into a question of being a relation a consequence of a system of relations is the omission of all parentheses and commas from the formulae of Φ_2; i. e. the use of ŁUKASIEWICZ's notation system.[24] For this purpose, we define a counterpart Φ_3 of the formal system Φ_2 as follows. *Symbols* of Φ_3 are 0, the variables $x, y, \ldots$, and the functors $F, \ldots, R, U, V$, as well as the equality symbol $=$. *Terms* of Φ_3 are (i) 0 and the variables $x, y, \ldots$; (ii) for any functor G, the sequence of symbols $G\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_s$ where $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ are terms of Φ_3 whose number s coincides with the number of arguments of G; (iii) nothing else. Obviously, omitting the parentheses and commas from a term $\mathbf{t}$ of Φ_2, we obtain a term $\bar{\mathbf{t}}$ of Φ_3 which we call the *simplification*[25] of $\mathbf{t}$; on the other part, each term of Φ_3 is the simplification of at least one (and, as we shall show in the sequel, only one) term of Φ_2. *Formulae* of Φ_3 are the sequences of symbols of the form $\mathbf{t} = \mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_3. If $\bar{\mathbf{t}}, \bar{\mathbf{u}}$ are the simplifications of the terms $\mathbf{t}, \mathbf{u}$ of Φ_2, respectively, then we call the formula $\bar{\mathbf{t}} = \bar{\mathbf{u}}$ of Φ_3 the simplification of the formula $\mathbf{t} = \mathbf{u}$ of Φ_2. Again, each formula

[24] See J. ŁUKASIEWICZ and A. TARSKI, Untersuchungen über den Aussagenkalkül, *Sprawozdania z posiedzeń Towarzystwa Naukowego Warszawskiego,* Wydz. III, 23 (1930), pp. 30—50. For a reader, who is acquainted with ŁUKASIEWICZ's notation system as well as with its properties, I could begin the proof in writing down the system E′ of equations in this notation system and then, continue with section **6.** Alternatively, I could menage without ŁUKASIEWICZ's notation system; then, in the formal system Φ_4, I should allow to multiply both sides of an equation by *any* symbol used, also (or , or), on the left or on the right. However, such a variant of the proof would not be simpler than that of the text, and it would be very strange for the algebraist; indeed, it would amount to consider also the parentheses (and) and the comma , as elements of an associative system, whereas our procedure amounts to consider, besides 0 (which is in no sense a "zero element"), the "indeterminates" $x, y, \ldots, w$ as well as the "operators" $F, \ldots, R, U, V$ as elements of an associative system and to consider the function value $G(x, y, \ldots, v)$ as the product of $G, x, y, \ldots$, and v (e. g., $F(x)$ as the product of the operator F and the indeterminate x), which is much more familiar to the algebraist.

[25] Alternatively, we could define the simplication of a term of Φ_2 thus: (i) the simplification of 0 or of a variable x is 0 or x itself, respectively; (ii) for any functor G and for any terms $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ of Φ_3 the number of which is the same as the number of arguments of G, the simplification of $G(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s)$ is $G\bar{\mathbf{t}}_1\bar{\mathbf{t}}_2\ldots\bar{\mathbf{t}}_s$, where, for $i = 1, 2, \ldots, s$, $\bar{\mathbf{t}}_i$ is the simplification of $\mathbf{t}_i$.

of Φ_3 is the simplification of a formula of Φ_2. *Axioms* of Φ_3 are the simplifications of the equations belonging to E, as well as the formulae of the form $\mathbf{t} = \mathbf{t}$ where $\mathbf{t}$ is a term of Φ_3. *Theorems* of Φ_3 are (i) the axioms of Φ_3; (ii) the result of substitution of 0 or of Fx for a variable x throughout a theorem $\mathbf{T}$ of Φ_3; (iii) the result of replacement of a particular occurrence of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$, where $\mathbf{T}$ and $\mathbf{t} = \mathbf{u}$ are theorems of Φ_3; (iv) nothing else.

The "rules of inference" (ii) and (iii) corresponding exactly to those of the formal system Φ_2, it is obvious that the simplification of a theorem of Φ_2 is a theorem of Φ_3. In particular, by lemma 1, $UF^k00 = 0$ is certainly a theorem of Φ_3 if the equation $R(k, y) = 0$ has a non-negative integer solution in y. (Here F^k0 is an abbreviation for the term $FF \cdots F0$ of Φ_3 with k symbols F; also, we shall use UF^k0^2 as an abbreviation of UF^k00.) However, the converse is not obvious at all; as a matter of fact, it is a consequence of the point of ŁUKASIEWICZ's notation system, viz. its unequivocality.[26] To prove it, the notion of the *valency* of a symbol (except $=$) of Φ_3, as defined below, will be useful. The valency of 0 or of one of the variables $x, y, \ldots$ is, by definition, the number -1; the valency of a functor G is, by definition, $s-1$, where s is the number of arguments of G. Also, we attach to a finite sequence of symbols, except $=$, of Φ_3, the sum of valencies of its members as its valency. Throughout this section, we call such a sequence a *word*; for any word $\alpha_1 \alpha_2 \cdots \alpha_l$ ($\alpha_1, \alpha_2, \ldots, \alpha_l$ symbols of Φ_3 except $=$), we call the words[27] e (the empty word), $\alpha_1, \alpha_1 \alpha_2, \ldots, \alpha_1 \alpha_2 \ldots \alpha_{l-1}$ the *proper sections of* $\alpha_1 \alpha_2 \ldots \alpha_l$. Then we have

LEMMA 3. *Each term of Φ_3 has the valency*[28] -1 *whereas its proper sections have non-negative valencies.*

Indeed, this holds obviously for 0 and the variables $x, y, \ldots$ Supposing, the statement of the lemma holds for the terms $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_s$ of Φ_3 and G is one of the functors $F, \ldots, R, U, V$, having s as the number of its argu-

[26] According to my knowledge ŁUKASIEWICZ did not publish any proof of the unequivocality of his notation system. The note of K. MENGER, Eine elementare Bemerkung über die Struktur logischer Formeln, *Ergebnisse eines math. Kolloquiums*, **3** (1930—31), pp. 22—23, furnishes a proof of the unequivocality in question for a particular case. The general case is treated in H. B. CURRY, *Leçons de logique algébrique* (Paris—Louvain, 1952), pp. 143—145 as well as in P. C. ROSENBLOOM, *Elements of mathematical logic* (New York, 1950), pp. 153—157 and 205. (The work of CURRY has appeared after the manuscript of this paper has been completed; as to the work of ROSENBLOOM, I got aware of its existence by a quotation of CURRY, loc. cit, p. 142.)

[27] We do not make any distinction between a symbol and the word formed of that symbol alone.

[28] In contrast to chemistry, here "saturated compounds" have the valency -1 and not 0, for they can serve at the same time as radicals out of which further compounds can be made (by means of a functor). For the same reason, a functor which needs s terms in order to get saturated, has the valency $s-1$ and not s as it would have in chemistry.

ments, it holds for $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s$ too. Indeed, the valency of this term is $s-1\overbrace{-1-1-\dots-1}^{s\text{ times}}=s-1-s=-1$, whereas each of its proper sections (other than e) has the form $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_{i-1}\mathbf{u}$ where $i=1,2,\dots$, or s, and $\mathbf{u}$ is a proper section of $\mathbf{t}_i$; hence, its valency is $s-1\overbrace{-1-1-\dots-1}^{i-1\text{ times}}+$ $+v=s-1-(i-1)+v=s-i+v\geqq v\geqq 0$, v denoting the valency of $\mathbf{u}$. Hence, lemma 3 holds for each term of Φ_3.

LEMMA 4. *Two different terms* $\mathbf{t}$ *and* $\mathbf{u}$ *of* Φ_2 *cannot have the same term* $\bar{\mathbf{t}}$ *of* Φ_3 *as their simplifications.*

This is obvious for terms $\bar{\mathbf{t}}$ of Φ_3 containing but one symbol. Assuming it to hold for terms $\bar{\mathbf{t}}$ of Φ_3 containing less than r symbols $(r=2,3,\dots)$, suppose the term $\bar{\mathbf{t}}$ of Φ_3 containing r symbols is the simplification of the terms $\mathbf{t}$ and $\mathbf{u}$ of Φ_2. Then, the first symbol of $\mathbf{t}$ and $\mathbf{u}$ must coincide with that of $\bar{\mathbf{t}}$, viz. a functor G. Thus, $\mathbf{t}$ and $\mathbf{u}$ have the forms $G(\mathbf{t}_1,\mathbf{t}_2,\dots,\mathbf{t}_s)$ and $G(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_s)$, respectively, where s is the number of arguments of G and $\mathbf{t}_1,\mathbf{t}_2,\dots,\mathbf{t}_s,\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_s$ are terms of Φ_2, each containing less than r symbols. Hence, denoting, for $i=1,2,\dots,s$, by $\bar{\mathbf{t}}_i$ and $\bar{\mathbf{u}}_i$ the simplifications of $\mathbf{t}_i$ and $\mathbf{u}_i$, respectively, $\bar{\mathbf{t}}$ is, on the one hand $G\bar{\mathbf{t}}_1\bar{\mathbf{t}}_2\dots\bar{\mathbf{t}}_s$, on the other hand $G\bar{\mathbf{u}}_1\bar{\mathbf{u}}_2\dots\bar{\mathbf{u}}_s$. Thus, for $i=1,2,\dots,s$, $\bar{\mathbf{t}}_i$ is identical with $\bar{\mathbf{u}}_i$. For in the opposite case, let $\bar{\mathbf{t}}_i$ be the first of $\bar{\mathbf{t}}_1,\bar{\mathbf{t}}_2,\dots,\bar{\mathbf{t}}_s$ which is different from the corresponding $\bar{\mathbf{u}}_i$. Then, the words $G\bar{\mathbf{t}}_1\bar{\mathbf{t}}_2\dots\bar{\mathbf{t}}_{i-1}$ and $G\bar{\mathbf{u}}_1\bar{\mathbf{u}}_2\dots\bar{\mathbf{u}}_{i-1}$ are identical, whereas $G\bar{\mathbf{t}}_1\bar{\mathbf{t}}_2\dots\bar{\mathbf{t}}_{i-1}\bar{\mathbf{t}}_i$ and $G\bar{\mathbf{u}}_1\bar{\mathbf{u}}_2\dots\bar{\mathbf{u}}_{i-1}\bar{\mathbf{u}}_i$, both sections of $\bar{\mathbf{t}}$ (i. e. either proper sections of $\bar{\mathbf{t}}$ or identical with $\bar{\mathbf{t}}$), are different. Hence, one of $\bar{\mathbf{t}}_i$ and $\bar{\mathbf{u}}_i$ would be a proper section of the other, which is impossible, for both $\bar{\mathbf{t}}_i$ and $\bar{\mathbf{u}}_i$ have, according to lemma 3, the valency -1, whereas their proper sections have non-negative valencies. By the induction hypothesis, the identity of $\bar{\mathbf{t}}_i$ with $\bar{\mathbf{u}}_i$ $(i=1,2,\dots,s)$ implies that of $\mathbf{t}_i$ with $\mathbf{u}_i$; hence, also $\mathbf{t}$ and $\mathbf{u}$ cannot be different.

As a corollary of lemma 4, we see that two different formulae of Φ_2 cannot have the same formula of Φ_3 as their simplifications.

Now, we can prove the converse of lemma 3, i. e.

LEMMA 5. *Each word whose valency is* -1 *while ist proper sections have non-negative valencies, is a term of* Φ_3.

Indeed, this holds for words containing a single letter (0, or one of the variables $x, y, \dots$). Supposing the statement of the lemma to hold for words containing less than r letters $(r=2,3,\dots)$, let $\mathbf{w}$ be a word containing r letters and satisfying the conditions of the lemma. Then the first letter of $\mathbf{w}$, having a non-negative valency, is a functor G. Denote by s the number of arguments of G, and, for $i=0,1,2,\dots,s$, by $\mathbf{w}_i$ the first (i. e., the shortest) non-empty section of $\mathbf{w}$ (proper section or $\mathbf{w}$ itself) whose valency is less than or equal to $s-1-i$. (There is such a section for $\mathbf{w}$ has the valency $-1=s-1-s \leqq s-1-i$.) Obviously, $\mathbf{w}_0$ is G, and $\mathbf{w}_s$ is $\mathbf{w}$. For $i= =0, 1, \ldots, s-1$, $\mathbf{w}_i$ is a section of $\mathbf{w}_{i+1}$, for the valency of $\mathbf{w}_{i+1}$ is less than or equal to $s-1-(i+1)<s-1-i$, thus $\mathbf{w}_{i+1}$ cannot be shorter than $\mathbf{w}_i$. We prove that, for $i=0, 1, \ldots, s$, the valency of $\mathbf{w}_i$ is exactly $s-1-i$ (hence, for $i=0, 1, \ldots, s-1$, $\mathbf{w}_i$ is a *proper* section of $\mathbf{w}_{i+1}$). Obviously, this holds for $i=0$. For $i=1, 2, \ldots, s$, $\mathbf{w}_i$ contains more than one letter, for the first letter G of $\mathbf{w}_i$ has the valency $s-1>s-1-i$. If $\mathbf{w}_i$ had not the valency $s-1-i$, its valency would be $s-2-i$ or less; hence, omitting its last letter, we should obtain a shorter word the valency of which would be $s-1-i$ or less (for the valency of a letter is -1 or more), which is impossible.

For $i=1, 2, \ldots, s$, let $\mathbf{t}_i$ be the word the subjoining of which to $\mathbf{w}_{i-1}$ yields $\mathbf{w}_i$. Then, $\mathbf{t}_i$ has the valency -1 whereas its proper sections have non-negative valencies; for in the opposite case, subjoining a proper section having a negative valency to $\mathbf{w}_{i-1}$, we should get a word shorter than $\mathbf{w}_i$ which has a valency $s-1-i$ or less. By the hypothesis, $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}$ are terms of Φ_3, hence the same holds for $\mathbf{w}$, which is $G\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_s$.

Now, we can prove the converse of the remark made after the definition of the theorems of Φ_3, i. e.

LEMMA 6. *Each theorem* $\overline{\mathbf{T}}$ *of* Φ_3 *is the simplification of a theorem* $\mathbf{T}$ *of* Φ_2.

Indeed, this holds for the axioms of Φ_3. Suppose it to hold for a theorem $\overline{\mathbf{T}}$ of Φ_3, then it holds for the result of substitution of 0 or of Fx for a variable x in $\overline{\mathbf{T}}$ too; for it is obviously the simplification of the result of substitution of 0 or of $F(x)$, respectively, for x in the theorem $\mathbf{T}$ of Φ_2 the simplification of which is $\overline{\mathbf{T}}$. Suppose now, the lemma holds for the theorems $\overline{\mathbf{T}}$ as well as $\overline{\mathbf{t}}=\overline{\mathbf{u}}$ of Φ_2; i. e. they are the simplifications of some theorems $\mathbf{T}$ and $\mathbf{t}=\mathbf{u}$ of Φ_3. Let $\overline{\mathbf{U}}$ be the result of replacement of a particular occurrence of $\overline{\mathbf{t}}$ in $\overline{\mathbf{T}}$ by $\overline{\mathbf{u}}$. Replace that occurrence of $\overline{\mathbf{t}}$ in $\overline{\mathbf{T}}$ by a variable v not occurring in $\overline{\mathbf{T}}$. Thus, we get an equation $\overline{\mathbf{V}}$ between two words which we prove to be a formula of Φ_3. Indeed, one of the sides of $\overline{\mathbf{V}}$ is identical with the corresponding side of $\overline{\mathbf{T}}$ whereas the other side differs by replacement of an occurrence of $\overline{\mathbf{t}}$ by v. Now, v has the valency -1 just as $\overline{\mathbf{t}}$; hence, the side in question of $\overline{\mathbf{V}}$ has the same valency as the corresponding side of $\overline{\mathbf{T}}$ and its proper sections have the same valencies as some proper sections of that. Thus, by lemmas 3 and 5, both sides of $\overline{\mathbf{V}}$ are terms of Φ_3, that is, $\overline{\mathbf{V}}$ is a formula of Φ_3, hence, the simplification of a formula $\mathbf{V}$ of Φ_2. Now, substituting $\mathbf{t}$ for v in $\mathbf{V}$, we get a formula of Φ_2 which has obviously $\overline{\mathbf{T}}$ as its simplification. By the above corollary of lemma 4, this formula is identical with the theorem $\mathbf{T}$ of Φ_2. Substituting, on the other hand, $\mathbf{u}$ for v in $\mathbf{V}$, we get another formula $\mathbf{U}$ of Φ_2 which has $\overline{\mathbf{U}}$ as its simplification. Now, **U** is plainly the result of replacement of an occurrence of **t** in **T** by **u**, hence, a theorem of Φ_2; consequently, lemma 6 holds for $\bar{\mathbf{U}}$ too.

As a corollary, we see that $UF^k0^2=0$ is a theorem of Φ_3 if and only if the formula $U(F^k(0),0)=0$ of Φ_2 having it as simplification is a theorem of Φ_2, that is, if we have $R(k,n)=0$ for some non-negative integer n. Hence, there is no algorithm by means of which, given any non-negative integer k, we could decide if $UF^k0^2=0$ is a theorem of Φ_3.

6. Remark that no theorem of Φ_3, except those of the form $\mathbf{t}=\mathbf{t}$, contains any variable which does not figure in at least one equation of E′. Hence, the class of theorems of Φ_3 does not change essentially if we modify system Φ_3 by allowing no variables other than those figuring in the equations of E′ (besides 0, $=$, and the functors $F,\ldots,R,U,V$) as symbols.[29]

By such a modification, the terms of Φ_3 become particular words formed of the letters of a *finite* alphabet, subject to the valency conditions of lemma 3 or 5. Our next step consists in removing those conditions, and, at the same time, in making the rules of inference more similar to those of Φ_S. For this purpose, we define a formal system Φ_4 as follows. *Symbols* of Φ_4 are 0, the variables $x,y,\ldots,w$ figuring in at least one equation of E′, the functors $F,\ldots,R,U,V,$ and the equality sign $=$. *Terms* of Φ_4 are arbitrary finite (possibly empty) sequences of symbols of Φ_4 except $=$, i. e. words formed of the letters of the alphabet $\{0,x,y,\ldots,w,F,\ldots,R,U,V\}$. *Formulae* of Φ_4 are the sequences of symbols of the form $\mathbf{t}=\mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_4. *Axioms* of Φ_4 are the simplifications of the equations belonging to E′, as well as the formula $e=e$ (e denoting the empty word). *Theorems* of Φ_4 are (i) the axioms of Φ_4; (ii) the result of substitution of 0 or of Fx for a variable x throughout a theorem **T** of Φ_4; (iii) $\alpha\mathbf{t}=\alpha\mathbf{u}$ and $\mathbf{t}\alpha=\mathbf{u}\alpha$ if $\mathbf{t}=\mathbf{u}$ is a theorem of Φ_4 and α a symbol of Φ_4 except $=$; (iv) $\mathbf{u}=\mathbf{v}$ if $\mathbf{t}=\mathbf{u}$ and $\mathbf{t}=\mathbf{v}$ are theorems of Φ_4; (v) nothing else.

As an easy consequence of (i), (iii) and (iv), we have the following

LEMMA 7. *For any term* **t** *of* Φ_4, $\mathbf{t}=\mathbf{t}$ *is a theorem of* Φ_4. *If* $\mathbf{t}=\mathbf{u}$ *is a theorem of* Φ_4 *then the same holds for* $\mathbf{u}=\mathbf{t}$, *as well as, for any term* **v** *of* Φ_4, *for* $\mathbf{vt}=\mathbf{vu}$ *and* $\mathbf{tv}=\mathbf{uv}$. *If* $\mathbf{t}=\mathbf{u}$ *and* **T** *are theorems of* Φ_4, *then the same holds for the result of replacement of a particular occurrence of* **t** *in* **T** by **u**.

Indeed, if $\mathbf{t}=\mathbf{u}$ is a theorem and **v** a term of Φ_4, then a repeated application of (iii) gives $\mathbf{vt}=\mathbf{vu}$ and $\mathbf{tv}=\mathbf{uv}$ as theorems of Φ_4. In particular, $\mathbf{t}=\mathbf{t}$ is a theorem of Φ_4 for each term **t** of Φ_4, for $e=e$ is a theorem of Φ_4. Further, we get for any theorem $\mathbf{t}=\mathbf{u}$ of Φ_4, by application of (iv) to the theorems $\mathbf{t}=\mathbf{u}$ and $\mathbf{t}=\mathbf{t}$ of Φ_4 (with **t** for **v**), $\mathbf{u}=\mathbf{t}$ as a theorem

[29] A similar remark applies for Φ_2 and (with E instead of E′) for Φ_1 too.

of Φ_4. Finally, suppose that $\mathbf{t}=\mathbf{u}$ and $\mathbf{T}$ are theorems of Φ_4 and that there is an occurrence of $\mathbf{t}$ in $\mathbf{T}$. Then T has one of the forms $\mathbf{t}_1\mathbf{t}\mathbf{t}_2=\mathbf{t}_3$ and $\mathbf{t}_3=\mathbf{t}_1\mathbf{t}\mathbf{t}_2$ with terms $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3$ of Φ_4. By what has been proved already, $\mathbf{t}_1\mathbf{t}\mathbf{t}_2=\mathbf{t}_3$ is in any case a theorem of Φ_4 and the same holds for $\mathbf{t}_1\mathbf{t}\mathbf{t}_2=\mathbf{t}_1\mathbf{u}\mathbf{t}_2$ too. Hence, by (iv), $\mathbf{t}_1\mathbf{u}\mathbf{t}_2=\mathbf{t}_3$ and $\mathbf{t}_3=\mathbf{t}_1\mathbf{u}\mathbf{t}_2$ are theorems of Φ_4; and, choosing $\mathbf{t}_1, \mathbf{t}_2$ and $\mathbf{t}_3$ appropriately, one of them is the result of replacement of the particular occurrence in question of $\mathbf{t}$ in $\mathbf{T}$ by $\mathbf{u}$.

As a corollary of lemma 7, we see that each theorem of Φ_3 is a theorem of Φ_4. Indeed, the axioms of Φ_3 are theorems of Φ_4 (for they are either axioms of Φ_4 or of the form $\mathbf{t}=\mathbf{t}$ with a term $\mathbf{t}$ of Φ_4); and the rules of inference of Φ_3 hold in Φ_4 too.

Now, we shall show that, in spite of allowing "meaningless" words (i. e. those which are no terms of Φ_3) as terms of Φ_4 and of admitting the new rules of inference, the class of the theorems did not change "essentially". For this purpose, we shall analyse the structure of the terms of Φ_4 as to their "meaningful" components.

A sequence of consecutive symbols of a term of Φ_4 is called a *sub-term* of it. A particular occurrence of a functor G in a term $\mathbf{t}$ of Φ_4 is called *saturated* (or G is called saturated at that occurrence), if there is a sub-term $\mathbf{u}$ of $\mathbf{t}$ beginning with that occurrence of G which is a term of Φ_3. In this case, $\mathbf{u}$ is uniquely determined, for of two different sub-terms of $\mathbf{t}$ beginning with the same occurrence of G, one is a proper section of the other, and, by lemma 3, a proper section of a term of Φ_3 cannot be a term of Φ_3. If a particular occurrence of G in $\mathbf{t}$ is not saturated, it is called *unsaturated* (or G is called unsaturated at that occurrence).

Now, we can describe the structure of the terms of Φ_4 by

LEMMA 8. *Each term* $\mathbf{t}$ *of* Φ_4 *can be decomposed in one and only one way into a product* (i. e., juxtaposition) *of "components" each of which is either a term of* Φ_3 *or an unsaturated occurrence of a functor.* (We call this decomposition the "standard decomposition" of $\mathbf{t}$.)

Indeed, this holds for a term of Φ_4 containing a single symbol (for it is either a term of Φ_3, viz. if it is 0 or one of the variables, or a functor in which case it is unsaturated). Suppose, the statement of the lemma holds for terms of Φ_4 containing less than r symbols ($r=2, 3, \ldots$) and let $\mathbf{t}$ be a term of Φ_4 containing r symbols. If the first symbol α of $\mathbf{t}$ is 0 or a variable, or an unsaturated functor, then $\mathbf{t}$ is $\alpha\mathbf{u}$ where $\mathbf{u}$ is a term of Φ_4 containing $r-1$ symbols; and $\mathbf{t}$ can be decomposed, as required, into α and the components of $\mathbf{u}$. This is the only such decomposition of $\mathbf{t}$ for there is no term of Φ_3 beginning with α and the standard decomposition of $\mathbf{u}$ is unique. If the first symbol of $\mathbf{t}$ is a saturated functor G, i. e. $\mathbf{t}$ is $\mathbf{u}\mathbf{v}$ where $\mathbf{u}$ is a term of Φ_3 beginning with G and $\mathbf{v}$ a term of Φ_4, containing less than r symbols, then

$\mathbf{t}$ can be decomposed as required into $\mathbf{u}$ and the components of $\mathbf{v}$. This is the only such decomposition, for by lemma 3, neither a proper section of $\mathbf{u}$ nor a term of Φ_4 having $\mathbf{u}$ as its proper section can be a term of Φ_3, and the standard decomposition of $\mathbf{v}$ is unique.

By means of lemma 8, we can prove the following necessary (and as easily seen, but not used in the sequel, sufficient) condition of being a theorem of Φ_4.

LEMMA 9. *If* $\mathbf{t} = \mathbf{u}$ *is a theorem of* Φ_4, *then* $\mathbf{t}$ *and* $\mathbf{u}$ *have the same number* r *of components. If* $\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_r$ *and* $\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_r$ *are the standard decompositions of* $\mathbf{t}$ *and* $\mathbf{u}$, *respectively, then, for* $i = 1, 2, \ldots, r$, $\mathbf{t}_i$ *and* $\mathbf{u}_i$ *are either both occurrences of the same unsaturated functor, or both are terms of* Φ_3 *and* $\mathbf{t}_i = \mathbf{u}_i$ *is a theorem of* Φ_3.

Indeed, this holds for the axioms of Φ_4, for, in case of the axiom $e = e$, e has no components, and, in case of the other axioms, each side of them has a single component and they are axioms of Φ_3. Supposing, the statement of the lemma holds for a theorem $\mathbf{t} = \mathbf{u}$ of Φ_4, let $\mathbf{t}'$ and $\mathbf{u}'$ denote the terms formed of $\mathbf{t}$ and $\mathbf{u}$, respectively, by substituting throughout 0 or Fx (the same in both cases) for a variable x. Let $\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_r$ and $\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_r$ be the standard decompositions of $\mathbf{t}$ and $\mathbf{u}$, respectively; and, for $i = 1, 2, \ldots, r$, let $\mathbf{t}_i'$ and $\mathbf{u}_i'$ denote the terms of Φ_4 formed of $\mathbf{t}_i$ and $\mathbf{u}_i$, respectively, by substituting throughout 0 or Fx, respectively, for x. Then, each of $\mathbf{t}_1', \mathbf{t}_2', \ldots, \mathbf{t}_r', \mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_r'$ is either an unsaturated functor, or a term of Φ_3. Indeed, a functor does not change and a term of Φ_3 does not cease being a term of Φ_3 by substitution of 0 or Fx for x. Also, an unsaturated functor cannot become saturated by this substitution. For, if an occurrence of the functor G, viz. one of $\mathbf{t}_1', \mathbf{t}_2', \ldots, \mathbf{t}_r'$ or of $\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_r'$, would be saturated in $\mathbf{t}'$ or $\mathbf{u}'$ respectively, then there would be a sub-term of $\mathbf{t}'$ or $\mathbf{u}'$, respectively, beginning with the said occurrence of G which is a term of Φ_3. However, we could regain a sub-term of $\mathbf{t}$ or $\mathbf{u}$, respectively, beginning with the corresponding occurrence of G, thus, with one of $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$, or of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, respectively, by means of replacing some of the 0's or Fx's by x according as 0 or Fx has been substituted for x. Now, as remarked above (see footnote [14]), we get by this replacement a term of Φ_3 again which is impossible, for the said occurrence of G in $\mathbf{t}$ or $\mathbf{u}$, respectively, is an unsaturated one. Hence $\mathbf{t}_1'\mathbf{t}_2'\ldots\mathbf{t}_r'$ and $\mathbf{u}_1'\mathbf{u}_2'\ldots\mathbf{u}_r'$ are the standard decompositions of $\mathbf{t}'$ and $\mathbf{u}'$, respectively. Now, if $\mathbf{t}_i$ and $\mathbf{u}_i$ are occurrences of some unsaturated functor, then the same holds for $\mathbf{t}_i'$ and $\mathbf{u}_i'$ too; if $\mathbf{t}_i$ and $\mathbf{u}_i$ are terms of Φ_3 such that $\mathbf{t}_i = \mathbf{u}_i$ is a theorem of Φ_3, then the same holds for $\mathbf{t}_i' = \mathbf{u}_i'$ too. Hence, the statement of the lemma holds for the theorem $\mathbf{t}' = \mathbf{u}'$ of Φ_4 too.

We have next to prove that if the statement of the lemma holds for the theorem $\mathbf{t} = \mathbf{u}$ of Φ_4, then, for an arbitrary symbol α of Φ_4 except $=$, it holds for the theorems $\alpha\mathbf{t} = \alpha\mathbf{u}$ as well as $\mathbf{t}\alpha = \mathbf{u}\alpha$ of Φ_4 too. Suppose again that $\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$ and $\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_r$ are the standard decompositions of $\mathbf{t}$ and $\mathbf{u}$, respectively. First we treat $\alpha\mathbf{t} = \alpha\mathbf{u}$. If α is 0 or a variable, or else a functor unsaturated in both $\alpha\mathbf{t}$ and $\alpha\mathbf{u}$, then our statement is trivial for then α is the first component of both $\alpha\mathbf{t}$ and $\alpha\mathbf{u}$ and their further components are $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_r$, and $\mathbf{u}_2, \mathbf{u}_2, \dots, \mathbf{u}_r$, respectively. Suppose, α is a functor G and let s be the number of arguments of G. In $\alpha\mathbf{t}$, that is $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$, the first G is saturated if and only if $r \geqq s$ and $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s$ are terms of Φ_3. Indeed, in this case, the sub-term $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s$ of $\alpha\mathbf{t}$ is a term of Φ_3; conversely, if the first G is saturated in $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$, i. e., if there is a sub-term of $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$ beginning with the first G, thus of the form $G\mathbf{t}'_1\mathbf{t}'_s\dots\mathbf{t}'_s$ with some terms $\mathbf{t}'_1, \mathbf{t}'_2, \dots, \mathbf{t}'_s$ of Φ_2, then $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$ is $G\mathbf{t}'_1\mathbf{t}'_2\dots\mathbf{t}'_s\mathbf{v}$ with some term $\mathbf{v}$ of Φ_4, thus $\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_r$ is $\mathbf{t}'_1\mathbf{t}'_2\dots\mathbf{t}'_s\mathbf{v}$, hence we have $r \geqq s$ and $\mathbf{t}_1$ is $\mathbf{t}'_1$, $\mathbf{t}_2$ is $\mathbf{t}'_2, \dots, \mathbf{t}_s$ is $\mathbf{t}'_s$ for the (unique) standard decomposition of $\mathbf{t}'_1\mathbf{t}'_2\dots\mathbf{t}'_s\mathbf{v}$ is $\mathbf{t}'_1\mathbf{t}'_2\dots\mathbf{t}'_s$ followed by the standard decomposition of $\mathbf{v}$. Taking into account that, for $i = 1, 2, \dots, r$, $\mathbf{t}_i$ and $\mathbf{u}_i$ are either both terms of Φ_3 or neither, we see that the first G is either saturated in both $\alpha\mathbf{t}$ and $\alpha\mathbf{u}$, or unsaturated in both. The second case being settled already, let us suppose that the first G is saturated in both $\alpha\mathbf{t}$ and $\alpha\mathbf{u}$, i. e. we have $r \geqq s$ and $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s$ are terms of Φ_3. Then, the components of $\mathbf{t}$ and $\mathbf{u}$ are $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s$, $\mathbf{t}_{s+1}, \mathbf{t}_{s+2}, \dots, \mathbf{t}_r$, and $G\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_s, \mathbf{u}_{s+1}, \mathbf{u}_{s+2}, \dots, \mathbf{u}_r$, respectively. We have only to show that $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s = G\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_s$ is a theorem of Φ_3, for if $i = s+1, s+2, \dots, r$, then either $\mathbf{t}_i$ and $\mathbf{u}_i$ are the same unsaturated functor or $\mathbf{t}_i = \mathbf{u}_i$ is a theorem of Φ_3. Now, $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s = G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s$ and $\mathbf{t}_1 = \mathbf{u}_1$, $\mathbf{t}_2 = \mathbf{u}_2, \dots, \mathbf{t}_s = \mathbf{u}_s$ being theorems of Φ_3, the same holds for $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s = G\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_s$ too, for it can be obtained from $G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s = G\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_s$ by replacement of the occurrences of $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s$ on the right-hand side by $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s$, respectively.

Now, let us examine $\mathbf{t}\alpha = \mathbf{u}\alpha$ where $\mathbf{t} = \mathbf{u}$ is a theorem of Φ_3 and α an arbitrary symbol of Φ_4 except $=$. If α is a functor, then it is obviously unsaturated, thus, the statement of the lemma is trivial again, the components of $\mathbf{t}\alpha$ and $\mathbf{u}\alpha$ being those of $\mathbf{t}$ and $\mathbf{u}$, respectively, and α. The same holds if α is 0 or a variable, provided no functor which is unsaturated in $\mathbf{t}$ or $\mathbf{u}$ becomes saturated by subjoining α. Suppose, some components of $\mathbf{t}$, say, which are unsaturated functors, are saturated in $\mathbf{t}\alpha$ and let $\mathbf{t}_i$ be the first of those components (from the left to the right). Then $\mathbf{t}_i$ is a functor G and $G\mathbf{t}_{i+1}\mathbf{t}_{i+2}\dots\mathbf{t}_r\alpha$ is a term of Φ_3 (for this is the only sub-term of $\mathbf{t}\alpha$ beginning with $\mathbf{t}_i$ which is not a sub-term of $\mathbf{t}$). Replacing the term $\mathbf{t}_j$ of Φ_3 by the term $\mathbf{u}_j$ of Φ_3 $(j = i+1, i+2, \dots, r)$, we get another term $G\mathbf{u}_{i+1}\mathbf{u}_{i+2}\dots\mathbf{u}_r\alpha$ of Φ_3 (see footnote [11]), which is a sub-term of $\mathbf{u}\alpha$ beginning with $\mathbf{t}_i$, i. e. $\mathbf{u}_i$; hence, $\mathbf{u}_i$ is saturated in $\mathbf{u}\alpha$. Also, $\mathbf{u}_i$ is the first component of $\mathbf{u}$ which is an unsaturated functor in $\mathbf{u}$ but saturated in $\mathbf{u}\alpha$, for a similar argument shows that if $\mathbf{u}_{i'}$ were another such component of $\mathbf{u}$ $(i' < i)$, then $\mathbf{t}_{i'}$ would be a component of $\mathbf{t}$ which is an unsaturated functor in $\mathbf{t}$ but saturated in $\mathbf{t}\alpha$, contrary to the fact that $\mathbf{t}_i$ is the first such component of $\mathbf{t}$. Hence, the components of $\mathbf{t}\alpha$ are $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_{i-1}$, and $G\mathbf{t}_{i+1}\mathbf{t}_{i+2}\ldots\mathbf{t}_r\alpha$, and those of $\mathbf{u}\alpha$ are $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}$, and $G\mathbf{u}_{i+1}\mathbf{u}_{i+2}\ldots\mathbf{u}_r\alpha$. Now, $G\mathbf{t}_{i+1}\mathbf{t}_{i+2}\ldots\mathbf{t}_r\alpha = G\mathbf{u}_{i+1}\mathbf{u}_{i+2}\ldots\mathbf{u}_r\alpha$ is a theorem of Φ_3, for it can be obtained from the theorem $G\mathbf{t}_{i+1}\mathbf{t}_{i+2}\ldots\mathbf{t}_r\alpha = G\mathbf{t}_{i+1}\mathbf{t}_{i+2}\ldots\mathbf{t}_r\alpha$ of Φ_3 by replacement of some of the $\mathbf{t}_j$, $j = i+1, i+2, \ldots, r$, viz. those which are terms of Φ_3, by the corresponding $\mathbf{u}_j$ and, for these j, $\mathbf{t}_j = \mathbf{u}_j$ is a theorem of Φ_3. Hence, the statement of the lemma holds for $\mathbf{t}\alpha = \mathbf{u}\alpha$ too.

We have still to prove that if the statement of the lemma holds for the theorems $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$ of Φ_4, then it holds for $\mathbf{u} = \mathbf{v}$ too. Let $\mathbf{t}_1\mathbf{t}_2\ldots\mathbf{t}_p$, $\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_q$ and $\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r$ be the standard decompositions of $\mathbf{t}, \mathbf{u}$ and $\mathbf{v}$, respectively. By the hypothesis, we have $p = q$ and $p = r$, thus $q = r$. Also, if, for some $i = 1, 2, \ldots, r$, $\mathbf{t}_i$ is an unsaturated functor in $\mathbf{t}$, then $\mathbf{u}_i$ and $\mathbf{v}_t$ are unsaturated functors in $\mathbf{u}$ and $\mathbf{v}$, respectively, and $\mathbf{u}_i$ is identical with $\mathbf{v}_i$, for both are identical with $\mathbf{t}_i$. If $\mathbf{t}_i$ is a term of Φ_3, the same holds for $\mathbf{u}_i$ and $\mathbf{v}_i$ too, and $\mathbf{t}_i = \mathbf{u}_i, \mathbf{t}_i = \mathbf{v}_i$ are theorems of Φ_3; hence, the same holds for $\mathbf{u}_i = \mathbf{v}_i$ too, for it can be obtained from $\mathbf{t}_i = \mathbf{v}_i$ by replacement ot the occurrence of $\mathbf{t}_i$ on the left-hand side by $\mathbf{u}_i$. Hence, the statement of the lemma holds for the theorem $\mathbf{u} = \mathbf{v}$ of Φ_4 too, which completes the proof of the lemma.

As a corollary, we see that if $\mathbf{t} = \mathbf{u}$ is a formula of Φ_3 and at the same time a theorem of Φ_4, then it is a theorem of Φ_3. Indeed, in this case, $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_3, thus, the only components of themselves.

In particular, by lemmas 7 and 9, we see that $UF^k0^2 = 0$ is a theorem of Φ_4 if and only if it is a theorem of Φ_3, that is, if we have $R(k, n) = 0$ for some non-negative integer n. Hence, there is no algorithm by means of which, given any non-negative integer k, we could decide if $UF^k0^2 = 0$ is a theorem of Φ_4.

7. The formal system Φ_4 is no particular case of Φ_S because of the rule of inference (ii), i. e. of the rule of substitution which is alien from Φ_S. However, we show that it is possible to dispense with this rule by introducing some new symbols and axioms. For this purpose, let us define a formal system Φ_5 as follows. *Symbols* of Φ_5 are those of Φ_4, i. e. 0, the variables $x, y, \ldots, w$, the functors $F, \ldots, R, U, V$, the equality sign $=$; and, in addition, to each of the variables $x, y, \ldots, w$, two new[30] symbols, $x_0, x_+, y_0, y_+, \ldots, w_0, w_+$, called *substitutors*, and a single new symbol a, called *absorptor*. *Terms* of Φ_5 are arbitrary finite (possibly empty) sequences of symbols of Φ_5 except $=$,

[30] I. e., we suppose that $x_0, x_+, y_0, y_+, \ldots, w_0, w_+$ are different from the symbols $0, x, y, \ldots, w, F, \ldots, R, U, V$ as well as from each other.

i. e. words formed of the letters of the alphabet $\{0, x, x_0, x_+, y, y_0, y_+, \ldots, \ldots, w, w_0, w_+, F, \ldots, R, U, V\}$. *Formulae* of Φ_5 are sequences of symbols of the form $\mathbf{t} = \mathbf{u}$ where $\mathbf{t}$ and $\mathbf{u}$ are terms of Φ_5; i. e., relations between such words. *Axioms* of Φ_5 are those of Φ_4 (i. e., the simplifications of the equations belonging to E' as well as $e = e$, e denoting the empty word); and, in addition, for each variable[31] x the relations

$$x_0\, x = 0 x_0\,, \tag{14}$$
$$x_+ x = F x x_+\,, \tag{15}$$
$$x_0\, a = e, \tag{16}$$
$$x_+ a = e, \tag{17}$$
$$x_0\, 0 = 0 x_0\,, \tag{18}$$
$$x_+ 0 = 0 x_+\,, \tag{19}$$

and, for each variable[32] y different from x and for each functor G,

$$x_0\;\, y = y x_0\,, \tag{20}$$
$$x_+\, y = y x_+\,, \tag{21}$$
$$x_0\, G = G x_0, \tag{22}$$
$$x_+ G = G x_+\,. \tag{23}$$

Theorems of Φ_5 are (i) the axioms of Φ_5; (ii) $\alpha\mathbf{t} = \alpha\mathbf{u}$ and $\mathbf{t}\alpha = \mathbf{u}\alpha$ if $\mathbf{t} = \mathbf{u}$ is any theorem of Φ_5 and α any symbol of Φ_5 except $=$; (iii) $\mathbf{u} = \mathbf{v}$ if $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$ are theorems of Φ_5; (iv) nothing else.

First we prove that the new symbols $a, x_0, x_+, y_0, y_+, \ldots, w_0, w_+$ together with the new axioms (14) to (23) suffice to replace the missing rule of substitution; i. e., we prove

LEMMA 10. *Each theorem of Φ_4 is a theorem of Φ_5.*

To prove this lemma we observe that lemma 7 holds for Φ_5 instead of Φ_4 too, for the rule of substitution, missing in Φ_5, has not been used in its proof.

Now, the statement of lemma 10 holds for the axioms of Φ_4 for they are axioms of Φ_5. Suppose, it holds for a theorem $\mathbf{T}$ of Φ_4; then it holds for the theorems $\mathbf{U}$ and $\mathbf{V}$ of Φ_4 obtained by substitution of 0 and Fx, respectively, for a variable x throughout $\mathbf{T}$. Indeed, suppose $\mathbf{T}$ is $\alpha_1 \alpha_2 \cdots \alpha_r = \beta_1 \beta_2 \cdots \beta_s$, where $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s$ are symbols of Φ_5, except $=$. Then $\mathbf{U}$ is $\alpha_1' \alpha_2' \cdots \alpha_r' = \beta_1' \beta_2' \cdots \beta_s'$ and $\mathbf{V}$ is $\alpha_1'' \alpha_2'' \cdots \alpha_r'' = \beta_1'' \beta_2'' \cdots \beta_s''$, α_i' denoting 0 or α_i and α_i'' denoting Fx or α_i according as α_i is x or not $(i = 1, 2, \ldots, r)$, and β_i' denoting 0 or β_i, β_i'' denoting Fx or β_i according

[31] "For *each* variable x" has been used here in an analogous sense as "for a variable x" (see footnote [14]).

[32] I. e., y can be replaced by any variable which is different from the variable by which x has been replaced.

as β_i is x or not ($i = 1, 2, \ldots, s$). Now, by (ii),

(24) $$x_0 \alpha_1 \alpha_2 \cdots \alpha_r a = x_0 \beta_1 \beta_2 \cdots \beta_s a$$

and

(25) $$x_+ \alpha_1 \alpha_2 \cdots \alpha_r a = x_+ \beta_1 \beta_2 \cdots \beta_s a$$

are theorems of Φ_5. By (14), (18), (20), and (22), the same holds for $x_0 \alpha_i = \alpha_i' x_0$ ($i = 1, 2, \ldots, r$), $x_0 \beta_i = \beta_i' x_0$ ($i = 1, 2, \ldots, s$), and by (15), (19), (21), and (23), for $x_+ \alpha_i = \alpha_i'' x_+$ ($i = 1, 2, \ldots, r$), $x_+ \beta_i = \beta_i'' x_+$ ($i = 1, 2, \ldots, s$) too. Hence, by lemma 7, last assertion, we can replace $x_0 \alpha_1$ by $\alpha_1' x_0$, then $x_0 \alpha_2$ by $\alpha_2' x_0, \ldots,$ finally $x_0 \alpha_r$ by $\alpha_r' x_0$ on the left-hand side of (24), and $x_0 \beta_1$ by $\beta_1' x_0$, then $x_0 \beta_2$ by $\beta_2' x_0, \ldots,$ finally $x_0 \beta_s$ by $\beta_s x_0$ on the right-hand side of (24), and, similarly, $x_+ \alpha_1$ by $\alpha_1'' x_+$, then $x_+ \alpha_2$ by $\alpha_2'' x_+, \ldots,$ finally $x_+ \alpha_r$ by $\alpha_r'' x_+$ on the left-hand side of (25), and $x_+ \beta_1$ by $\beta_1'' x_+$, then $x_+ \beta_2$ by $\beta_2'' x_+, \ldots,$ finally $x_+ \beta_s$ by $\beta_s'' x_+$ on the right-hand side of (25). Thus, we see that

$$\alpha_1' x_2' \cdots \alpha_r' x_0 a = \beta_1' \beta_2' \cdots \beta_s' x_0 a$$

and

$$\alpha_1'' \alpha_2'' \cdots \alpha_r'' x_+ a = \beta_1'' \beta_2'' \cdots \beta_s'' x_+ a$$

are theorems of Φ_5. Here, by (16) and (17), we can replace $x_0 a$ and $x_+ a$ by e, i. e. we can omit them, on both sides; thus, we get $\alpha_1' \alpha_2' \cdots \alpha_r' = \beta_1' \beta_2' \cdots \beta_s'$ and $\alpha_1'' \alpha_2'' \cdots \alpha_r'' = \beta_1'' \beta_2'' \cdots \beta_s''$, i. e. **U** and **V** as theorems of Φ_5.

The rest of the proof of lemma 10 is quite trivial for the rules of inference (iii) and (iv) of Φ_4 are those (viz. (ii) and (iii)) of Φ_5 too.

Now, we shall examine the relation of the theorems of Φ_5 to those of Φ_4 in order to show that a theorem of Φ_5 which is a formula of Φ_4 is a theorem of Φ_4.

We call a sequence of consecutive symbols of a term of Φ_5 a *sub-term* of it. We call a term of Φ_5 *reduced* if it does not contain any occurrence of a substitutor preceding an occurrence of a symbol which is no substitutor. To each term **t** of Φ_5, we attach a reduced term $\bar{\mathbf{t}}$, called the *reductum* of **t**, by the following *reduction process.* If **t** is reduced, then $\bar{\mathbf{t}}$ is **t**. If **t** is not reduced, let α be the last occurrence of a substitutor in **t** preceding an occurrence of a symbol which is no substitutor. If there is an occurrence of the absorptor a preceded by α, then perform as first "*reduction step*" the substitution indicated by α (i. e. of 0 for x if α is x_0, and of Fx for x if α is x_+, say) throughout the sub-term of **t** between α and the next occurrence of a after α, and cancel both α and this occurrence of a. If there is no occurrence of a preceded by α, then perform as first reduction step the substitution indicated by α throughout the sub-term of **t** between α and the next occurrence of a substitutor after α, or between α and the end of **t**, if such occurrence does not exist, and transplant α to the place immediately after this sub-term. By means of this reduction step, we obtain another term **u** of Φ_5; and the reductum of $\mathbf{t}$ is, by definition, the reductum of $\mathbf{u}$. This definition has a sense, for in $\mathbf{u}$, the number of the occurrences of substitutors, preceding an occurrence of a symbol which is not a substitutor, is one less than in $\mathbf{t}$ so that the iterated application of the said reduction step comes to an end in a finite number of steps.

A reduced term $\mathbf{t}$ has obviously the form $\mathbf{t}_0 a \mathbf{t}_1 a \mathbf{t}_2 \ldots a \mathbf{t}_r \mathbf{v}$ where $\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$ are terms of Φ_4 (possibly the empty word e) and $\mathbf{v}$ is a finite (possibly empty) sequence of substitutors. We call $\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$ the *intervals* and $\mathbf{v}$ the *suffix* of $\mathbf{t}$. For a non-reduced term $\mathbf{t}$, we call the intervals and the suffix of its reductum also the intervals and the suffix, respectively, of $\mathbf{t}$.

Now, the relation between the theorems of Φ_5 and those of Φ_4 is displayed in the following lemma which gives a necessary (and, as easily seen, but not used in the sequel, sufficient) condition of being a theorem of Φ_5.

LEMMA 11. *If* $\mathbf{t} = \mathbf{u}$ *is a theorem of* Φ_5, *then* $\mathbf{t}$ *and* $\mathbf{u}$ *have the same number* r *of intervals. If* $\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$ *and* $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ *are in succession the intervals of* $\mathbf{t}$ *and* $\mathbf{u}$, *respectively, then for* $i = 0, 1, 2, \ldots, r$, $\mathbf{t}_i = \mathbf{u}_i$ *is a theorem of* Φ_4. *Also, the suffixes of* $\mathbf{t}$ *and* $\mathbf{u}$ *are identical.*

Indeed, the statement of the lemma holds for the axioms of Φ_5, for a term of Φ_4 is its only interval and has an empty suffix, and, for any variable x, $x_0 x, x_+ x, x_0 a, x_+ a, x_0 \alpha, x_+ \alpha$ (for any symbol α of Φ_5 except x, a, and $=$) have the reducta $0 x_0, F x x_+, e, e, \alpha x_0, \alpha x_+$, respectively. Suppose the statement of the lemma holds for a theorem $\mathbf{t} = \mathbf{u}$ of Φ_5. Then, for any symbol α of Φ_5 except $=$, it holds for $\alpha \mathbf{t} = \alpha \mathbf{u}$ too. Indeed, if α is no substitutor, then the result of each reduction step, performed on $\alpha \mathbf{t}$, is the same as that of the corresponding reduction step, performed on $\mathbf{t}$, except that α is prefixed; hence, if α is not the absorptor a either, the first interval of $\alpha \mathbf{t}$ is $\alpha \mathbf{t}_0$, the others are, in succession, $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$, ($\mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$ denoting the intervals of $\mathbf{t}$) and the suffix of $\alpha \mathbf{t}$ is the same as that of $\mathbf{t}$. Analogously, the intervals of $\alpha \mathbf{u}$ are, in succession, $\alpha \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ denoting the intervals of $\mathbf{u}$, and the suffix of $\alpha \mathbf{u}$ is the same as that of $\mathbf{u}$. Now, $\mathbf{t}_1 = \mathbf{u}_1, \mathbf{t}_2 = \mathbf{u}_2, \ldots, \mathbf{t}_r = \mathbf{u}_r$ are, by the hypothesis, theorems of Φ_4; and the same holds for $\alpha \mathbf{t}_0 = \alpha \mathbf{u}_0$ too, for $\mathbf{t}_0 = \mathbf{u}_0$ is, by the hypothesis, a theorem of Φ_4. Also, the suffixes of $\alpha \mathbf{t}$ and $\alpha \mathbf{u}$ are, by the hypothesis, the same; hence, the statement of the lemma holds for the theorem $\alpha \mathbf{t} = \alpha \mathbf{u}$ of Φ_5 too. If α is the absorptor a, then the number of the intervals of $\alpha \mathbf{t}$ and $\alpha \mathbf{u}$ is one more than that of the intervals of $\mathbf{t}$ and $\mathbf{u}$, respectively, the first interval of both $\alpha \mathbf{t}$ and $\alpha \mathbf{u}$ being e, and the others, in succession, the same as those of $\mathbf{t}$ and $\mathbf{u}$, respectively; also, the suffixes of $\alpha \mathbf{t}$ and $\alpha \mathbf{u}$ are the same as those of $\mathbf{t}$ and $\mathbf{u}$, respectively. Hence, the statement of the lemma holds for the theorem $\alpha \mathbf{t} = \alpha \mathbf{u}$ of Φ_5 too. If α is a substitutor, x_0 or x_+, say, then the result of each but the last reduction step, performed on $\alpha\mathbf{t}$, differs but in a prefixed α from that of the corresponding reduction step, performed on $\mathbf{t}$. As to the last reduction step, its effect is substitution of 0 or Fx, respectively, for x in the first interval of $\mathbf{t}$ and cancellation of the first occurrence of a in the reductum of $\mathbf{t}$, if $\mathbf{t}$ has more than one interval; and substitution of 0 or Fx, respectively, for x in the single interval of $\mathbf{t}$ and prefixing of α to the suffix of $\mathbf{t}$ if there is but one interval of $\mathbf{t}$. Hence, the intervals of $\alpha\mathbf{t}$ are $\mathbf{t}_0'\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \ldots, \mathbf{t}_r$ and similarly, those of $\alpha\mathbf{u}$ are $\mathbf{u}_0'\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_r$ in the first case, $\mathbf{t}_0'$ and $\mathbf{u}_0'$, respectively, in the second case, $\mathbf{t}_0'$ and $\mathbf{u}_0'$ denoting the result of substitution of 0 or of Fx for x, according as α is x_0 or x_+, in $\mathbf{t}_0$ and $\mathbf{u}_0$, respectively. Now, $\mathbf{t}_0 = \mathbf{u}_0, \mathbf{t}_1 = \mathbf{u}_1, \mathbf{t}_2 = \mathbf{u}_2, \ldots, \mathbf{t}_r = \mathbf{u}_r$ being theorems of Φ_4, the same holds for $\mathbf{t}_0' = \mathbf{u}_0'$, and, by lemma 7, for $\mathbf{t}_0'\mathbf{t}_1 = \mathbf{u}_0'\mathbf{u}_1$ too (for $\mathbf{t}_0'\mathbf{t}_1 = \mathbf{t}_0'\mathbf{t}_1$ is a theorem of Φ_4 and by replacement of $\mathbf{t}_0'$ by $\mathbf{u}_0'$ on the right-hand side we get $\mathbf{t}_0'\mathbf{t}_1 = \mathbf{u}_0'\mathbf{t}_1$, then, by replacement of $\mathbf{t}_1$ by $\mathbf{u}_1$ on the right-hand side, $\mathbf{t}_0'\mathbf{t}_1 = \mathbf{u}_0'\mathbf{u}_1$). Also, the suffixes of $\alpha\mathbf{t}$ and $\alpha\mathbf{u}$ are the same in both cases; hence, the statement of lemma 11 holds for the theorem $\alpha\mathbf{t} = \alpha\mathbf{u}$ of Φ_5 also if α is a substitutor.

Suppose again, the statement of the lemma holds for a theorem $\mathbf{t} = \mathbf{u}$ of Φ_5; we have next to prove that, for any symbol α of Φ_5 except $=$, it holds for the theorem $\mathbf{t}\alpha = \mathbf{u}\alpha$ of Φ_5 too. If α is a substitutor, then the result of each reduction step, performed on $\mathbf{t}\alpha$, differs from that of the corresponding reduction step, performed on $\mathbf{t}$, but in a suffixed α to the suffix of $\mathbf{t}$; hence, the intervals of $\mathbf{t}\alpha$ are the [same as those of $\mathbf{t}$ and the suffix of $\mathbf{t}\alpha$ differs from that of $\mathbf{t}$ but in a suffixed α. An analogous assertion holding for $\mathbf{u}\alpha$, the statement of the lemma holds for the theorem $\mathbf{t}\alpha = \mathbf{u}\alpha$ of Φ_5 too. If α is the absorptor a, then its effect in the result of the successive reduction steps is suffixing a to the result of the corresponding reduction steps, performed on $\mathbf{t}$, so far as no substitutor appears on the end of this result; afterwards its effect is cancelling the last substitutor appearing on the end of this result. Hence, the intervals of $\mathbf{t}a$ are the same as those of $\mathbf{t}$ and the suffix of $\mathbf{t}a$ is the same as that of $\mathbf{t}$ except that the last substitutor has been cancelled, provided the suffix of $\mathbf{t}$ is not empty; and, if the suffix of $\mathbf{t}$ is empty, the intervals of $\mathbf{t}a$ are the same as those of $\mathbf{t}$ and, in addition, the empty word e, whereas, in this case, the suffix of $\mathbf{t}a$ is empty as well. Hence, the statement of the lemma holds for the theorem $\mathbf{t}a = \mathbf{u}a$ of Φ_5 too. Finally, if α is 0, a variable, or a functor, then its effect in the result of thə successive reduction steps is suffixing α to the result of the corresponding reduction steps, performed on $\mathbf{t}$, so far as no substitutor appears on the end of this result; and afterwards, suffixing the result of the substitutions in α, indicated by the substitutors standing on the end of this result, to the last interval of $\mathbf{t}$. Hence, the intervals of $\mathbf{t}\alpha$ are the same as those of $\mathbf{t}$, except that the last one is suffixed by the result α' of substitutions in α, indicated by the suffix of $\mathbf{t}$ (i. e. by α, if α is 0 or a functor; by $F^l 0$ if α is a variable x and there are l occurrences of the substitutor x_+ preceded by the last occurrence of the substitutor x_0 in the suffix of $\mathbf{t}$; and by $F^l x$ if α is a variable x, there are l occurrences of the substitutor x_+ and no occurrence of the substitutor x_0 in the suffix of $\mathbf{t}$); and the suffix of $\mathbf{t}\alpha$ is the same as that of $\mathbf{t}$. An analogous assertion holds for $\mathbf{u}\alpha$; hence, the statement of the lemma holds for the theorem $\mathbf{t}\alpha = \mathbf{u}\alpha$ of Φ_5 in this case too, for, together with $\mathbf{t}_r = \mathbf{u}_r$, $\mathbf{t}_r\alpha' = \mathbf{u}_r\alpha'$ is a theorem of Φ_4.

We have still to prove that if the statement of the lemma holds for the theorems $\mathbf{t} = \mathbf{u}$ and $\mathbf{t} = \mathbf{v}$ of Φ_5, then it holds for $\mathbf{u} = \mathbf{v}$ too. Now, if $\mathbf{t}$ and $\mathbf{u}$ have the same number of intervals, and the same holds for $\mathbf{t}$ and $\mathbf{v}$ too, then also $\mathbf{u}$ and $\mathbf{v}$ have the same number of intervals. Let $r+1$ be the number of these intervals, and let, in succession, $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_r$ be the intervals of $\mathbf{t}$, $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_r$ those of $\mathbf{u}$ and $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_r$ those of $\mathbf{v}$. Then, by the hypothesis, $\mathbf{t}_0 = \mathbf{u}_0, \mathbf{t}_1 = \mathbf{u}_1, \ldots, \mathbf{t}_r = \mathbf{u}_r$ as well as $\mathbf{t}_0 = \mathbf{v}_0, \mathbf{t}_1 = \mathbf{v}_1, \ldots, \mathbf{t}_r = \mathbf{v}_r$ are theorems of Φ_4; hence, $\mathbf{u}_0 = \mathbf{v}_0, \mathbf{u}_1 = \mathbf{v}_1, \ldots, \mathbf{u}_r = \mathbf{v}_r$ too. Also, the suffixes of $\mathbf{t}$ and $\mathbf{u}$ are identical, and those of $\mathbf{t}$ and $\mathbf{v}$ too; hence, $\mathbf{u}$ and $\mathbf{v}$ have the same suffixes, so that the statement of the lemma holds for the theorem $\mathbf{u} = \mathbf{v}$ of Φ_5 too, which concludes the proof of the lemma.

As a corollary, we see that if $\mathbf{t} = \mathbf{u}$ is a theorem of Φ_5 and, at the same time, a formula of Φ_4, then it is a theorem of Φ_4, for then, $\mathbf{t}$ and $\mathbf{u}$ are the only intervals of themselves. Hence, a formula of Φ_4 is a theorem of Φ_5 if and only if it is a theorem of Φ_4. In particular $UF^k 0^2 = 0$ is a theorem of Φ_5 if and only if it is a theorem of Φ_4, i. e. if we have $R(k, n) = 0$ for some non-negative integer n. Hence, there is no algorithm by means of which, given any non-negative integer k, we could decide if $UF^k 0^2 = 0$ is a theorem of Φ_5.

8. Now, the formal system Φ_5 is a particular case of Φ_S. Indeed, as immediately seen, it is the formal system Φ_S belonging to the particular system S of relations formed of the simplifications of the equations E′ by subjoining the relations (14) to (23) (for each variable x figuring in at least one equation of E′ as well as for each such variable y different from x, and for each functor G figuring in at least one equation of E′). Hence, *there is no algorithm by means of which, given any relation in* $\{0, x, x_0, x_+, y, y_0, y_+, \ldots, w, w_0, w_+, F, \ldots, R, U, V\}$, *we could decide if it is a consequence of this particular system* S *of relations; i. e., the word problem for associative systems, relative to this particular system* S *of relations, is unsolvable by any algorithm.* Thus, the Markov—Post theorem has been proved.

9. By a slight modification of the proof method, we can prove some further results of MARKOV concerning the impossibility of some algorithms for associative systems. Indeed, for the word problem for associative systems, relative to the system of relations in $\{0, x, x_0, x_+, y, y_0, y_+, \ldots, w, w_0, w_+, F, \ldots, R\}$, formed of the simplifications of the equations E (instead of E′) by subjoining the relations (14) to (23) for each pair of different variables x and y as well as for each functor G figuring in at least one equation of E (instead of E′), a solving algorithm can be given by means of methods of the proof theory. On the other hand, the question of the existence of a word $\mathbf{w}$ formed of the letters of the above alphabet, for which $RF^k 0\,\mathbf{w} = 0$ is a consequence of this system of relations, can be proved to be equivalent to the question of the existence of a non-negative integer solution of the equation $R(k, y) = 0$ in y. Hence, *for this system of relations, the word problem can be solved by an algorithm, whereas the (left-hand side) "divisibility problem" cannot be solved by any algorithm.* The existence of a system of relations with essentially[33] the same property has been first proved by MARKOV.[34]

On the other hand, subjoining the relation[35]

$$SxP0y = x$$

to the simplifications of the equations of E, with functors S and P not figuring[36] in the equations of E, and introducing new substitutors indicating, for each pair of variables x and y, substitution of y for x, as well as those indicating, for each functor G, the substitution of $G(x, y, \ldots, v)$ for x (where the number of variables $x, y, \ldots, v$ is the same as that of the arguments of G), we get a system of relations such that in the associative system generated by it every element $\mathbf{t}$ is a right-hand divisor of every element $\mathbf{u}$ (indeed, $S\mathbf{u}P0\mathbf{t} = \mathbf{u}$ is a consequence of that system of relations), whereas $RF^k 0$ is a left-hand divisor of 0 if and only if there is a non-negative integer solution of the equation $R(k, y) = 0$ in y. For this system of relations, the word problem can again be solved by an algorithm; hence, *we have a system of relations, for which both the word problem and the one hand divisibility problem can be solved by an algorithm whereas the other hand divisibility problem not.* The existence of such a system of relations has been first proved also by MARKOV.[37]

To a detailed exposition of the ideas sketched in this section, I shall come back in another publication.

(Received 3 April 1952)

[33] In MARKOV, right-hand side divisibility has been considered instead of left-hand side one. However, this makes no difference, for from any associative system we get another by reinterpreting ab as ba and, in the latter, right-hand side divisibility means the same as left-hand side divisibility in the former.

[34] See loc. cit. [2], theorem 2.

[35] Its intuitive meaning is $x + 0 . y = x$ (Sxy and Pxy standing for $x + y$ and xy, respectively).

[36] Or figuring "in the sense" of the (two member) addition and (two factor) multiplication functor.

[37] See loc. cit. [4].

НОВОЕ ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ МАРКОВА—POST'A

Л. КАЛЬМАР (Сегед)

(Резюме)

Автор дает новое доказательство теоремы, доказанной одновременного и независимо друг от друга Марковым и Post'ом, согласно которой существует конечно-образованная ассоциативная система, для которой проблема тождества неразрешима (конечным общерекурсивным) алгоритмом. Новое доказательство не опирается на теорию λ-конверсируемости Church'a как доказательства Маркова и Post'a. Вместо этого доказательство основано на теореме Kleene, согласно которой существует двухпеременная функция $R(x, y)$ для которой нет такого альгоритма, с помощью которого, если дано неотрицательное целое число k, можно узнать, разрешимо-ли уравнение $R(k, y) = 0$ по y в неотрицательных целых числах.

Основные идеи доказательства состоят в следующем: Напишем определяющую систему E уравнений функции R в обозначении без скобок Лукашиевича. Вместе с этими уравнениями рассмотрим также уравнения

$$(1)\quad Uxy = VUxFyRxy,$$

$$(2)\quad VUxFy0 = 0,$$

$$(3)\quad VUxFyFz = UxFy,$$

где U и V— два новых функтори и F— знак последования, значит $Fx = x+1$. Легко видеть, что уравнение $R(k, y) = 0$ разрешимо в неотрицателных целых числах тогда и только тогда, если уравнение $UF^k00 = 0$ является следствием вышеуказанной системы E' уравнений с помощью шагов, допустимых по ходу вычисления значений общерекурсивных функций, т. е. с помощью замен какого-либо переменного x на Fx или 0 (повторение которого допускает замену x также на $F0$, $FF0 = F^2 0$, $FFF0 = F^3 0$ и т. д.) и с помощью замещения одного члена некоторого уравнения другим. Поставим в соотношение с каждой переменной $x, y, \ldots, w$ два „субститутора" $x_0, x_+, y_0, y_+, \ldots, w_0, w_+$, дальше один общий „абсорптор" a и обширим системы уравнений E' с уравнениями

$$(4)\quad x_0x = 0x_0, \qquad (5)\quad x_+x = Fxx_+,$$

$$(6)\quad x_00 = 0x_0, \qquad (7)\quad x_+0 = 0x_+,$$

$$(8)\quad x_0y = yx_0, \qquad (9)\quad x_+y = yx_+,$$

$$(10)\quad x_0g = gx_0, \qquad (11)\quad x_+g = gx_+,$$

$$(12)\quad x_0a = e, \qquad (13)\quad x_+a = e$$

и

$$(14)\quad e = e,$$

где e— пустое слово, x— любая переменная фигурирующая в E', y— любая такая-же переменная, азнящаяся от x и g— любой функтор фигурирующий в E'. Тогда вместо замена x на Fx или 0 в некотором уравнении можно ее умножить на x_0 соответственно x_+ слева и на a с права и после этого применить уравнения (4), (6), (8), (10), (12) соответственно (5), (7), (9), (11), (13). Из этого следует, что проблема тождества ассоциативной системы, образованной над алфавитом $\{0, x, x_0, x_+, y, y_0, y_+, \ldots, w, w_0, w_+, a, F, \ldots, R, U, V\}$ (где $F, \ldots, R, U, V$— функторы фигурирующие в E') с помощью уравнений E' и уравнений (4)—(14)

неразрешима в алгоритмах. В самом деле, в противном случае существовал бы алгоритм посредством которого для каждого k можно было бы узнать, следует-ли соотношение $UF_k00=0$ из вышеупомянутых образующих уравнений.

Метод с легкими модификациями может быть применен для доказательства теоремы Маркова, согласно которой существует конечно-образованная ассоциативная система, проблема тождества которой разрешима, но одна из односторонних проблем делимости неразрешима с помощью алгоритма. По существу тот же метод дает доказательство следующей теоремы Маркова: существует конечно-образованная система с разрешимой с помощью алгоритма проблемой тождества и одной из односторонних проблем делимости но с неразрешимой другой проблемой делимости.

SUR L'ORDRE DE GRANDEUR DE L'APPROXIMATION D'UNE FONCTION PÉRIODIQUE PAR LES SOMMES DE FEJÉR

Par

GEORGES ALEXITS (Budapest), membre de l'Académie

1. Introduction

Soit E un espace métrique de fonctions périodiques définies dans l'intervalle $0 \leqq x \leqq 2\pi$. Désignons par $\varrho_E(f, \varphi)$ la distance des points $f(x)$ et $\varphi(x)$ d'un espace fonctionnel E contenant les polynomes trigonométriques. Pour représenter les fonctions plus ou moins compliquées par des fonctions élémentaires, accessibles à une recherche directe, on peut se demander, quelle est la grandeur de la distance $\varrho_E(f, T_n)$ où $T_n(x)$ est un polynome trigonométrique d'ordre n. Inversement, on peut se demander, quelles sont les propriétés différentielles de $f(x)$ qu'on peut garantir par la connaissance de l'ordre de grandeur des distances $\varrho_E(f, T_n)$ pour une suite $\{T_n(x)\}$ de polynomes trigonométriques. Le premier de ces problèmes remonte en essence jusqu'à TCHÉBYCHEV, tandis que le deuxième a été posé et en grande partie résolu par SERGE BERNSTEIN. L'ensemble des théorèmes concernant ces deux questions constitue une théorie bien large qui fait partie de l'ainsi dite théorie constructive des fonctions.

Désignons par $\sigma_n(x)$ la n-ième somme de FEJÉR de la série de Fourier

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

et par $\tilde{\sigma}_n(x)$ les sommes de FEJÉR de la série conjuguée

$$\tilde{f}(x) \sim c + \sum_{n=1}^{\infty} (b_n \cos nx - a_n \sin nx) \tag{2}$$

où la fonction conjuguée $\tilde{f}(x)$ est univoquement déterminée, à une constante près, par la fonction $f(x)$. Le fait que la fonction $f(x)$ satisfait à une condition de Lipschitz d'ordre α, c'est à dire que

$$|f(x+\delta) - f(x)| \leqq K\delta^\alpha \qquad (0 < \alpha \leqq 1),$$

sera désigné par $f \in \mathrm{Lip}\,\alpha$. Si l'on a

$$\left\{ \int_0^{2\pi} |f(x+\delta) - f(x)|^p \, dx \right\}^{\frac{1}{p}} \leqq K\delta^\alpha \qquad (0 < \alpha \leqq 1, p \geqq 1),$$

nous écrirons $f \in \mathrm{Lip}\,(\alpha, p)$. Les théorèmes suivants sont connus:

I. *La condition nécessaire et suffisante pour que* $f \in \mathrm{Lip}\,\alpha$ *est que, dans l'espace C des fonctions périodiques continues, on ait*

$$\varrho_C(f, \sigma_n) = O\left(\frac{1}{n^\alpha}\right)$$

pourvu que $0 < \alpha < 1$. *Si* $\alpha = 1$, *on n'obtient que*

$$\varrho_C(f, \sigma_n) = O\left(\frac{\log n}{n}\right).$$

L'existence d'une suite $\{T_n(x)\}$ *telle que*

$$\varrho_C(f, T_n) = O\left(\frac{1}{n}\right)$$

permet seulement à conclure que le module de continuité de la fonction $f(x)$ *est de la grandeur* $\delta \log 1/\delta$.

II. *La condition nécessaire et suffisante pour que* $f \in \mathrm{Lip}\,(\alpha, p)$ *est l'existence d'une suite* $\{T_n(x)\}$ *de polynomes trigonométriques,* $T_n(x)$ *ayant l'ordre n, tels que l'on ait*

$$\varrho_{L^p}(f, T_n) = O\left(\frac{1}{n^\alpha}\right)$$

pourvu que $0 < \alpha < 1$. *Quant au cas*

$$\varrho_{L^p}(f, T_n) = O\left(\frac{1}{n}\right),$$

on n'en peut conclure qu'à ce que

$$\left\{\int_0^{2\pi} |f(x+\delta) - f(x)|^p dx\right\}^{\frac{1}{p}} = O(x \log 1/x).$$

Le théorème I est dû à S. BERNSTEIN,[1] le théorème II se trouve, même dans une forme plus générale, dans le livre d'ACHYESER.[2]

Il y a une douzaine d'années que j'ai démontré[3] le théorème suivant:

III. *Pour les éléments f de l'espace C, la relation*

$$\varrho_C(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right)$$

est nécessaire et suffisante pour que $f \in \mathrm{Lip}\,1$.

[1] S. BERNSTEIN, Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné, *Mémoires Acad. Belgique* (2), 4 (1912), p. 1—104.

[2] Н. И. Ахиезер, Лекции по теории аппроксимации (Москва—Ленинград, 1947), p. 226.

[3] G. ALEXITS, A Fourier-sor Cesàro-közepeivel való approximáció nagyságrendjéről, *Mat. Fiz. Lapok,* **48** (1941), p. 410—421 (hongrois) et p. 421—422 (français).

Ce théorème est un complément du théorème I dans le cas critique $\alpha = 1$. En suite, BÉLA SZ.-NAGY[4] a déterminé aussi la fonction $f(x)$ pour laquelle la distance $\varrho_C(\tilde{f}, \tilde{\sigma}_n)$ est maximale. Ces résultats, parus d'abord seulement en hongrois, sont restés tout à fait inaperçus pendant la guerre. Une partie du théorème III, notamment la nécessité de la condition, a été découverte, dans le cadre de recherches générales, aussi par S. M. NIKOLSKY[5] et ZYGMUND.[6] C'est après la guerre que M. ZAMANSKY[7] a recommencé la recherche de ce problème; ses résultats et ses méthodes de démonstration ont plusieurs points de contact avec les miennes.

La méthode que j'ai employé dans la démonstration, est applicable non seulement à la question de l'approximation dans l'espace C, mais aussi aux problèmes d'approximation dans l'espace $L^p (p \geqq 1)$, ce qui nous permet de démontrer des compléments analogues du théorème II, comme le théorème III est, dans le cas critique $f \in \text{Lip } 1$, un complément du théorème I. Parmi les résultats obtenus, le cas de l'approximation dans l'espace $L = L^1$ paraît avoir un intérêt spécial. C'est que dans ce cas il s'agit de l'approximation des fonctions à variation bornée; classe de fonctions dont l'approximation — tant que je sais — n'était que très peu recherchée de ce point de vue malgré la grande importance de cette classe. Le résultat que nous obtiendrons est le suivant:[8]

Pour que la fonction $f(x) \in L$ soit à variation bornée, il faut et il suffit que l'on ait

$$\varrho_L(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right).$$

[4] B. SZ.-NAGY, Függvények megközelítése Fourier-soruk számtani közepeivel, *Mat. Fiz. Lapok*, **49** (1942), p. 123—138 (hongrois) et p. 138 (allemand). Approximation der Funktionen durch die arithmetischen Mittel ihrer Fourierschen Reihen, *Acta Sci. Math. Szeged*, **11** (1946–1948), p. 71–84.

[5] С. М. Никольский, Приближение периодических функций тригонометрическими многочленами, Труди Матем. Инст. им. В. А. Стеклова, **15** (1945), p. 1—58

[6] A. ZYGMUND, On the degree of approximation of functions by Fejér means, *Bulletin American Math. Soc.*, **51** (1945), p. 274—278.

[7] M. ZAMANSKY, Sur l'approximation des fonctions continues périodiques, I. *C. R. Acad. Sci. Paris*, **227** (1948), p. 1011—1013; II. *Ibid.*, **228** (1949), p. 460—461. Classes de saturation de certains procédés d'approximation des séries de Fourier des fonctions continues et applications à quelques problèmes d'approximation, *Ann. Sci. Ecole Norm. Sup.*, (3), **66** (1949), p. 19—93. Sur la sommation des séries de Fourier, *C. R. Acad. Sci. Paris*, **231** (1950), p. 1118—1120.

[8] Puisque deux fonctions ne différant qu'en un ensemble de mesure nulle engendrent la même série de Fourier, nous identifions toutes ces fonctions équivalentes. Nous démontrerons donc, à vrai dire, seulement que la fonction $f(x)$ ne diffère qu'en un ensemble de mesure nulle d'une fonction à variation bornée. Or si $f(x)$ est continue, elle est univoquement déterminée par cette propriété.

Il est remarquable que la fonction conjugée $\tilde{f}(x)$ n'est pas, en général, à variation bornée et les sommes de Fejér $\tilde{\sigma}_n(x)$ peuvent diverger en certains points. L'approximation de $\tilde{f}(x)$ par $\tilde{\sigma}_n(x)$ dans l'espace M des fonctions bornées est donc inutile. Ainsi par exemple les sommes de FEJÉR de la série

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n} \sim -\frac{1}{2} \log 4 \sin^2 \frac{x}{2}$$

ne convergent pas au point $x=0$, bien que c'est la série conjuguée du développement

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} \sim \frac{\pi - x}{2}.$$

On voit de ces recherches que la fonction conjuguée $\tilde{f}(x)$ et les moyennes de FEJÉR $\tilde{\sigma}_n(x)$ jouent un rôle important dans la théorie de l'approximation des fonctions satisfaisant à une condition de Lipschitz proprement dite ou bien généralisée. Or ce rôle n'est pas du tout formel et n'est pas borné seulement à la condition de Lipschitz d'ordre 1. Au contraire, en désignant par L^{∞} l'espace C et par $\mathrm{Lip}(\alpha, \infty)$ la condition $\mathrm{Lip}\,\alpha$, on obtient un résultat (théorème 3) qui contient l'essentiel de tous les théorèmes concernant l'approximation des fonctions satisfaisantes à une condition de Lipschitz quelconque.

La condition nécessaire et suffisante pour que la fonction $f \in L^p$ satisfasse à la condition $\mathrm{Lip}(\alpha, p)$, *les valeurs $p=1$ et $p=\infty$ y comprises, est que*

$$\varrho_{L^p}(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n^{\alpha}}\right).$$

2. Deux lemmes sur la sommation des séries de fonctions

Soit E un espace vectoriel métrique complet de fonctions définies sur un ensemble quelconque. Les points $\varphi(x)$ de E sont donc considérés comme des vecteurs dont les longueurs seront désignées par $\|\varphi\|$; la distance $\varrho_E(f, \varphi)$ des points $f(x)$ et $\varphi(x)$ est alors $\|f-\varphi\|$. Désignons par $S_n^*(x)$ la n-ième moyenne arithmétique de la série

$$\varphi_1(x) + \varphi_2(x) + \cdots + \varphi_n(x) + \cdots$$

et par $S_n(x)$ la n-ième moyenne de la série

$$\varphi_1(x) + \frac{\varphi_2(x)}{2} + \cdots + \frac{\varphi_n(x)}{n} + \cdots$$

où les fonctions $\varphi_1(x), \varphi_2(x), \ldots, S_n^*(x)$ et $S_n(x)$ appartiennent à l'espace E.

LEMME 1. *S'il existe une constante K telle que $\|S_n^*\| \leqq K$ pour $n = 1, 2, \ldots$, alors la suite $\{S_n\}$ converge selon la métrique de E vers un poit $S \in E$ et*

$$\varrho_E(S, S_n) \leq \frac{3K}{n+1}.$$

En effet, puisque

$$S_\nu(x)-S_{\nu-1}(x)=\frac{\sum_{k=1}^{\nu}\varphi_k(x)}{\nu(\nu+1)},$$

nous obtenons par une transformation d'Abel

$$\varrho_E(S_m,S_n)=\left\|\sum_{\nu=n+1}^{m}(S_\nu-S_{\nu-1})\right\|\leq\sum_{\nu=n+1}^{m-1}\frac{2}{\nu(\nu+1)(\nu+2)}\left\|\sum_{l=1}^{\nu}\sum_{k=1}^{l}\varphi_k\right\|+$$
$$+\frac{\left\|\sum_{l=1}^{m}\sum_{k=1}^{l}\varphi_k\right\|}{(m+1)(m+2)}+\frac{\left\|\sum_{l=1}^{n}\sum_{k=1}^{l}\varphi_k\right\|}{(n+1)(n+2)}.$$

D'après l'hypothèse on a

$$\frac{\left\|\sum_{l=1}^{\nu}\sum_{k=1}^{l}\varphi_k\right\|}{\nu+1}=\|S_\nu^*\|\leq K,$$

par conséquent

$$\varrho_E(S_m,S_n)\leq\sum_{\nu=n+1}^{m-1}\frac{2K}{\nu(\nu+2)}+\frac{K}{m+2}+\frac{K}{n+2}<\frac{3K}{n+1}+\frac{K}{m+2}.$$

La distance $\varrho_E(S_m,S_n)$ devient donc pour tout $m\geqq n+2$ aussi petite que l'on veut pourvu que n soit assez grand. Il en résulte, l'espace E étant complet, l'existence d'un point $S\in E$ tel que

$$\varrho_E(S,S_n)=\lim_{m\to\infty}\varrho_E(S_m,S_n).$$

On obtient donc de l'inégalité précédente:

$$\varrho_E(S,S_n)=\frac{3K}{n+1}, \qquad \text{c. q. f. d.}$$

LEMME 2. *S'il existe une constante K telle que*

$$\varrho_E(S,S_n)\leq\frac{K}{n+1},$$

alors $\|S_n^*\|\leq 4K$.

Désignons par $\bar S_n(x)$ la n-ième moyenne $(C,2)$ de la série $\varphi_1(x)+ +\frac{\varphi_2(x)}{2}+\cdots$. En posant $S_0(x)\equiv 0$, on peut écrire

$$\bar S_n(x)=\frac{\sum_{k=0}^{n}\binom{k+1}{k}S_k(x)}{\binom{n+2}{n}},$$

par conséquent

$$\varrho_E(S, \bar{S}_n) \leq \frac{\sum_{k=0}^{n} \binom{k+1}{k} \|S - S_k\|}{\binom{n+2}{n}} \leq \frac{\sum_{k=0}^{n} \binom{k+1}{k} \frac{K}{k+1}}{\binom{n+1}{n}} =$$

$$= \frac{K(n+1)}{\binom{n+2}{n}} = \frac{2K}{n+2}.$$

Vu que $S_n^* = (n+2)(S_n - \bar{S}_n)$, il s'ensuit

$$\|S_n^*\| \leq (n+2)(\|S - S_n\| + \|S - \bar{S}_n\|) \leq (n+2)\left(\frac{K}{n+1} + \frac{2K}{n+2}\right) \leq 4K, \text{c. q. f. d.}$$

3. Ordre de grandeur de l'approximation par les sommes de Fejér de la série conjuguée

Les deux lemmes démontrés au paragraphe précédent contiennent l'essentiel de l'approximation des fonctions satisfaisantes à la condition Lip 1 ou Lip$(1, p)$.

THÉORÈME 1. *La condition nécessaire et suffisante pour que* $f \in \text{Lip}\,1$, *ou* $f \in \text{Lip}\,(1, p)$ $(p > 1)$, *est que l'on ait, selon le cas,*

$$\varrho_C(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right) \quad \text{ou} \quad \varrho_{L^p}(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right). \tag{3}$$

En effet, si $f \in \text{Lip}\,1$, la dérivée $f'(x)$ existe presque partout et appartient à l'espace M. Si $f \in \text{Lip}\,(1, p)$, la fonction $f(x)$ est, d'après un théorème de HARDY et LITTLEWOOD,[9] équivalente à une fonction $g(x)$ qui est l'intégrale indéfinie de sa dérivée $g'(x)$ appartenant à l'espace L^p. Dans tous les deux cas on a donc

$$f'(x) \sim \sum_{k=1}^{\infty} k(b_k \cos kx - a_k \sin kx) \sim g'(x).$$

Si $f' \in M$, les valeurs absolues des sommes de FEJÉR $S_n^*(x)$ de cette série ne surpassent pas la valeur $K = \text{Max}\,|f'(x)|$. En posant donc $\|S_n^*(x)\| = \text{Max}\,|S_n^*(x)|$, on obtient d'après le lemme 1:

$$|\tilde{f}(x) - \tilde{\sigma}_n(x)| \leq \frac{3K}{n+1},$$

ce qui équivaut à (3) dans l'espace C. — Si $g' \in L^p$, on a[10]

$$\|S_n^*\| = \left\{ \int_0^{2\pi} |S_n^*(x)|^p \, dx \right\}^{\frac{1}{p}} \leq K;$$

[9] G. H. HARDY—J. E. LITTLEWOOD, Some properties of fractional integrals, I. *Math. Zeitschrift*, 27 (1928), p. 565—606.

[10] Voir A. ZYGMUND, *Trigonometrical series* (Warszawa - Lwów, 1935), p. 84.

il s'ensuit donc, $f(x)$ étant équivalent à $g(x)$ et par conséquent $\tilde{f}(x)$ à $\tilde{g}(x)$, d'après le lemme 1:

$$\left\{ \int_0^{2\pi} |\tilde{f}(x) - \tilde{\sigma}_n(x)|^p \, dx \right\}^{\frac{1}{p}} \leq \frac{3K}{n+1},$$

ce qui équivaut à (3) dans l'espace L^p.

Admettons, inversement, que la condition (3) est satisfaite. Il en résulte dans l'espace C, en vertu du lemme 2, l'existence d'une constante K telle que

$$(4) \qquad \|S_n^*\| = \operatorname{Max} |S_n^*(x)| \leq 4K.$$

On obtient de la même manière dans l'espace L^p:

$$(5) \qquad \left\{ \int_0^{2\pi} |S_n^*(x)|^p \, dx \right\}^{\frac{1}{p}} \leq 4K.$$

La relation (4) est la condition suffisante pour que la série

$$(6) \qquad \sum_{k=1}^{\infty} k(b_k \cos kx - a_k \sin kx)$$

soit le développement d'une fonction bornée. Par conséquent, la fonction $f(x)$ figurant dans la relation (1) appartient à la classe Lip 1. — La relation (5) entraîne[11] que (6) soit le développement d'une fonction appartenant à l'espace L^p; la fonction $f(x)$ est par conséquent équivalente à l'intégrale indéfinie d'une fonction appartenant à l'espace L^p, ce qui implique, d'après le théorème cité de HARDY et LITTLEWOOD, $f \in \operatorname{Lip}(1, p)$. Notre théorème est donc entièrement démontré.

THÉORÈME 2. *Une condition nécessaire et suffisante pour que la fonction $f(x) \in L$ soit équivalente à une fonction à variation bornée, est que*

$$(7) \qquad \varrho_L(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right).$$

Si $f(x)$ est équivalente à une fonction à variation bornée, les sommes de FEJÉR de la série (6) possèdent la propriété[12]

$$(8) \qquad \|S_n^*\| = \int_0^{2\pi} |S_n^*(x)| \, dx \leq K.$$

Il s'ensuit donc d'après le lemme 1:

$$(9) \qquad \int_0^{2\pi} |\tilde{f}(x) - \tilde{\sigma}_n(x)| \, dx = O\left(\frac{1}{n}\right).$$

La condition (7) de notre théorème est donc nécessaire. — Elle est aussi suffisante, parce qu'elle équivaut à (9), ce qui entraîne, d'après le lemme 2,

[11] Loc. cit [10], p. 84.
[12] Loc. cit. [10], p. 79.

l'inégalité (8); or (8) équivaut [13] à la condition que (1) soit le développement d'une fonction $f(x)$ équivalente à une fonction à variation bornée.

Remarque. Le théorème 2 est un élargissement du théorème 1 au cas $p=1$. En effet, d'après le théorème cité de HARDY et LITTLEWOOD, la propriété que $f(x)$ soit équivalente à une fonction à variation bornée équivaut á la propriété $f \in \mathrm{Lip}(1, 1)$. En désignant donc par L^∞ l'espace C et par $\mathrm{Lip}(1, \infty)$ la condition Lip 1, on peut unifier les théorèmes 1 et 2 dans la proposition suivante: *La condition nécessaire et suffisante pour que la fonction $f(x) \in L^p$ où $1 \leq p \leq \infty$ appartienne à* $\mathrm{Lip}(1,p)$ *est que*

$$\varrho_{L^p}(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right).$$

Mais cette proposition se laisse encore élargir de façon qu'elle contienne aussi les théorèmes I et II. C'est que la fonction conjuguée d'une fonction $f \in \mathrm{Lip}\,\alpha$ ou $f \in \mathrm{Lip}(\alpha, p)$ satisfait pour $0<\alpha<1$ à la même condition.[14] On peut donc remplacer dans les théorèmes I et II les différences $f(x)-\sigma_n(x)$, respectivement $f(x)-T_n(x)$ par les différences conjuguées $\tilde{f}(x)-\tilde{\sigma}_n(x)$, respectivement $\tilde{f}(x)-\tilde{T}_n(x)$. Après cela, toutes les propositions se rapportent à l'approximation par les sommes de FEJÉR de la série conjuguée. Nous pouvons alors énoncer le théorème suivant qui unifie tous les autres:

THÉORÈME 3. *La condition nécessaire et suffisante pour que la fonction $f(x) \in L^p$ appartienne à* $\mathrm{Lip}(\alpha, p)$ *est, pour tout $0<\alpha \leq 1$ et $1 \leq p \leq \infty$,*

$$\varrho_{L^p}(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n^\alpha}\right).$$

3. L'ordre de grandeur de l'approximation par les suites partielles de la série de Fourier et de sa série conjuguée

Le théorème II a été énoncé sous une forme plus générale que celle que nous avons donnée à l'introduction: au lieu de $f \in \mathrm{Lip}(\alpha, p)$ on aurait pu exiger que la r-ième dérivée $f^{(r)}(x)$ de $f(x)$ satisfasse à la condition $f^{(r)} \in \mathrm{Lip}(\alpha, p)$; l'ordre de grandeur $O\left(\frac{1}{n^\alpha}\right)$ est alors à remplacer par $O\left(\frac{1}{n^{r+\alpha}}\right)$. Or cette forme du théorème II ne se laisse pas étendre à l'approximation par les sommes de FEJÉR. Envisageons, en effet, les fonctions conjuguées $f(x) = \cos x$ et $\tilde{f}(x) = \sin x$. Les sommes de FEJÉR sont

$$\sigma_n(x) = \left(1-\frac{1}{n+1}\right)\cos x, \qquad \tilde{\sigma}_n(x) = \left(1-\frac{1}{n+1}\right)\sin x,$$

[13] Loc. cit. [10], p. 79.
[14] Loc. cit. [2], p. 225 et 230.

donc

$$f(x)-\sigma_n(x)=\frac{\cos x}{n+1},\qquad \tilde{f}(x)-\tilde{\sigma}_n(x)=\frac{\sin x}{n+1};$$

c'est à dire que l'ordre de grandeur de l'approximation atteint en tout espace L^p — les valeurs $p=1$ et ∞ y comprises — les valeurs

$$\varrho_{L^p}(f,\sigma_n)=O\left(\frac{1}{n}\right),\qquad \varrho_{L^p}(\tilde{f},\tilde{\sigma}_n)=O\left(\frac{1}{n}\right)$$

et cet ordre de grandeur ne se laisse pas améliorer, bien que toutes les dérivées des fonctions $\cos x$ et $\sin x$ existent et appartiennent à toutes les classes $\mathrm{Lip}\,(\alpha,p)$. On voit donc que le théorème 3 ne donne non plus lieu à une amélioration de l'ordre de grandeur $O\left(\frac{1}{n}\right)$. Or la situation est tout autre lorsque l'approximation est effectuée par les sommes partielles de la série de Fourier ou celles de la série conjuguée. Pour la rechercher, nous démontrons d'abord le lemme suivant:

LEMME 3. *Si l'on a, dans un espace vectoriel métrique complet E, pour tout $m \geqq n+2$,*

$$\left\|\sum_{k=n+1}^{m}\varphi_k(x)\right\|=O(\lambda_n)$$

où $0<\lambda_n$ et $\lambda_n=o(n)$, alors $\sum\frac{\varphi_n(x)}{n}$ converge selon la métrique de E vers un élément $s(x)\in E$ et, en posant $s_n(x)=\sum_{k=1}^{n}\frac{\varphi_k(x)}{k}$, on a

$$\varrho_E(s,s_n)=O\left(\frac{\lambda_n}{n}\right).$$

On obtient, en effet, par une transformation d'Abel

$$\varrho_E(s_m,s_n)\leqq\sum_{\nu=n+1}^{m-1}\frac{\left\|\sum_{k=n+1}^{\nu}\varphi_k(x)\right\|}{\nu(\nu+1)}+\frac{\left\|\sum_{k=n+1}^{m}\varphi_k(x)\right\|}{m}=$$

$$=O(\lambda_n)\sum_{\nu=n+1}^{m-1}\frac{1}{\nu(\nu+1)}+O\left(\frac{\lambda_n}{m}\right)=O\left(\frac{\lambda_n}{n}\right).$$

D'après l'hypothèse $\lambda_n=o(n)$, par conséquent $\varrho_E(s_m,s_n)=o(1)$ pour tout $m\geqq n+2$, il existe donc dans l'espace complet E un point $s(x)$ tel que

$$\varrho_E(s,s_n)=\lim_{m\to\infty}\varrho_E(s_m,s_n);$$

il s'ensuit alors de l'inégalité précédente

$$\varrho_E(s,s_n)=O\left(\frac{\lambda_n}{n}\right),\qquad \text{c. q. f. d.}$$

Désignons maitenant par $s_n(x)$ la n-ième somme partielle de la série de Fourier (1) et par $\tilde{s}_n(x)$ celle de la série conjuguée (2); $f^{(r)}(x)$ est la r-ième dérivée de la fonction $f(x)$ et nous posons $f^{(0)}(x)=f(x)$. Les sommes partielles de la série de Fourier de $f^{(r)}(x)$ seront désignées par $s_n^r(x)$, celles de sa série conjuguée par $\tilde{s}_n^r(x)$.

THÉORÈME 4. *Pour tout $f \in L^p$ où $p>1$, une condition nécessaire et suffisante pour que $f^{(r)} \in \mathrm{Lip}(1,p)$, est que*

$$\varrho_{L^p}(f, s_n) = o\left(\frac{1}{n^{r+1}}\right), \qquad \varrho_{L^p}(\tilde{f}, \tilde{s}_n) = o\left(\frac{1}{n^{r+1}}\right).$$

Si $f^{(r)} \in \mathrm{Lip}(1,p)$, alors $f^{(r)}(x)$ est équivalent à une fonction $g(x)$ dont la dérivée $g'(x)$ existe presque partout et appartient à l'espace L^p (théorème cité de HARDY et LITTLEWOOD). Les fonctions $f^{(r)}(x)$ et $g(x)$ ne différant qu'en un ensemble de mesure nulle, leur séries de Fourier sont identiques et, puisque $g(x)$ est absolument continue, la série de Fourier de $g'(x)$ est la dérivée formelle de la série de Fourier de $f^{(r)}(x)$. C'est à dire que la n-ième somme partielle de la série de Fourier de $g'(x)$ est $s_n^{r+1}(x)$ ou $\tilde{s}_n^{r+1}(x)$ selon le cas que r est pair ou impair. Or $g' \in L^p$ entraîne les relations [15]

$$\varrho_{L^p}(g', s_n^{r+1}) = o(1), \qquad \varrho_{L^p}(\tilde{g}', \tilde{s}_n^{r+1}) = o(1).$$

Il s'ensuit d'après le lemme 3:

$$\varrho_{L^p}(f^{(r)}, s_n^r) = o\left(\frac{1}{n}\right), \qquad \varrho_{L^p}(\tilde{f}^{(r)}, \tilde{s}_n^r) = o\left(\frac{1}{n}\right).$$

En répétant r-fois cette application du lemme 3, on obtient la proposition de notre théorème.

4. Approximation à l'exception d'un ensemble de mesure nulle

Si on ne cherche pas à obtenir des conditions nécessaires et suffisantes pour qu'une fonction appartienne à une certaine classe, on peut rechercher point par point l'ordre de grandeur de l'approximation. On pourrait caractériser de cette façon la rapidité de la convergence sous des conditions données. Désignons à ce but par $O_x(\lambda_n)$ et $o_x(\lambda_n)$ le fait que l'approximation au point x est de l'ordre $O(\lambda_n)$ ou $o(\lambda_n)$. On voit immédiatement que les lemmes 1 et 3 conservent leur validité, si l'on y remplace l'approximation dans un espace vectoriel par l'approximation en un point, en remplaçant aussi le symbole O par O_x. En conservant donc les notations de ces lemmes, on peut énoncer le

LEMME 4. *Si les relations $|S_n^*(x)| = O_x(1)$, respectivement $\left|\sum_{k=n+1}^{m} \varphi_k(x)\right| = O_x(\lambda_n)$ $(m \geq n+2)$ subsistent pour un point x, les suites $\{S_n(x)\}$, respective-*

[15] M. RIESZ, Sur les fonctions conjuguées, *Math. Zeitschrift,* **27** (1927), p. 218—214.

ment $\{s_n(x)\}$ *convergent vers des valeurs* $S(x)$, *respectivement* $s(x)$, *et la rapidité de leur convergence peut être mesurée par*

$$|S(x)-S_n(x)| = O_x\left(\frac{1}{n}\right), \qquad |s(x)-s_n(x)| = O_x\left(\frac{\lambda_n}{n}\right).$$

On en tire aisément des conséquences concernant l'ordre de grandeur de l'approximation par les sommes de FEJÉR.

THÉORÈME 5. *Si* $f \in \mathrm{Lip}\,(1, p)$ — *la valeur* $p = 1$ *y comprise* — *on a presque partout*

$$|f(x)-\sigma_n(x)| = O_x\left(\frac{1}{n}\right), \qquad |\tilde{f}(x)-\tilde{\sigma}_n(x)| = O_x\left(\frac{1}{n}\right).$$

En effet, si $p > 1$, les dérivées formelles $\sigma_n'(x)$ et $\tilde{\sigma}_n'(x)$ sont les sommes de FEJÉR des dérivées $\tilde{g}'(x)$ et $g'(x)$ existant presque partout et, comme nous avons déjà observé, appartenant à la classe L^p. Les suites $\{\sigma_n'(x)\}$ et $\{\tilde{\sigma}_n'(x)\}$ convergent donc presque partout, ce qui entraîne presque partout $|\sigma_n'(x)| = O_x(1)$ et $|\tilde{\sigma}_n'(x)| = O_x(1)$. La proposition de notre théorème s'ensuit donc du lemme 4. — Si $p = 1$, la fonction $f(x)$ est, comme nous avons déja remarqué, équivalente à une fonction à variation bornée; $\tilde{\sigma}_n'(x)$ est donc la n-ième somme de FEJÉR de la dérivée formelle d'une série de Fourier d'une fonction à variation bornée. Par conséquent, la suite $\{\tilde{\sigma}_n'(x)\}$ ainsi que la suite conjuguée $\{\sigma_n'(x)\}$ converge presque partout.[16] Il s'ensuit $|\sigma_n'(x)| = O_x(1)$ et $|\tilde{\sigma}_n'(x)| = O_x(1)$ presque partout, la proposition de notre théorème résulte donc dans ce cas aussi du lemme 4.

En ce qui concerne la rapidité de la convergence des suites partielles $s_n(x)$ et $\tilde{s}_n(x)$, on peut démontrer le

THÉORÈME 6. *Si* $f^{(r)}(x)$ *est à variation bornée, on a presque partout*

$$|f(x)-s_n(x)| = o_x\left(\frac{\log n}{n^{r+1}}\right), \qquad |\tilde{f}(x)-\tilde{s}_n(x)| = o_x\left(\frac{\log n}{n^{r+1}}\right).$$

En effet, on sait que $s_n^{r+1}(x) = o_x(\log n)$ et $\tilde{s}_n^{r+1}(x) = o_x(\log n)$ presque partout.[17] On peut donc prendre $\lambda_n = o(\log n)$ et la proposition résulte immédiatement par l'application r-fois répétée du lemme 4.

Ce théorème montre, en posant $r = 0$, que le critère de convergence de DIRICHLET—JORDAN est très faible, car il rend, sauf un ensemble de mesure nulle, beaucoup plus qu'il ne faudrait. Or, en admettant que la fonction $f^{(r)}(x)$, au lieu d'étre à variation bornée — ce qui équivaut à $f^{(r)} \in \mathrm{Lip}\,(1, 1)$ — appartienne à la classe $\mathrm{Lip}\,(1, 2)$, le théorème 6 se laisse encore considérablement améliorer:

16 Loc. cit. [10], p. 59.
17 Loc. cit. [10], p, 32.

THÉORÈME 7. *Si* $f^{(r)} \in \mathrm{Lip}(1, 2)$, *on a presque partout*

$$|f(x)-s_n(x)| = o_x\left(\frac{\sqrt{\log n}}{n^{r+1}}\right), \qquad |\tilde{f}(x)-\tilde{s}_n(x)| = o_x\left(\frac{\sqrt{\log n}}{n^{r+1}}\right).$$

En effet, dans ce cas, $f^{(r+1)} \in L^2$ et $\tilde{f}^{(r+1)} \in L^2$, les suites partielles $s_n^{r+1}(x)$ et $\tilde{s}_n^{r+2}(x)$ de leur séries de Fourier ont donc presque partout l'ordre de grandeur [18] $o_x(\sqrt{\log n})$. En prenant $\lambda_n = o(\sqrt{\log n})$ notre proposition résulte par l'application r-fois répétée du lemme 4.

(Reçu le 18 avril 1952.)

[18] A. KOLMOGOROFF — G. SELIVERSTOFF, Sur la convergence des séries de Fourier, *C. R. Acad. Sci. Paris,* **178** (1925). p 303—305. A. PLESSNER, Über Konvergenz von trigonometrischen Reihen, *Journ. f. reine u. angew. Math.,* **155** (1925), p. 15—25.

О ПОРЯДКЕ ВЕЛИЧИНЫ АППРОКСИМАЦИИ ПЕРИОДИЧЕСКИХ ФУНКЦИЙ ФЕЙЕРОВСКИМИ СРЕДНИМИ

Г. АЛЕКСИЧ (Будапешт)

(Резюме)

Пусть $f(x)$ функция периода 2π. Обозначим через $\sigma_n(x)$ n-ое Фейеровское среднее ряда Фурье функции $f(x)$. Если в метрическом функциональном пространстве E обозначить через $\varrho_E(f, \varphi)$ расстояние элементов $f \in E$ и $\varphi \in E$, то одной из основных теорем С. Н. Бернштейна можно дать следующую формулировку: в случае $0 < \alpha < 1$ необходимым и достаточным условием для того чтобы было $f \in \mathrm{Lip}\, 1$, является

$$\varrho_C(f, \sigma_n) = O\left(\frac{1}{n^\alpha}\right),$$

где C есть пространство непрерывных функций периода 2π; нельзя однако эту теорему распространить на случай $\alpha = 1$: в этом случае для аппроксимации с $\sigma_n(x)$ имеет силу более слабая оценка. Ахиезер дал следующее обобщение этого результата на пространство L^p: если $0 < \alpha < 1$ и $p > 1$, то для того чтобы r-ая производная $f(x)$ принадлежала классу $\mathrm{Lip}\,(\alpha, p)$, необходимо и достаточно, чтобы для всякого n существовал тригонометрический многочлен $T_n(x)$ порядка n, удовлетворяющий соотношению

$$\varrho_{L^p}(f, T_n) = O\left(\frac{1}{n^{r+\alpha}}\right);$$

на случай $\alpha = 1$ достаточность условия нельзя распространить. Если вместо $f(x)$ и $\sigma_n(x)$ рассматриваются сопряженные величины $\tilde{f}(x)$ и $\tilde{\sigma}_n(x)$, то теорему Бернштейна можно распространить на критический случай $\alpha = 1$ следующим образом: $f \in \mathrm{Lip}\, 1$ тогда и только тогда, если

$$\varrho_C(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right).$$

Этот результат был получен в 1941 г. Алексичем, а в 1942 г. Бела С.-Надь дал ему дальнейшее развитие. Вполне независимо от этого подобные результаты были получены во время войны Никольским и Зигмундом, в послевоенное же время Заманский в ряде работ опять трактовал эту тему.

В настоящей работе дается сосредоточение всех известных до сих пор результатов в единственную теорему, и их распространение на дальнейшие случаи. Вводя обозначения $\mathrm{Lip}\,(\alpha, \infty) = \mathrm{Lip}\,\alpha$ и $L^\infty = C$, на основе двух очень простых лемм об общих функциональных рядах доказываем следующую теорему:

Если $f \in L^p$, то для того чтобы было $f \in \mathrm{Lip}\,(\alpha, p)$ необходимым и достаточным является условие

$$\varrho_{L^p}(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n^\alpha}\right)$$

со включением и случаев $\alpha = 1$, $p = 1$ и $p = \infty$.

Если $\alpha < 1$ и $p = \infty$, то ввиду известной теоремы Привалова, следует теорема Бернштейна, а если $\alpha < 1$ и $1 < p < \infty$ то получается упомянутая теорема Ахиезера $(r = 0)$; в случае же $\alpha = 1$ и $p = \infty$ мы придем к упомянутой теореме Алексича—Никольского—Зигмунда. Особый интерес представляет случай в котором $\alpha = 1$ и $p = 1$, так как на основе одной из теорем Харди—Литтльвуда утверждение нашей теоремы в этом случае может формулироваться и следующим образом:

Функция $f \in L$ тогда и только тогда эквивалентна функции с ограниченным изменением, если

$$\varrho_L(\tilde{f}, \tilde{\sigma}_n) = O\left(\frac{1}{n}\right).$$

В упомянутой теореме Ахиезера в случае $r > 1$ нельзя писать $\sigma_n(x)$ вместо $T_n(x)$ (пример на противоположное $f(x) = \cos x$). Если однако через $s_n(x)$ обозначать n-ую частную сумму ряда Фурье для $f(x)$, а через $\tilde{s}_n(x)$ соответствующую сумму сопряженного ряда, то применяя одну теорему М. Рисса мы получим следующую теорему:

В случае $1 < p < \infty$ необходимым и достаточным для того, чтобы было $f \in \mathrm{Lip}\,(1, p)$, является условие

$$\varrho_{L^p}(f, s_n) = O\left(\frac{1}{n^{r+1}}\right) \quad \text{и} \quad \varrho_{L^p}(\tilde{f}, \tilde{s}_n) = O\left(\frac{1}{n^{r+1}}\right).$$

Наш метод делает возможным исследование скорости сходимости ряда Фурье и сопряженного ряда в отдельных точках. Мы доказываем, например, что если $f^{(r)}(x)$ функция с ограниченным изменением, то за исключением множества меры нуль

$$|f(x) - s_n(x)| = o_x\left(\frac{\log n}{n^{r+1}}\right) \quad \text{и} \quad |\tilde{f}(x) - \tilde{s}_n(x)| = o_x\left(\frac{\log n}{n^{r+1}}\right).$$

Случай $r = 0$ показывает, что классический критерий сходимости Дирихле слишком груб, ведь при его выполненности о ряде Фурье можно сказать значительно больше, чем надобно для наших целей.

ÜBER DEN ANNÄHERUNGSGRAD DER ORTHOGONALPOLYNOMENTWICKLUNGEN

Von

GEORG ALEXITS (Budapest), Mitglied der Akademie

1. Sei $w(x) \geqq 0$ eine im endlichen abgeschlossenen Intervall $[a, b]$ definierte L-integrierbare Funktion. Sie bestimmt bekanntlich bis auf das Vorzeichen eindeutig ein orthogonales und normiertes Polynomsystem $\{p_n(x)\}$. Hierbei bedeutet $p_n(x)$ ein Polynom von genau n-tem Grad, welches bezüglich $w(x)$ die Orthonormaleigenschaft

$$\int_a^b w(x)p_i(x)p_j(x)\,dx = \begin{cases} 0 & (i \neq j), \\ 1 & (i = j) \end{cases}$$

besitzt. Ist $\sqrt{w}f \in L^2$, so läßt sich die Entwicklung

$$f(x) \sim c_0 p_0(x) + c_1 p_1(x) + \cdots + c_n p_n(x) + \cdots \tag{1}$$

formal bestimmen. Bezeichne $s_n(x)$ die n-te Partialsumme der Entwicklung (1). Es entsteht die Frage: Wie groß ist der Annäherungsgrad

$$\varrho_n(a, b) = \operatorname*{Max}_{a \leqq x \leqq b} |f(x) - s_n(x)|,$$

wenn $f(x)$ zu einer bestimmten Funktionsklasse gehört? Die Beantwortung dieser Frage hat außer ihrem theoretischen Wert auch eine gewisse praktische Bedeutung, da es in der theoretischen Physik Funktionen gibt, welche eben durch ihre Entwicklungen in Orthogonalpolynomreihen dargestellt werden. Trotzdem scheint dieses Problem ganz allgemein durchaus nicht gelöst zu sein. Im Folgenden soll gezeigt werden, daß der Annäherungsgrad $\varrho_n(a', b')$ in jedem inneren Teilintervall $[a', b']$ von (a, b) für alle, überhaupt praktisch in Betracht kommenden Orthogonalpolynomentwicklungen im Wesen derselbe ist, wie der durch die Fourierentwicklung von $f(x)$ erreichbare Annäherungsgrad.

2. Die Abschätzung der sogenannten Lebesgueschen Funktionen des Orthogonalpolynomsystems $\{p_n(x)\}$ ergibt eine recht einfache Methode für die Abschätzung des Annäherungsgrades $\varrho_n(a', b')$. Dabei versteht man unter

$$L_n(x) = \int_a^b w(t) \left| \sum_{k=0}^{n} \varphi_k(t)\varphi_k(x) \right| dt$$

die n-te Lebesguesche Funktion des Systems $\{p_n(x)\}$. Da die Beziehung

$$P_n(x)-s_n(x)=\int_a^b [P_n(t)-f(t)]\, w(t) \sum_{k=0}^{n} p_k(t)p_k(x)\, dt$$

für jedes Polynom $P_n(x)$ genau n-ten Grades gilt, folgt

$$|f(x)-s_n(x)|\leq |f(x)-P_n(x)|+|P_n(x)-s_n(x)|\leq$$

$$\leq |f(x)-P_n(x)|+\int_a^b |P_n(t)-f(t)|\, w(t) \left| \sum_{k=0}^{n} p_k(t)p_k(x)\right| dt.$$

Wählt man also für $P_n(x)$ jenes Polynom n-ten Grades, welches $f(x)$ in $[a, b]$ am besten annähert und ist dieser Annäherungsgrad $E_n(f)$, so gilt

(2) $$\varrho_n(a', b')\leq E_n(f)\,[1+\underset{a'\leq x\leq b'}{\operatorname{Max}} L_n(x)].$$

Da $E_n(f)$ für viele wichtige Funktionsklassen bekannt ist, hängt das Weitere nur von der Abschätzung der Lebesgueschen Funktionen $L_n(x)$ ab.

3. *Gibt es zu einem inneren Teilintervall* $[a+h, b-h]$ *von* (a, b) *Konstanten* P_h *und* W_h *derart, daß für* $x\in\left[a+\frac{h}{2},\ b-\frac{h}{2}\right]$ *die Abschätzungen*

$$|p_n(x)|\leq P_h, \quad 0\leq w(x)\leq W_h$$

gelten, so besteht für $n\geq n_h$, *wo* $n_h>\frac{2-h}{h}$ *ist, die Ungleichung*

$$\underset{a+h\leq x\leq b-h}{\operatorname{Max}} L_n(x)\leq \frac{8cP\sqrt{W}}{h}+2P_h^2 W_h+4cP_h^2 W_h \log (b-a)(n+1),$$

wo zur Abkürzung $W=\int_a^b w(x)\, dx$ *gesetzt wurde.*

Zum Beweis bedienen wir uns eines von BÉLA SZ.-NAGY[1] herrührenden Gedankenganges, der eine Verbesserung eines von mir zu ähnlichen Zwecken verwendeten Verfahrens[2] ist. Teilen wir das Integrationsintervall $[a, b]$ in fünf Teile:

$$\int_a^b=\int_a^{a+\frac{h}{2}}+\int_{a+\frac{h}{2}}^{x-\frac{1}{n+1}}+\int_{x-\frac{1}{n+1}}^{x+\frac{1}{n+1}}+\int_{x+\frac{1}{n+1}}^{b-\frac{h}{2}}+\int_{b-\frac{h}{2}}^{b}.$$

Die bekannte Christoffel—Darbouxsche Formel lautet:

$$\sum_{k=0}^{n} p_k(t)p_k(x)=\frac{\alpha_n}{\alpha_{n+1}}\,\frac{p_{n+1}(t)p_n(x)+p_n(t)p_{n+1}(x)}{t-x},$$

[1] B. SZ.-NAGY, Über die Konvergenz von Reihen orthogonaler Polynome, *Math. Nachrichten,* **4** (1951), S. 50—55.

[2] G. ALEXITS, Sur la convergence des séries de polynomes orthogonaux, *Commentarii Math. Helvetici,* **16** (1944), S. 200—208.

wo α_n den (positiven) Koeffizienten von x^n in $p_n(x)$ bedeutet. Wegen

$$\int_a^b w_n(x) p_{n+1}(x) p_n(x) \cdot x \, dx = \int_a^b w(x) p_{n+1}(x) (\alpha_n x^{n+1} + \cdots) \, dx =$$

$$= \frac{\alpha_n}{\alpha_{n+1}} \int_a^b w(x) p_{n+1}^2(x) \, dx = \frac{\alpha_n}{\alpha_{n+1}}$$

folgt aus der Buniakowski—Schwarzschen Ungleichung

$$\frac{\alpha_n}{\alpha_{n+1}} \leq \left\{ \int_a^b w(x) p_{n+1}^2(x) \, dx \int_a^b w(x) p_n^2(x) \cdot x^2 \, dx \right\}^{\frac{1}{2}} \leq c = \operatorname{Max}(|a|, |b|),$$

also ist

$$\left| \sum_{k=0}^{n} p_k(t) p_k(x) \right| \leq c \left| \frac{p_{n+1}(t) p_n(x) + p_n(t) p_{n+1}(x)}{t-x} \right|.$$

Da $x \in [a+h, b-h]$, gilt

$$\int_a^{a+\frac{h}{2}} \leq \frac{2cP_h}{h} \int_a^{a+\frac{h}{2}} (|p_n(t)| + |p_{n+1}(t)|) \, w(t) \, dt \leq$$

$$\leq \frac{2cP_h}{h} \left\{ \left[\int_a^{a+\frac{h}{2}} w(t) \, dt \int_a^{a+\frac{h}{2}} w(t) p_n^2(t) \, dt \right]^{\frac{1}{2}} + \left[\int_a^{a+\frac{h}{2}} w(t) \int_a^{a+\frac{h}{2}} w(t) p_{n+1}^2(t) \, dt \right]^{\frac{1}{2}} \right\} \leq$$

$$\leq \frac{4cP_h \sqrt{W}}{h}.$$

Auf dieselbe Weise ergibt sich

$$\int_{b-\frac{h}{2}}^{b} \leq \frac{4cP_h \sqrt{W}}{h}.$$

Für das zweite Integral erhalten wir

$$\int_{x+\frac{h}{2}}^{x-\frac{1}{n+1}} \leq 2cP_h^2 W_h \int_{a+\frac{h}{2}}^{x-\frac{1}{n+1}} \frac{dt}{|t-x|} \leq 2cP_h^2 W_h \log(b-a)(n+1)$$

und ähnlich folgt

$$\int_{x+\frac{1}{n+1}}^{b-\frac{h}{2}} \leq 2cP_h^2 W_h \log(b-a)(n+1).$$

Für das dritte Integral erhalten wir endlich

$$\int\limits_{x-\frac{1}{n+1}}^{x+\frac{1}{n+1}} \leq P_h^2 W_h (n+1) \int\limits_{x-\frac{1}{n+1}}^{x+\frac{1}{n+1}} dx = 2 P_h^2 W_h .$$

Aus diesen fünf Ungleichungen ergibt sich unmittelbar unsere Behauptung.

4. Bezeichne nun $f^{(r)}(x)$ die r-te Ableitung der Funktion $f(x)$, wobei $f^{(0)}(x) = f(x)$ bedeuten soll. Das Stetigkeitsmaß der Ableitung $f^{(r)}(x)$ im Intervall $[a, b]$ bezeichnen wir mit $\omega_r(x)$. Das ist bekanntlich die obere Grenze der Differenz $|f^{(r)}(x+t) - f^{(r)}(x)|$ für alle $|t| \leq \delta$, wenn x das Intervall $[a, b]$ durchläuft. Es ist bekannt,[3] daß für $n > r$ die Abschätzung

$$E_n(f) \leq \frac{12 \cdot 6^r r^r}{r!} \frac{(b-a)^r}{n^r} \omega_r \left(\frac{b-a}{2(n-r)} \right)$$

gilt. Setzen wir also

$$C_r = \frac{12 \cdot 6^r r^r}{r!} (b-a)^r, \quad K_h = 1 + \frac{8 c P_h \sqrt{W}}{h} + 2 P_h^2 W_h ,$$

so erhalten wir aus (2) den folgenden Satz:

Gibt es Konstanten P_h und W_h derart, daß in $\left[a + \frac{h}{2}, b - \frac{h}{2}\right]$ die Abschätzungen $|p_n(x)| \leq P_h$ $(n = 0, 1, \cdots)$ und $w(x) \leq W_h$ gelten, und existiert $f^{(r)}(x)$ in $[a, b]$ überall, so läßt sich der Annäherungsgrad für $n \geq$ Max $(n_h, r+1)$ im Intervall $[a+h, b-h]$ wie folgt abschätzen:

$$\varrho_n(a+h, b-h) \leq 4 c P_h^2 W_h \log (b-a)(n+1) \cdot \frac{\omega_r \left(\frac{b-a}{2(n-r)} \right)}{n^r} + C_r K_h \frac{\omega_r \left(\frac{b-a}{2(n-r)} \right)}{n^r} .$$

5. Aus diesem Satz lassen sich leicht Annäherungssätze für Funktionen, deren Ableitungen einer Lipschitzbedingung oder einer anderen der üblichen Bedingungen entsprechen, eventuell auch mit etwas besseren Konstanten herleiten, worauf wir aber nicht eingehen wollen. Wir begnügen uns mit der Bemerkung, daß der Annäherungsgrad auch in diesen Fällen für die wichtigsten in Betracht kommenden Orthogonalpolynomreihen, wie z. B. die nach den klassischen Jacobischen Polynomen fortschreitenden Entwicklungen, mit explizite bestimmten Konstanten aus unserem Satz ohne Weiteres herauslesbar ist. In den Anwendungen spielt aber oft eher die mittlere Abweichung der Partialsumme $s_n(x)$ von $f(x)$, also der quadratische Annäherungsgrad

$$\varrho_n = \left\{ \int_a^b [f(x) - s_n(x)]^2 w(x)\, dx \right\}^{\frac{1}{2}},$$

[3] И. П. Натансон, Конструктивная теория функций (Москва—Ленинград, 1949), S. 164.

eine Rolle, als der Annäherungsgrad $\varrho_n(a, b)$. Um ϱ_n zu bestimmen, beschränken wir uns einfachkeitshalber auf das Intervall $[-1, 1]$ und führen $\vartheta = \arccos x$ als neue Variable ein. Setzen wir dann

$$a_n = \frac{2}{\pi}\int_0^\pi f(\cos\vartheta)\cos n\vartheta\, d\vartheta,$$

so ergibt sich für den quadratischen Annäherungsgrad ϱ_n eine gute Abschätzung aus dem folgenden Lemma:[4] *Ist im Intervall* $[-1, 1]$ *fast überall* $w(x) \leqq W/\sqrt{1-x^2}$, *so gilt*

$$\varrho_n \leqq \sqrt{\frac{\pi W}{2}\sum_{k=n+1}^{\infty} a_k^2}\,. \tag{3}$$

Setzen wir nun voraus, es sei

$$\max_{-1\leqq x\leqq 1}|f^{(\nu)}(x)| \leqq M \qquad (\nu = 1, 2, \ldots, r).$$

Nach Einführung von $\vartheta = \arccos x$ ergibt sich

$$\frac{df}{d\vartheta} = -\frac{df}{d(\cos\vartheta)}\sin\vartheta,$$

$$\frac{d^2 f}{d\vartheta^2} = \frac{d^2 f}{d(\cos\vartheta)^2}\sin^2\vartheta - \frac{df}{d(\cos\vartheta)}\cos\vartheta,$$

usw. Wegen $\dfrac{d^\nu f}{d(\cos\vartheta)^\nu} = f^{(\nu)}(x)$ folgt also nach unserer Annahme: $\left|\dfrac{df}{d\vartheta}\right| \leqq M$, $\left|\dfrac{d^2 f}{d\vartheta^2}\right| \leqq 2M$, usw. Um $\left|\dfrac{d^r f}{d\vartheta^r}\right|$ für ein beliebiges r abzuschätzen, nehmen wir an, daß $\dfrac{d^\nu f}{d\vartheta^\nu}$ für $\nu \leqq r-1$ höchstens aus $3^{\nu-1}$ Gliedern besteht und jedes Glied von der Form

$$\alpha_i \frac{d^j f}{d(\cos\vartheta)^j}\sin^k\vartheta\cos^l\vartheta$$

mit $|\alpha_i| \leqq (\nu-1)!$ und $j, k, l \leqq \nu$ ist. Diese Annahme trifft für $\nu = 1$ zu, weiter schließen wir durch Induktion. Für $\nu = r$ besteht dann $\dfrac{d^r f}{d\vartheta^r}$ aus höchstens 3^r Gliedern, welche von der Form

$$\beta_{i'}\frac{d^{j'} f}{d(\cos\vartheta)^{j'}}\sin^{k'}\cos^{l'}\vartheta$$

[4] G. Alexits, A. a. O. [2], Vgl. 203. S. auch die Bemerkung von B. Sz.-Nagy in der Fußnote meiner Arbeit: Sur la convergence et la sommabilité presque partout des séries de polynomes orthogonaux, *Acta Sci. Math.*, 12 B (1950), S. 223—225.

sind, wo $|\beta_{i'}| \leq (r-1)\,|\alpha_i| \leq (r-1)!$ und $j', k', l' \leq j+1, k+1, l+1 \leq r$ ist. Damit ist der Induktionsschluß zu Ende geführt, woraus

$$\left|\frac{d^r f}{d\vartheta^r}\right| \leq (r-1)!\; 3^r M$$

folgt. Es ergibt sich somit die Abschätzung

$$|a_n| \leq \frac{2}{n^r \pi} \int_0^{2\pi} \left|\frac{d^r f}{d\vartheta^r}\right| \left|\begin{matrix}\cos\\ \sin\end{matrix} n\vartheta\right| d\vartheta \leq \frac{(r-1)!\; 3^r M}{n^r}.$$

Setzen wir diesen Wert in (5) ein, so erhalten wir den folgenden Satz:

Ist im Intervall $[-1, 1]$ *fast überall* $w(x) \leq W/\sqrt{1-x^2}$ *und gilt* $\operatorname{Max} |f^{(\nu)}(x)| \leq M$ *für alle* $\nu = 1, 2, \ldots, r$, *so läßt sich der quadratische Annäherungsgrad durch*

$$\varrho_n \leq \sqrt{\frac{\pi W}{2}}\, \frac{(r-1)!\; 3^r M}{n^{r-\frac{1}{2}}}$$

abschätzen.

(Eingegangen am 18. April 1952.)

О ПОРЯДКЕ ВЕЛИЧИНЫ АППРОКСИМАЦИИ РЯДАМИ ОРТОГОНАЛЬНЫХ МНОГОЧЛЕНОВ

Г. АЛЕКСИЧ (Будапешт)

(Резюме)

Пусть $w(x) \geqq 0$ и $\{p_n(x)\}$ ортогональная нормированная система многочленов, вполне определенная на конечном интервале $a \leqq x \leqq b$ весовой функцией $w(x)$. Пусть $\sqrt{w}.f \in L^2$, и обозначим через $s_n(x)$ n-ую частную сумму разложения

$$f(x) \sim c_0 p_0(x) + c_1 p_1(x) + \ldots + c_n p_n(x) + \ldots$$

Кроме того, обозначим через $f^{(r)}(x)$ r-ую производную функции $f(x)$ $(f^{(0)} = f)$ а через $\omega_r(\delta)$ модуль непрерывности производной $f^{(r)}(x)$.

Если в интервале $\left[a + \frac{h}{2}, b + \frac{h}{2}\right]$, $w(x) \leqq W_h$, $|p_n(x)| \leqq P_h$, с постоянными W_h и P_h, то существуют такие явно задаваемые постоянные C_{hr}, K_{hr}, что для всякого индекса $n \geqq \operatorname{Max}\left(\frac{2-h}{2}, r+1\right)$ имеет место соотношение

$$\operatorname*{Max}_{a+h \leqq x \leqq b-h} |f(x) - s_n(x)| \leqq C_{hr}\, \omega_r\left(\frac{b-a}{2(n-r)}\right) \cdot \frac{\log n}{n^r} + K_{hr}\, \omega_r\left(\frac{b-a}{2(n-r)}\right) \cdot \frac{1}{n^r}.$$

ON THE STABILITY OF THE INDEX OF UNBOUNDED LINEAR TRANSFORMATIONS

By

BÉLA SZ.-NAGY (Szeged), corresponding member of the Academy

1. In a recent paper, KREĬN and KRASNOSEL'SKIJ[1] have considered linear transformations A with domain $\mathfrak{D}(A)$ in a Banach space E and with range $\mathfrak{R}(A)$ in a Banach space E', satisfying the following conditions:

a) A is a closed transformation and its range $\mathfrak{R}(A)$ is a closed linear manifold (i. e. a subspace of E');

b) The set of the elements $f \in E$ for which $Af = 0$ (the "null-space" of A), is of a finite number $\alpha(A)$ ($\geqq 0$) of dimensions;

c) The factor subspace $E'/\mathfrak{R}(A)$ is of a finite number $\beta(A)$ ($\geqq 0$) of dimensions.

Condition *c)* is equivalent to the following:

c)* Those continuous linear functionals in E' which vanish on $\mathfrak{R}(A)$ form a linear manifold of a finite number $\beta(A)$ of dimensions.

The difference

$$\nu(A) = \alpha(A) - \beta(A)$$

is called the *index* of A.

Transformations of this type, with the additional property of *boundedness*, were studied previously by ATKINSON[2] who called them *of generalized Fredholm type* (g. F. t.); let us call them likewise in the unbounded case too. The denomination is motivated by the fact that if K is a completely continuous linear transformation of a Banach space E into itself, and if I is the identical transformation of E, then it follows from the FREDHOLM—RIESZ theory[3] that $A = I - K$ satisfies the above conditions, moreover, $\nu(A) = 0$.

[1] М. Г. Крейн—М. А. Красносельский, Устойчивость индекса неограниченного оператора, Матем. сборник, **30 (72)** (1952), pp. 219—224.

[2] Ф. В. Аткинсон, Нормальная разрешимость линейных уравнений в нормированных пространствах, Матем. сборник, **28 (70)** (1951), pp. 3—14.

[3] See for instance F. RIESZ—B. SZ.-NAGY, *Leçons d'Analyse fonctionnelle* (Budapest, 1952), pp. 176—185, 214—216

ATKINSON proved,[2] for *bounded* transformations A of g. F. t., the following theorems:

THEOREM I. *There exists a positive constant ϱ (depending on A) such that, for every bounded linear transformation B of E into E' satisfying the inequality*

$$\|B\| \leqq \varrho, \tag{1}$$

the transformation $A+B$ is also of g. F. t. and $\nu(A+B) = \nu(A)$.

THEOREM II. *For any completely continuous linear transformation B of E into E', $A+B$ is also of g. F. t. and $\nu(A+B) = \nu(A)$.*

In their cited paper, KREĬN and KRASNOSEL'SKIJ have extended these theorems to the case of an *unbounded* A; their proofs are independent of ATKINSON's. They establish, however, theorem II only under some additional hypotheses: they suppose indeed either that B is of *finite rank* (i. e. such that $\mathfrak{R}(B)$ is of a finite number of dimensions), or that the space E' has a *base*.

The aim of this Note is to remark that these additional hypotheses are superfluous, and that, moreover, the case of unbounded transformations may be reduced in both theorems, by a simple device, to the case of bounded transformations, i. e. to the case considered by ATKINSON. It turns out that if A is unbounded, the class of the transformations B figuring in theorem I may be enlarged so as to include some unbounded transformations too.

2. Now we prove these assertions. Let A be a transformation of g. F. t. with $\mathfrak{D}(A) \subseteq E$ and $\mathfrak{R}(A) \subseteq E'$. Let us distinguish the norms in E and E' by writing $\| \ \|$ and $\| \ \|'$, respectively. For the elements f of $\mathfrak{D}(A)$ we introduce a second norm as follows:

$$|\mathbf{f}| = \|f\| + \|Af\|'.$$

The fact that A is *closed* implies that $\mathfrak{D}(A)$ becomes, by this definition of the norm, a *complete* normed linear space, i. e. a Banach space, which we shall denote by $\boldsymbol{E}$. We have then

$$\|Af\|' \leqq \|f\| + \|Af\|' = |\mathbf{f}|,$$

i. e. A is a *bounded* linear transformation of $\boldsymbol{E}$ into E'. Since

$$|\mathbf{f}| = 0 \quad \text{if and only if} \quad \|f\| = 0,$$

linearly independent elements of $\mathfrak{D}(A)$ are linearly independent also when considered in $\boldsymbol{E}$, and conversely. Thus the "null-space" of A in $\boldsymbol{E}$ is of the same dimension as in E, i. e. of dimension $\alpha(A)$. Since the introduction of the new norm does not alter at all the situation in E', we see that A, as a transformation of $\boldsymbol{E}$ in E', is *bounded*, of g. F. t., and has the same characteristic numbers $\alpha(A)$, $\beta(A)$.

Applying ATKINSON's theorem I, it results that there exists a positive number ϱ (which we may and shall choose < 1) such that, for each linear transformation B of $\boldsymbol{E}$ in E' satisfying the inequality

$$\|Bf\|' \leq \varrho \,|\!|\!| f |\!|\!| \qquad \text{for all } f \in \boldsymbol{E},$$

the transformation $A+B$ of $\boldsymbol{E}$ in E' is also of g. F. t. and has the same index as A. Now let B be a linear transformation of E into E' with domain $\mathfrak{D}(B) \supseteq \mathfrak{D}(A)$ and such that

$$\|Bf\|' \leq \varrho(\|f\| + \|Af\|') \qquad \text{for all } f \in \mathfrak{D}(A); \tag{2}$$

it generates a linear transformation of $\boldsymbol{E}$ into E' of the above type. It remains only to show that $A+B$, as a transformation of E in E', is also of g. F. t., and that it has the same index as when considered from $\boldsymbol{E}$ in E'. Since

$$\mathfrak{D}(A+B) = \mathfrak{D}(A) \cap \mathfrak{D}(B) = \mathfrak{D}(A),$$

there is no problem with the range of $A+B$, with $\beta(A+B)$, and with $\alpha(A+B)$, so that it remains only to show that $A+B$, as a transformation from E in E', is *closed*. Now we have

$$\|(A+B)f\|' \begin{cases} \leq \|Af\|' + \|Bf\|' \leq \|Af\|' + \varrho(\|f\| + \|Af\|'), \\ \geq \|Af\|' - \|Bf\|' \geq \|Af\|' - \varrho(\|f\| + \|Af\|'), \end{cases}$$

thus

$$(1-\varrho)\,\|Af\|' \leq \|(A+B)f\|' + \varrho\|f\| \leq (1+\varrho)\|Af\|' + 2\varrho\,\|f\|.$$

Since A is closed and $0 < \varrho < 1$, it results from these inequalities that $A+B$ is also closed. This concludes the proof of theorem I in the unbounded case, *for all B satisfying* (2).

Let now B be a completely continuous linear transformation of E in E'. The inequality $|\!|\!| f |\!|\!| \geqq \|f\|$ implies that each bounded set in $\boldsymbol{E}$ is bounded also when considered as a set in E; hence B is completely continuous also as a transformation of $\boldsymbol{E}$ in E'. Applying ATKINSON's theorem II, it results that $A+B$, as a transformation of $\boldsymbol{E}$ in E', is of g. F. t. and has the same index as A. One shows then as above that this remains true for $A+B$ also when considered as a transformation from E; the closure of $A+B$ is obvious because A is closed and B is continuous. This concludes the proof of theorem II in the unbounded case.

(Received 16 June 1952)

[4] To be exact, he has considered only the case when $E' = E$, but, as one easily checks, his proofs run also when $E' \neq E$.

ОБ УСТОЙЧИВОСТИ ИНДЕКСА НЕОГРАНИЧЕННЫХ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЙ

Б. С.-НАДЬ (Сегед)

(Резюме)

В своей статье [2] Аткинсон определил индекс $\nu(A)$ для некоторых ограниченных линейных преобразований A банахова пространства E в банахово пространство E', и показал, что $\nu(A+B)=\nu(A)$ (I) при всяком линейном преобразовании B с достаточно малым $\|B\|$, (II) при всяком вполне непрерывном линейном преобразовании B.

Крейн и Красносельский [1] дали другое доказательство этой же теоремы, справедливое и при неограниченных преобразованиях, а в случае (II) они предполагали, что пространство E' имеет базис (а если E' не имеет базиса, то B конечномерно).

В настоящей статье доказывается, что теоремы Аткинсона могут быть обобщёны просто и непосредственно для неограниченных преобразований, без предположения существовании в E' базиса.

A SPECTRAL PROBLEM FOR COMPLETELY CONTINUOUS OPERATORS

By

F. V. ATKINSON (Ibadan, Nigeria)

(*Presented by* B. Sz.-Nagy)

1. Let f and g denote elements of a vector-space, and V a linear operator on this space into itself; let λ denote a scalar parameter. The theory of the solubility of an equation of the form

$$f+\lambda Vf=g \tag{1. 1}$$

is in many cases characterised by a discrete spectrum, and by certain solubility conditions which are sometimes referred to as the "alternative", to wit the following two assertions:

(i) the corresponding homogeneous equation

$$f+\lambda Vf=0 \tag{1. 2}$$

and its transposed or adjoint version have both the same finite number (possibly zero) of linearly independent solutions,

(ii) the non-homogeneous equation (1. 1) is soluble for f if and only if g is orthogonal (in a suitable sense) to all solutions of the transposed homogeneous equation.

Familiar cases for which these results hold are the case of linear equations in a finite number of unknowns, the case of linear integral equations of Fredholm's type, and the case of systems of such equations. These are of course all special cases of the theory of completely continuous operators, as developed initially by F. Riesz, with subsequent contributions from Hildebrandt and Schauder.[1]

My object in this paper is to extend the theory to equations involving higher powers of λ, namely of the form

$$f+\sum_{r=1}^{n}\lambda^r V_r f=g, \tag{1. 3}$$

[1] The references are F. Riesz (*Acta Math.*, **41** (1918), pp. 71—98), T. H. Hildebrandt (*Acta Math.*, **51** (1928), pp. 311—318) and J. Schauder (*Studia Math.*, **2** (1930), pp. 183—196). A later account is given by S. Banach, *Théorie des opérations linéaires* (Warsaw, 1932), pp. 151—164.

where f, g are elements of a complex Banach space, and $V_1, \ldots, V_n$ are all completely continuous linear operators on the space into itself. Now so far as the "alternative" is concerned this introduces no novelty. Since a linear combination of completely continuous operators is itself completely continuous the validity of the "alternative" for (1. 3) is assured by the Riesz—Hildebrandt—Schauder theory. The task of this paper is therefore that of proving that (1. 3) has so to speak a discrete spectrum, that is to say that the set of values of λ for which

$$f+\sum_{r=1}^{n}\lambda^r V_r f=0 \tag{1.4}$$

has a non-trivial solution has no finite limit-point.

2. I allow myself some remarks by way of justification of this investigation.

In certain special cases the discreteness of the spectrum of (1. 3) can be proved for $n>1$ in much the same way as in the ordinary case $n=1$. For example, for finite-dimensional sets of linear equations such as

$$x_p+\sum_{r=1}^{n}\lambda^r\sum_{q=1}^{m}a_{pq}^{(r)}x_q=y_p \qquad (p=1,\ldots,m),$$

the "eigen-values" will be the roots of an algebraic equation. Again, for integral equations of the form

$$f(x)+\sum_{r=1}^{n}\lambda^r\int_a^b K_r(x,y)f(y)\,dy=g(x),$$

wherein the kernels are suitably continuous, the "eigen-values" will be the zeros of an integral function constructed in FREDHOLM's manner. It might thus be expected that the argument used by F. RIESZ to prove the discreteness of the spectrum of (1. 1) would extend to the case of (1. 3), but I have been unable make this extension. The proof which I give here proceeds on quite different lines, and may have interest whether RIESZ's argument can be extended or not.

The proof which I give here differs in an important logical respect to that given by RIESZ for the case $n=1$. In considering the equation of the form (1. 1) the fact of V being completely continuous was used by RIESZ at a number of points, including also in the proof of the discreteness of the spectrum. In the present argument the complete continuity of $V_1, \ldots, V_n$ is used only in that it establishes the validity of the "alternative" for (1. 3). It may be that the validity of the "alternative" is a weaker assumption than that of complete continuity.

3. I next introduce certains notations. Let $\mathfrak{R}$ denote the complex Banach space in question, and $\mathfrak{R}^*$ the adjoint space. It will incidentally not be assumed that $\mathfrak{R}$ is either separable or reflexive. Let now T denote some bounded linear operator, which maps R onto the whole or part of itself. Following the notations of a previous paper,[2] I introduce in association with T certain linear manifolds and integer-valued functionals, as follows:

(*a*) $\mathfrak{A}(T)$ is to denote the linear manifold of solutions of the homogeneous equation

$$Tf = 0, \quad f \in \mathfrak{R}, \tag{3. 1}$$

(*b*) $\mathfrak{B}(T)$ is to denote the linear manifold of solutions of the transposed homogeneous equation, which I write as[3]

$$l\,T = 0, \quad l \in \mathfrak{R}^*. \tag{3. 2}$$

In cases with which this paper is concerned (3. 1—2) will each have at most a finite number of linearly independent solutions, and I then define

(*c*) $\alpha(T)$, $\beta(T)$ as the dimension-numbers of the linear manifolds $\mathfrak{A}(T)$, $\mathfrak{B}(T)$. These numbers will then be finite positive integers or zero, and will moreover both be equal in cases considered here.[4]

Now let us write for brevity

$$T(\lambda) = I + \sum_{r=1}^{n} \lambda^r V_r,$$

where I denotes the identity operator on $\mathfrak{R}$, and $V_1, \ldots, V_n$ are bounded linear completely continuous operators on $\mathfrak{R}$ into $\mathfrak{R}$. Then for any finite λ, real or complex, $T(\lambda)$ has the form $I+V$, where V is completely continuous. It therefore follows in virtue of the Riesz—Hildebrandt—Schauder theory that the numbers

$$\alpha(T(\lambda)), \quad \beta(T(\lambda))$$

both exist, both being equal for any particular λ-value either to zero or to some positive integer. What we have to prove is that they are both zero for all λ with the possible exception of a set of λ-values with no finite limit-point.

4. In this section I outline the construction of certain projection operators and a generalised inverse associated with operators of the form considered here.[5]

[2] Ф. В. Аткинсон, Нормальная разрешимость линейных уравнений в нормированных пространствах, Матем. сборник, **28** (**70**) (1951), pp. 3—14.

[3] I depart here from Banach's notation, in which this equation would have been written $\overline{T}(l) = 0$.

[4] Cases in which they are not necessarily equal formed the subject of my above-cited paper.

[5] Similar procedures have been employed by С. М. Никольский, Известия А. Н. СССР, сер. матем., **7**, No. 3 (1943), pp. 147—166, and by myself in my paper mentioned above. See also S. Banach, loc. cit.

Let λ_0 be any selected λ-value, and write for brevity T_0 for $T(\lambda_0)$, and similarly $\mathfrak{A}_0, \mathfrak{B}_0, \alpha_0$ and β_0 in place of $\mathfrak{A}(T_0), \mathfrak{B}(T_0), \alpha(T_0)$ and $\beta(T_0)$. We need first to construct two projection operators P_1, P_2 such that

$$P_1\mathfrak{R} = \mathfrak{A}_0, \quad \mathfrak{R}^* P_2 = \mathfrak{B}_0.$$

The construction of a projection operator whose range is to be a given closed linear manifold is of course more difficult in a Banach space than in a Hilbert space, but is nevertheless certainly possible when the manifold is finite-dimensional. For example for P_1 let g_r $(r=1, 2, \ldots, \alpha_0)$, $g_r \in \mathfrak{R}$, be linearly independent and form a basis of $\mathfrak{A}_0$. We then define a set of bounded linear functionals l_r $(r=1, 2, \ldots, \alpha_0)$, $l_r \in \mathfrak{R}^*$. These are defined on $\mathfrak{A}_0$ by the requirements

$$l_r(g_s) = \begin{cases} 1 & (r=s), \\ 0 & (r \neq s), \end{cases}$$

and may be extended to the whole of $\mathfrak{R}$. We then define

$$P_1 f = \sum_{r=1}^{\alpha_0} l_r(f) g_r \quad \text{for all } f \in \mathfrak{R}.$$

Then P_1 has the required properties

$$P_1\mathfrak{R} = \mathfrak{A}_0, \quad P_1^2 = P_1.$$

The construction of P_2 is entirely analogous. Since P_1 and P_2 are projections into $\mathfrak{A}_0, \mathfrak{B}_0$ we have the equations

$$T_0 P_1 = 0, \quad P_2 T_0 = 0.$$

I now assert that the operator T_0 effects a one-to-one continuous mapping of the space $(I-P_1)\mathfrak{R}$ onto the space $(I-P_2)\mathfrak{R}$. The continuity of the mapping follows of course from the boundedness of T_0. That the mapping is onto $(I-P_2)\mathfrak{R}$ follows from the fact that if

$$g = T_0 f, \qquad g, f \in \mathfrak{R},$$

then

$$P_2 g = P_2 T_0 f = 0, \quad g = (I-P_2) g \in (I-P_2)\mathfrak{R}.$$

As regards the mapping being one-to-one, suppose that

$$T_0 f = T_0 f', \qquad f, f' \in (I-P_1)\mathfrak{R}.$$

It then follows that $T_0(f-f')=0$, so that we have simultaneously

$$(f-f') \in P_1\mathfrak{R}, \quad (f-f') \in (I-P_1)\mathfrak{R},$$

and since P_1 is a projection operator, this implies that $f-f'=0$, so that the mapping is indeed one-to-one in so far as no two distinct elements of $(I-P_1)\mathfrak{R}$ can be mapped into the same element of $(I-P_2)\mathfrak{R}$. It has also to be shown that $(I-P_1)\mathfrak{R}$ is mapped into the whole of $(I-P_2)\mathfrak{R}$, that is to say that to each $g \in (I-P_2)\mathfrak{R}$ there is an $f \in (I-P_1)\mathfrak{R}$ such that $T_0 f = g$.

Now if $g \in (I-P_2)\mathfrak{N}$, then $P_2 g = 0$, so that $lg = 0$ for all $l \in \mathfrak{B}_0$. This is however precisely the condition for the solubility of the equation $T_0 f = g$, according to the Riesz—Hildebrandt—Schauder theory. Here it will be possible to take $f \in (I-P_1)\mathfrak{N}$, replacing f if necessary by $(I-P_1)f$. This completes the proof.

We now deduce that the inverse linear mapping of $(I-P_2)\mathfrak{N}$ onto $(I-P_1)\mathfrak{N}$ is also one-to-one and continuous. We denote this bounded linear operator by U_0. So far U_0 is defined only on the domain $(I-P_2)\mathfrak{N}$; we extend its domain to the whole of $\mathfrak{N}$ by specifying that $U_0 P_2 = 0$. Since the range of U_0 is $(I-P_1)\mathfrak{N}$ we have also $P_1 U_0 = 0$. The operator U_0 is in fact a generalised inverse of T_0. We have the further equations

$$U_0 T_0 = I - P_1, \quad T_0 U_0 = I - P_2. \tag{4.1}$$

5. I now initiate the study of the behaviour of the functions $\alpha(T(\lambda))$, $\beta(T(\lambda))$ in a neighbourhood of λ_0. Let λ be some value of the parameter, not equal to λ_0, which we shall in fact restrict to λ a neighbourhood of λ_0, such as a circle, centre λ_0. I write for brevity

$$T(\lambda) = T(\lambda_0) - A(\lambda) = T_0 - A,$$

so that $A = A(\lambda)$ is a polynomial in λ, vanishing for $\lambda = \lambda_0$. We have thus to consider the behaviour of $\alpha(T_0 - A)$, $\beta(T_0 - A)$ in a neighbourhood of λ_0.

Let us assume that the operator $(I - U_0 A)$ has a bounded inverse. This will certainly be the case if $\|A\|$ is sufficiently small, which may be achieved by making $|\lambda - \lambda_0|$ sufficiently small. With this assumption we have, using (4. 1),

$$(I - U_0 A)^{-1} U_0 (T_0 - A) = (I - U_0 A)^{-1} (U_0 T_0 - U_0 A) =$$
$$= (I - U_0 A)^{-1} (I - P_1 - U_0 A) = I - (I - U_0 A)^{-1} P_1.$$

From this equation it follows that

$$(T_0 - A) f = 0, \quad f \in \mathfrak{N},$$

implies

$$f = (I - U_0 A)^{-1} P_1 f,$$

or, in other terminology, we have

$$\mathfrak{A}(T_0 - A) \subseteq (I - U_0 A)^{-1} P_1 \mathfrak{N}. \tag{5.1}$$

We note here two features of the operator $(I - U_0 A)^{-1} P_1$, firstly that it is of the same dimensionality as P_1, i. e. $\alpha(T_0) = \alpha_0$, and secondly that it is a projection operator. The latter property follows from the equation $P_1 U_0 = 0$. From the former property and from (5. 1) we see incidentally that

$$\alpha(T_0 - A) \leqq \alpha(T_0). \tag{5.2}$$

6. We need however a more precise analysis of (5. 1). Let us write

$$k = k(\lambda) = \alpha(T_0) - \alpha(T_0 - A),$$

and select two sets of elements of $\mathfrak{R}$,

$$x_q \quad (q=1,\ldots,\alpha(T_0-A)), \qquad y_r \quad (r=1,\ldots,k),$$

in such a way that the first set forms a basis of $\mathfrak{A}(T_0-A)$ and the two sets together form a basis of $(I-U_0A)^{-1}P_1\mathfrak{R}$. It then follows that the first set of elements will be annulled by the operator (T_0-A), and that no linear combination formed from the second set will be annulled by this operator unless it vanishes identically. It then follows that the elements

$$(T_0-A)y_r \qquad (r=1,\ldots,k)$$

will be linearly independent and will form a basis of the linear manifold

$$(T_0-A)(I-U_0A)^{-1}P_1\mathfrak{R},$$

which will therefore be of dimensionality $k=k(\lambda)$.

Summing up so far we have, if $|\lambda-\lambda_0|$ is sufficiently small,

$$\alpha(T_0-A)=\beta(T_0-A)=\alpha(T_0)-k(\lambda)=\beta(T_0)-k(\lambda),$$

where

$$k(\lambda)=\dim\{(T_0-A)(I-U_0A)^{-1}P_1\}.$$

7. As the next and penultimate stage in the proof I show that $k(\lambda)$ is constant in a neighbourhood of λ_0, with the possible exception of λ_0 itself.

Let us write

$$F(\lambda)=(T_0-A)(I-U_0A)^{-1}P_1,$$

so that, in a certain neighbourhood of λ_0, $F(\lambda)$ will be an analytic function of λ, with values in the ring of bounded operators. Let λ_1 be a λ-value in such a neighbourhood of λ_0. Then $F(\lambda_1)$ will be of dimensionality $k(\lambda_1)$, and it will therefore be possible to find linearly independent elements

$$z_r \quad (r=1,\ldots,k(\lambda_1)), \qquad z_r\in\mathfrak{R},$$

such that the linearly independent elements

$$F(\lambda_1)z_r \qquad (r=1,\ldots,k(\lambda_1))$$

constitute a basis of $F(\lambda_1)\mathfrak{R}$. We then determine functionals

$$m_r\in\mathfrak{R}^* \qquad (r=1,\ldots,k(\lambda_1)),$$

such that

$$m_r(F(\lambda_1)z_s)=\begin{cases}1 & (r=s),\\ 0 & (r\neq s).\end{cases}$$

Now let us consider, as a function of λ, the determinant, of order $k(\lambda_1)$,

$$G(\lambda)=|m_r(F(\lambda)z_s)| \qquad (r,s=1,\ldots,k(\lambda_1)).$$

According to our construction we have $G(\lambda_1)=1$. Also, since

$$F(\lambda_0)=T_0P_1=0$$

we have $G(\lambda_0)=0$. Furthermore, $G(\lambda)$ is an analytic function, in the ordinary sense of the term, at any rate in a certain neighbourhood of λ_0 including the point λ_1. It then follows that there exists a positive number ϱ such that $G(\lambda) \neq 0$ for all λ in the region $0 < |\lambda - \lambda_0| < \varrho$.

Since now $G(\lambda) \neq 0$ in this region it follows that in this region the elements

$$F(\lambda) z_r \qquad (r = 1, \ldots, k(\lambda_1))$$

are linearly independent, so that in the same region the dimensionality of $F(\lambda)$ is at least $k(\lambda_1)$. Thus for $0 < |\lambda - \lambda_0| < \varrho$ we have $k(\lambda) \geqq k(\lambda_1)$.

There are then two possibilities. Either we have $k(\lambda) = k(\lambda_1)$ throughout the region $0 < |\lambda - \lambda_0| < \varrho$, or else there is in this region a value λ_2 such that $k(\lambda_2) > k(\lambda_1)$. However in this latter case we have merely to repeat the argument, which would then show that there was a positive number ϱ' such that $k(\lambda) \geqq k(\lambda_2)$ for all λ with $0 < |\lambda - \lambda_0| < \varrho'$. But the process cannot be repeated indefinitely, since $k(\lambda_1), k(\lambda_2), \ldots$ would form an increasing sequence of integers, whose members cannot exceed $\alpha(T_0)$, by the definition of $k(\lambda)$. Hence finally we must arrive at a positive number ϱ^*, and a non-negative integer k^*, such that $k(\lambda) = k^*$ for $0 < |\lambda - \lambda_0| < \varrho^*$.

This proves the assertion that $k(\lambda)$ is constant in a certain neighbourhood of λ_0, excluding possibly the value λ_0 itself. It follows that $\alpha(T(\lambda))$ and $\beta(T(\lambda))$ are also constant in such a region.

8. I now pass to the main assertion of this paper. Let $\mathfrak{C}$ denote the resolvent set of $T(\lambda)$, that is to say the set of λ-values for which $(T(\lambda))^{-1}$ exists as a bounded operator. This set will be identical with the set of λ-values for which the functions $\alpha(T(\lambda))$, $\beta(T(\lambda))$ vanish, and is an open non-empty set including at the least a certain neighbourhood of $\lambda = 0$. If $\mathfrak{C}$ does not consist of the entire λ-plane, let μ denote a λ-value lying in the boundary of $\mathfrak{C}$, that is to say in the closure $\bar{\mathfrak{C}}$ of $\mathfrak{C}$ but not in $\mathfrak{C}$ itself. There will then be a sequence of λ-values tending to μ for which $\alpha(T(\lambda)) = \beta(T(\lambda)) = 0$. But by what has been proved, $\alpha(T(\lambda))$ and $\beta(T(\lambda))$ have constant values near μ in a region of the form $0 < |\lambda - \mu| < \varrho$, for some positive ϱ. Hence there is a region of the form $0 < |\lambda - \mu| < \varrho$ in which $\alpha(T(\lambda)) = \beta(T(\lambda)) = 0$, so that this region consists entirely of points of $\mathfrak{C}$. This proves that the boundary of $\mathfrak{C}$ consists at most of isolated points, so that the points which form the boundary of $\mathfrak{C}$ have no finite limit-point. From this it is easily proved that $\mathfrak{C}$ covers the entire λ-plane, with the possible exception of an enumerable set with no finite limit-point. This was the result to be proved.

UNIVERSITY COLLEGE,
IBADAN, NIGERIA

(Received 15 July 1951, in a revised form 14 March 1952)

Note added in proof. (8 September 1952.) Since the above paper was originally drafted a number of closely-related papers have appeared. I. C. GOHBERG (И. Ц. Гохберг) *(Doklady Akad. Nauk SSSR,* **76** (1951), pp. 9—12, *ibid.,* **76** (1951), pp. 477—480, *ibid.,* **78** (1951), pp. 629—632, *Uspehi Mat. Nauk,* **7** (48) (1952), pp. 149—156) has investigated similar problems by different methods (see also the above-cited paper of the present author), while M. G. KREĬN and M. A. KRASNOSEL'SKIJ *(Mat. Sbornik,* **30** (72) (1952), pp. 219—224) have given extensions of some of the present author's previous results to the case of non-bounded operators.

СПЕКТРАЛЬНАЯ ПРОБЛЕМА С ВПОЛНЕ НЕПРЕРЫВНЫМ ОПЕРАТОРОМ

Ф. В. АТКИНСОН (Ибадан, Нигерия)

(Резюме)

Пусть I единичный оператор, и $V_1, \ldots, V_n$ — вполне непрерывные линейные операторы на пространстве типа Банаха. Автор предлагает доказательство того, что собственные числа многочленного оператора $\left(I + \sum_{r=1}^{n} \lambda^r V_r\right)$ суть счетное множество без конечной предельной точки. Для частного случая $n = 1$ это утверждение — известный результат Ф. Рисс'а. Однако метод Рисс'а, повидимому, в своем непосредственном виде неприменим в общем случаем $n > 1$ и предлагаемый метод пользуется совершенно отличными способами.

ON A SPECTRAL PROBLEM OF ATKINSON

By

BÉLA SZ.-NAGY (Szeged), corresponding member of the Academy

§ 1

In the same issue of these *Acta,* ATKINSON[1] has considered polynomials of the complex variable

(1) $$T(\lambda)=I+\lambda V_1+\lambda^2 V_2+\ldots+\lambda^p V_p,$$

whose coefficients V_i are completely continuous linear transformations of a Banach space $\mathfrak{R}$ into itself; I denotes the identical transformation of $\mathfrak{R}$. Calling λ_0 a *zero* of $T(\lambda)$ if the equation

$$T(\lambda_0)f=0$$

has a non-zero solution $f\in\mathfrak{R}$, ATKINSON proved the theorem that *the zeros of the polynomial* $T(\lambda)$ *have no finite limit-point.*

In the case $p=1$ this is a well-known fact proved in this general situation by F. RIESZ [see [2] or [3]]; RIESZ' method, however, does not seem to apply for $p>1$. This is not surprising as even for ordinary polynomials, to which our problem reduces if $\mathfrak{R}$ is of dimension 1, the problem of the zeros lies much deeper for general p's than for $p=1$.

In this paper we give another proof of ATKINSON's theorem, which, if not shorter, seems in some respect simpler than the original one. The main difference between the two methods is that we use spectral manifolds while ATKINSON uses eigen-manifolds, and that we apply the convenient tool of contour integration.

Before going into the proof we recall, for sake of convenient reference, some known facts and definitions:

For an arbitrary bounded (or at least closed) linear transformation T of the Banach space $\mathfrak{R}$ into itself, the *resolvent set* $\varrho(T)$ is defined as the set

[1] F. V. ATKINSON, A spectral problem for completely continuous operators, *Acta Math. Acad. Sci. Hungaricae,* **3** (1952), pp. 53—60.

[2] F. RIESZ, Über lineare Funktionalgleichungen, *Acta Math.,* **41** (1917), pp. 71—98.

[3] F. RIESZ—B. SZ.-NAGY, *Leçons d'Analyse fonctionnelle* (Budapest, 1952), nº 79 and 89. This book will be cited in the sequel as R.—Sz.-N.

of those points z of the plane of complex numbers, for which

$$(T-zI)^{-1}$$

exists as a bounded and everywhere defined transformation of $\mathfrak{R}$. The complementary set $\sigma(T)$ is called the *spectrum* of T.

(i) If C is a simple closed and rectifiable curve in the complex plane, passing within $\varrho(T)$, then

$$P=-\frac{1}{2\pi i}\oint_C (T-zI)^{-1}\,dz$$

exists (as a limit in norm of sums of Cauchy—Riemann type), is bounded and idempotent: $P^2=P$. The space $\mathfrak{R}$ splits off into the vectorial sum of its disjoint subspaces

$$\mathfrak{M}=\mathfrak{M}(T;C)=P\mathfrak{R},\quad \mathfrak{N}=\mathfrak{N}(T;C)=(I-P)\mathfrak{R}:$$

P is the *projection of* $\mathfrak{R}$ *on* $\mathfrak{M}$ *parallel to* $\mathfrak{N}$. $\mathfrak{M}$ and $\mathfrak{N}$ both reduce T; denoting by T' and T'' the transformations induced by T in $\mathfrak{M}$ and in $\mathfrak{N}$ respectively, $\sigma(T')$ coincides with that part of $\sigma(T)$ which lies inside C, and $\sigma(T'')$ coincides with that part of $\sigma(T)$ which lies outside C.[4]

(ii) If, in addition, T is completely continuous, then $\sigma(T)$ has no limit-point $\neq 0$. If the point 0 lies outside the curve C, then the subspace $\mathfrak{M}(T;C)$ is of finite dimension.[5]

(iii) If $\varphi_1, \varphi_2, \ldots, \varphi_r$ are linearly independent elements of the Banach space $\mathfrak{R}$, then there exist elements $\Phi_1, \Phi_2, \ldots, \Phi_r$ of the conjugate space $\mathfrak{R}^*$ such that the φ_i and Φ_k form a biorthogonal set, i. e.

$$(\varphi_i, \Phi_k)=\delta_{ik}=\begin{cases} 1 & \text{if } i=k, \\ 0 & \text{if } i\neq k.^6 \end{cases}$$

§ 2

Our proof will be based on the following

LEMMA. *Let* $S_0, S_1, S_2, \ldots$ *be bounded linear transformations of a Banach space* $\mathfrak{R}$ *into itself, suppose that* S_0 *is completely continuous and that there is a constant* a *such that*

$$\|S_n\|\leq a^n \qquad (n=1,2,\ldots). \tag{2}$$

Then any point $\mu\neq 0$ *of the plane of complex numbers has a neighborhood such that, for sufficiently small values of the complex parameter* ε, *the spectrum of the transformation*

$$S(\varepsilon)=S_0+\varepsilon S_1+\varepsilon^2 S_2+\cdots+\varepsilon^n S_n+\cdots \tag{3}$$

[4] Cf. R.—Sz.-N., n° 147.

[5] Cf. R.—Sz.-N., p. 411.

[6] This follows from the lemma on R.—Sz.-N., p. 213.

consists, in this neighborhood of μ, of the characteristic roots of some finite square matrix whose elements $c_{ik}(\varepsilon)$ are regular analytic functions of ε.[7]

PROOF. Observe first of all that (2) implies the convergence in norm of the series (3) at least for $|\varepsilon| < a^{-1}$.

Since S_0 is completely continuous, μ is not a limit-point of $\sigma(S_0)$, thus we can trace in the complex plane a sufficiently small circle C with μ as centre, which passes entirely in $\varrho(S_0)$ and leaves the point 0 outside. The corresponding spectral manifold

$$\mathfrak{M}_0 = \mathfrak{M}(S_0; C)$$

is, according to (ii), of finite dimension, of dimension r say ($r \geqq 0$).

We show that, for sufficiently small ε, C is contained also in $\varrho(S(\varepsilon))$ and that the corresponding spectral manifold

$$\mathfrak{M}(\varepsilon) = \mathfrak{M}(S(\varepsilon); C)$$

is of the same dimension r.

Since

$$R_z = (S_0 - zI)^{-1}$$

is a continuous function (in the norm topology) of the variable z on C, there is a constant $M > 0$ such that

$$\|R_z\| \leq M \quad \text{for} \quad z \in C.$$

Then we have, uniformly in $z \in C$,

$$\|(S(\varepsilon) - S_0)R_z\| = \left\| \sum_{n=1}^{\infty} \varepsilon^n S_n R_z \right\| \leq \sum_{n=1}^{\infty} |\varepsilon|^n a^n M < 1$$

for sufficiently small ε. Then $[I + (S(\varepsilon) - S_0)R_z]^{-1}$ exists,[8] and since

$$S(\varepsilon) - zI = [I + (S(\varepsilon) - S_0)R_z]\ (S_0 - zI),$$

we have

$$R_z(\varepsilon) = [S(\varepsilon) - zI]^{-1} = R_z[I + (S(\varepsilon) - S_0)R_z]^{-1} =$$

$$= R_z \sum_{\nu=0}^{\infty} \Big[-(S(\varepsilon) - S_0)R_z\Big]^{\nu} = R_z \sum_{\nu=0}^{\infty} \Big[-\sum_{n=1}^{\infty} \varepsilon^n S_n R_z\Big]^{\nu}.$$

An easy calculation with majorising functions, obtained by using the inequalities $\|R_z\| \leq M$ and $\|S_n\| \leq a^n$, shows that, for sufficiently small ε, it is legitimate to rearrange the last expansion in a power series

$$R_z(\varepsilon) = R_{z,0} + \varepsilon R_{z,1} + \ldots + \varepsilon^n R_{z,n} + \ldots \tag{4}$$

[7] This would follow from some more general results of the author, cf. BÉLA SZ.-NAGY, Perturbations des transformations linéaires fermées, *Acta Sci. Math. Szeged,* **14** (1951), pp. 125—137; but we prefer to give here a complete proof.

[8] Here, and always in the sequel, if we say that an inverse transformation *exists*, we mean also that it is bounded and everywhere defined in $\mathfrak{R}$.

with $R_{z,0}=R_z$, and that we have

$$\|R_{z,n}\| \leq \frac{M^2}{1+M}[a(1+M)]^n \qquad (n=1,2,\ldots). \tag{5}$$

The series (4) being, for sufficiently small values of ε, uniformly convergent with respect to $z\in C$, it may be integrated term by term, and so it results that

$$P(\varepsilon)=-\frac{1}{2\pi i}\oint_C R_z(\varepsilon)\,dz=P_0+\varepsilon P_1+\ldots+\varepsilon^n P_n+\ldots \tag{6}$$

with

$$\|P_n\|=\left\|-\frac{1}{2\pi i}\oint_C R_{z,n}\,dz\right\| \leqq \frac{|C|}{2\pi}\frac{M^2}{1+M}[a(1+M)]^n \qquad (n=1,2,\ldots) \tag{7}$$

where $|C|$ denotes the perimeter of C.

For sufficiently small ε this implies the inequality

$$\|P(\varepsilon)-P_0\|<1$$

which guarantees that the transformations

$$A(\varepsilon)=I-(P(\varepsilon)-P_0) \quad\text{and}\quad B(\varepsilon)=I+(P(\varepsilon)-P_0)$$

can be boundedly inverted. Thus, in particular, $A(\varepsilon)$ maps $\mathfrak{R}$ on the *whole* of $\mathfrak{R}$, and, consequently, $P(\varepsilon)A(\varepsilon)$ maps $\mathfrak{R}$ on the *whole* of $P(\varepsilon)\mathfrak{R}$. But since $P^2(\varepsilon)=P(\varepsilon)$ [see (i)], we have $P(\varepsilon)A(\varepsilon)=P(\varepsilon)P_0$, and since

$$P_0\mathfrak{R}=P(0)\mathfrak{R}=\mathfrak{M}(S_0;C), \quad P(\varepsilon)\mathfrak{R}=\mathfrak{M}(S(\varepsilon);C),$$

it results that

$$P(\varepsilon)\mathfrak{M}_0=\mathfrak{M}(\varepsilon).$$

Moreover, this mapping of $\mathfrak{M}_0$ on $\mathfrak{M}(\varepsilon)$ is biunivoque and bicontinuous, for if $f\in\mathfrak{M}_0$ is carried into $g\in\mathfrak{M}(\varepsilon)$, we have

$$B(\varepsilon)f=f+P(\varepsilon)f-P_0f=f+g-f=g,$$

thus, putting $C(\varepsilon)=[B(\varepsilon)]^{-1}$, we have $f=C(\varepsilon)g$.

It results from all these that if $\varphi_1,\varphi_2,\ldots,\varphi_r$ are linearly independent elements forming a base of the r-dimensional subspace $\mathfrak{M}_0$, then the elements

$$\varphi_i(\varepsilon)=P(\varepsilon)\varphi_i \qquad (i=1,2,\ldots,r) \tag{8}$$

are also linearly independent and form a base of $\mathfrak{M}(\varepsilon)$; moreover, we have

$$\varphi_i=C(\varepsilon)\varphi_i(\varepsilon). \tag{9}$$

Thus we have proved not only that $\mathfrak{M}_0$ and $\mathfrak{M}(\varepsilon)$ have the same dimension, but also that one can pass from a base of $\mathfrak{M}_0$ to a base of $\mathfrak{M}(\varepsilon)$, and back again, by transformations which are power series of ε; indeed, for $P(\varepsilon)$ we have the power series expansion (6), and for $C(\varepsilon)$ we have

$$C(\varepsilon)=\sum_{\nu=0}^{\infty}[-(P(\varepsilon)-P_0)]^\nu=\sum_{\nu=0}^{\infty}\left[-\sum_{k=1}^{\infty}\varepsilon^k P_k\right]^\nu;$$

using the inequalities (7), an easy calculation with majorising functions shows that, for sufficiently small ε, it is legitimate to rearrange the last expression in a power series:

$$C_0+\varepsilon C_1+\cdots+\varepsilon^n C_n+\cdots.$$

Since $S(\varepsilon)$ transforms $\mathfrak{M}(\varepsilon)$ into itself [see (i)], we have

$$S(\varepsilon)\varphi_k(\varepsilon)=\sum_{i=1}^{r} c_{ik}(\varepsilon)\varphi_i(\varepsilon) \qquad (k=1,2,\ldots,r) \tag{10}$$

with numerical coefficients depending on ε. They can be determined as follows. Applying $C(\varepsilon)$ on both sides of (10) we get, by virtue of (8) and (9),

$$C(\varepsilon)S(\varepsilon)P(\varepsilon)\varphi_k=\sum_{i=1}^{r} c_{ik}(\varepsilon)\varphi_i;$$

if $\{\Phi_k\}$ is a set of elements of $\mathfrak{N}^*$ forming with $\{\varphi_k\}$ a biorthogonal set (see (iii)], then we obtain that

$$(C(\varepsilon)S(\varepsilon)P(\varepsilon)\varphi_k,\Phi_i)=c_{ik}(\varepsilon) \qquad (i,k=1,2,\ldots,r).$$

$C(\varepsilon)$, $S(\varepsilon)$ and $P(\varepsilon)$ being, for sufficiently small ε, power series of ε, so is their product (this results again by using majorising functions). Hence the $c_{ik}(\varepsilon)$ are also power series of ε.

Now that part of the spectrum of $S(\varepsilon)$ which lies inside the circle C, coincides with the spectrum of $S(\varepsilon)$ as a transformation of $\mathfrak{M}(\varepsilon)$ [see (i)], thus it consists of the characteristic roots of the matrix $c_{ik}(\varepsilon)$. This concludes the proof of the lemma.

§ 3

Now let us return to our original problem. Denoting by Λ the set of the zeros of $T(\lambda)$, and by Λ' the set of the finite limit-points of Λ, we have to prove that Λ' is void. It suffices to show to this end that Λ' is at the same time *closed* and *open,* and that it *does not fill out the whole complex plane.* The closure is obvious, for Λ' is a derived set. The third assertion follows from the fact that, for sufficiently small λ,

$$\|T(\lambda)-I\|=\|\lambda V_1+\lambda^2 V_2+\ldots+\lambda^p V_p\|<1,$$

so that $[T(\varepsilon)]^{-1}$ exists; thus Λ and consequently Λ' do not have points in some circle with centre $\lambda=0$. It remains only to prove that Λ' is open, and this may be done as follows.

We show that if Λ' contains a point λ_0, it contains some neighborhood of λ_0 too. We have certainly $\lambda_0\neq 0$. Let $\lambda_1,\lambda_2,\ldots$ be a sequence of zeros of $T(\lambda)$, all different from λ_0 and tending to λ_0. Put

$$S(\varepsilon)=V_1+(\lambda_0+\varepsilon)V_2+\ldots+(\lambda_0+\varepsilon)^{p-1}V_p=S_0+\varepsilon S_1+\ldots+\varepsilon^{p-1}S_{p-1},$$

then

$$T(\lambda_0+\varepsilon)=I+(\lambda_0+\varepsilon)S(\varepsilon). \tag{11}$$

Since, for $n=1,2,\ldots$, λ_n is a zero of $T(\lambda)$, it follows that $-\frac{1}{\lambda_n}$ is an eigen-value of $S(\lambda_n-\lambda_0)$, i. e., putting

$$\varepsilon_n=\lambda_n-\lambda_0,$$

$-\frac{1}{\lambda_0+\varepsilon_n}$ is an eigen-value of $S(\varepsilon_n)$. We have $\varepsilon_n\neq 0$ and $\varepsilon_n\to 0$.

But, by our Lemma, the spectrum of $S(\varepsilon)$, in the neighborhood of $\mu=-\frac{1}{\lambda_0}$, consists, for all sufficiently small values of ε, of the characteristic roots of some finite square matrix whose elements $c_{ik}(\varepsilon)$ are regular analytic functions of ε. The determinant

$$\det\left(c_{ik}(\varepsilon)+\frac{1}{\lambda_0+\varepsilon}\delta_{ik}\right)$$

is then also a regular analytic function of ε in a neighborhood of $\varepsilon=0$, and since it vanishes in (almost all) points ε_n, it vanishes identically. This implies that, for *all* sufficiently small values of ε, for $|\varepsilon|<\eta$ say, $S(\varepsilon)$ has the eigen-value $-\frac{1}{\lambda_0+\varepsilon}$, hence, by (11), $\lambda_0+\varepsilon$ is a zero of $T(\lambda)$.

Thus if Λ' contains the point λ_0, then a whole neighborhood of λ_0 belongs to Λ and hence to Λ'. This proves that Λ' is an open set and thus concludes the proof of the theorem.

BOLYAI INSTITUTE,
SZEGED UNIVERSITY.

(Received 16 June 1952)

ОБ ОДНОЙ СПЕКТРАЛЬНОЙ ПРОБЛЕМЕ АТКИНСОНА

Б. С.-НАДЬ (Сегед)

(Р е з ю м е)

Пусть $T(\lambda)=I+\lambda V_1+\ldots+\lambda^p V_p$ многочлен комплексного переменного λ, коэффициенты V_i которого — вполне непрерывные преобразования банахова пространства $\mathfrak{R}$ в само себя, а I — единичное преобразование на $\mathfrak{R}$. Пусть λ_0 называется нулём многочлена $T(\lambda)$, если равенство

$$T(\lambda_0)f=0$$

имеет нетривиальное решение $f\in\mathfrak{R}$. А т к и н с о н доказал, что множество всех нулей не имеет конечной предельной точки.

В настоящей статье дается алтернативное, более простое доказательство того же утверждения. Это доказательство отличается от оригинального особенно в том, что здесь используются спектральые — а не собственные — подпространства, и что проекции на эти подпространства представляются интегралами в комплексной плоскости.

INHALTSABSCHÄTZUNG EINES SPHÄRISCHEN POLYGONS

Von

J. MOLNÁR (Veszprém)

(*Vorgelegt von* G. HAJÓS)

HABICHT und VAN DER WAERDEN haben in einer Arbeit[1] folgenden Satz bewiesen:

Bedeutet T_n ein sphärisches n-Eck, dessen jede Seite gleich a ist und in welchem der Abstand je zweier Eckpunkte $\geqq a$ ausfällt, so ist[2]

$$T_n \geqq (n-2)\, T_3. \tag{1}$$

Aus diesem Satz folgt in einfacher Weise eine Abschätzungsformel von FEJES TÓTH:[3]

$$d \leqq \arc\cos \frac{\operatorname{cotg}^2 \omega - 1}{2}, \quad \omega = \frac{n}{n-2} \frac{\pi}{6},$$

die sich auf den sphärischen Mindestabstand d von n Punkten der Einheitskugel bezieht. Eine weitere Anwendung der Ungleichung (1) findet in einer Arbeit von SCHÜTTE und VAN DER WAERDEN[4] statt.

Der Beweis von HABICHT und VAN DER WAERDEN beruht auf mechanischen Überlegungen, die sich auf Stangenvierecke beziehen. Im Folgenden geben wir einen einfachen, elementargeometrischen Beweis der Ungleichung (1).[5] Wir werden T_n in wenigstens $n-2$ Dreiecke zerlegen, deren Inhalt $\geqq T_3$ ist.

Wir dürfen $n > 3$ voraussetzen und ergänzen das System der Eckpunkte $E_1, \ldots, E_n$ von T_n durch weitere Punkte $E_{n+1}, \ldots, E_m$ der Kugelfläche, sodaß $E_i E_j \geqq a$ $(i, j = 1, \ldots, m)$ sei, und daß eine weitere derartige Ergänzung nicht mehr möglich sei. Wir bemerken zunächst, daß mit Rücksicht auf die Voraussetzung $n > 3$, die Punkte $E_1, \ldots, E_m$ nicht alle auf einer Halbkugel

[1] W. HABICHT und B. L. VAN DER WAERDEN, Lagerung von Punkten auf der Kugel, *Math. Annalen,* **123** (1951), S. 223—234.

[2] Wir bezeichnen ein Gebiet und seinen Inhalt mit demselben Symbol.

[3] L. FEJES TÓTH, Über eine Abschätzung des kürzesten Abstandes zweier Punkte eines auf einer Kugelfläche liegenden Punktsystems, *Jber. deutscher Math.-Ver.,* **53** (1943) S. 66—68.

[4] K. SCHÜTTE und B. L. VAN DER WAERDEN, Auf welcher Kugel haben 5, 6, 7, 8 oder 9 Punkte mit Mindestabstand Eins Platz? *Math. Annalen,* **123** (1951), S. 96—124.

[5] Der Beweis enthält Gedanken aus dem Aufsatz: L. FEJES TÓTH, On the densest packing of spherical caps, *Amer. math. Monthly,* **56** (1949), S. 330—331.

liegen können. Liegen sie nämlich auf der südlichen Hemisphäre, so bilden sie mit dem Nordpol ein System von wenigstens fünf Punkten mit einem Mindestabstand $> 90^0$, was unmöglich ist.[6] Folglich ist die konvexe Hülle H der Punkte $E_1, \ldots, E_m$ ein konvexes Polyeder mit den Ecken $E_1, \ldots, E_m$, das den Kugelmittelpunkt in seinem Inneren enthält. Wir können H als ein Dreieckspolyeder betrachten, da eine mehr als dreieckige Fläche in Dreiecke zerlegt werden kann. Wir betrachten das sphärische Netz N von H, das durch Zentralprojektion der Kanten von H auf die Kugelfläche entsteht. Die Kugelfläche wird durch N in sphärische Dreiecke zerlegt. Wir zeigen, daß der Inhalt dieser Dreiecke $\geqq T_3$ ist.

Es sei $E_i E_j E_k$ ein Dreieck des Netzes N. Wir legen zunächst fest, daß im Inneren des durch die Ecken von $E_i E_j E_k$ hindurchgehenden Kreises, also außerhalb H, kein Punkt des Systems $\{E_1, \ldots, E_m\}$ liegen kann. Folglich ist der sphärische Radius dieses Kreises $< a$, da sonst das Punktsystem $\{E_1, \ldots, E_m\}$ durch den Mittelpunkt dieses Kreises ergänzt werden könnte.

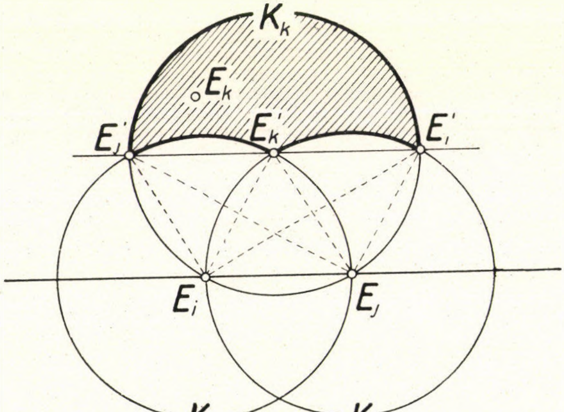


Fig. 1

Es sei $E_j E_k \geqq E_k E_i \geqq E_i E_j \geqq a$. Wir setzen zunächst $E_i E_j \leqq 90^0$ voraus. Wir betrachten das gleichseitige Dreieck $E_i E_j E_k'$, wobei E_k und E_k' auf derselben, etwa auf der „nördlichen" Seite des „Äquators" $E_i E_j$ liegen. Der Inhalt dieses Dreiecks ist offenbar $\geqq T_3$. Wir schlagen um die Punkte E_i, E_j und E_k' je einen Kreis K_i, K_j, K_k' vom Radius $E_i E_j$ und bezeichnen den von E_i verschiedenen Schnittpunkt von K_j und K_k' mit E_i' und den von E_j verschiedenen Schnittpunkt von K_k' und K_i mit E_j' (s. Fig. 1). Da $E_i E_j E_i' E_k'$ und $E_i E_j E_k' E_j'$ sphärische Rhomben sind, sind die Dreiecke $E_i E_j E_k'$, $E_i E_j E_i'$ und $E_i E_j E_j'$ inhaltsgleich. Folglich ist der auf der „nördlichen" Hemisphäre liegende Bogen des Kreises $E_i' E_j' E_k'$ der Ort derjenigen Punkte, die mit den festen

[6] Vgl. die unter [3] und [4] angeführten Aufsätze, oder die Aufgabe 35 von H. DAVENPORT und G. HAJÓS in *Matematikai Lapok*, 2 (1951), S. 68.

Punkten E_i und E_j inhaltsgleiche Dreiecke bilden. Nun liegt der Punkt E_k einerseits außerhalb oder auf dem Rand der Kreise K_i, K_j, andererseits innerhalb des Kreises K'_k, da sonst der Radius des Kreises $E_iE_jE_k$ im Gegensatz zu der obigen Bemerkung $\geqq a$ wäre. Folglich liegt E_k „oberhalb" oder auf dem Rand des Lexellschen Kreises $E'_iE'_jE'_k$, womit unsere Behauptung für $E_iE_j \leqq 90^0$ bewiesen ist.

Es sei nun $E_iE_j > 90^0$. Da wir den Fall $n = 3$ ausgeschlossen haben, ist andererseits $E_iE_j < 120^0$, folglich können E'_k und die Kreise K_i, K_j wie oben definiert werden. Wir betrachten das sphärische Dreieck $E'_kE^*_iE^*_j$, wobei E^*_i und E^*_j die antipodischen Punkte von E_i und E_j bezeichnen. Jeder auf der „nördlichen" Hemisphäre liegende Punkt dieses Dreiecks bildet mit E_i und E_j als Ecken ein Dreieck, das $E_iE_jE'_k$ enthält. Das Dreieck $E'_kE^*_iE^*_j$ enthält wegen $E_iE'_k = E_jE'_k > 90^0$ das außerhalb K_i und K_j liegende Teilgebiet der nördlichen Hemisphäre (s. Fig. 2). Da aber E_k in diesem Teilgebiet liegt, enthält $E_iE_jE_k$ das Dreieck $E_iE_jE'_k$, sodaß $E_iE_jE_k \geqq E_iE_jE'_k \geqq T_3$ ausfällt.

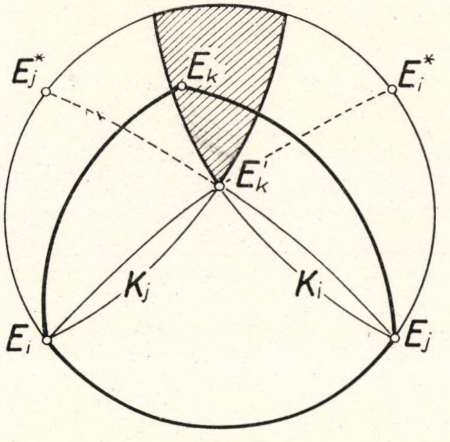


Fig. 2

Wir zeigen nun, daß die Seiten von T_n Kanten des Netzes N sind. Wir schlagen etwa um die sphärische Strecke E_1E_2 als Durchmesser einen Kreis K. Die Abstände der übrigen Ecken $E_3, \ldots, E_m$ von E_1 sind $\geqq E_1E_2 = a$. Diese Ecken liegen also offenbar außerhalb K. Folglich ist die Ebene von K eine Stützebene von H, die nur die Strecke E_1E_2 mit H gemein hat. Daraus folgt aber, daß E_1E_2 tatsächlich eine Kante von H ist.

Bezeichnen wir die Flächen-, Kanten- und Eckenzahl des in T_n enthaltenen Teilnetzes von N mit f, k bzw. e, so haben wir offenbar $2k = 3f + n$ und $e \geqq n$. Nach dem Eulerschen Satz $e - k + f = 1$ folgt hieraus

$$f = 2e - n - 2 \geqq n - 2.$$

Damit ist der Beweis beendet.

(Eingegangen am 10. Dezember 1951.)

ОБ ОЦЕНКЕ ПОВЕРХНОСТЕЙ СФЕРИЧЕСКИХ МНОГОУГОЛЬНИКОВ

И. МОЛЬНАР (Веспрем)

(Резюме)

Автор даёт элементарное доказательство следующей теоремы Габихта и ван дер Вардена.

Пусть дан равносторонний сферический n-угольник со стороной равной a, расстояние любых двух вершин которого $\geqq a$. Если его площадь T_n, то

$$T_n \geqq (n-2)\, T_3.$$

Доказательство основывается на том, что многоугольник разбивается на $n-2$ треугольника с площадью $\geqq T_3$ каждый.

SUR CERTAINS AUTOMORPHISMES A POINTS FIXES DES SURFACES FERMÉES ORIENTABLES

Par

G. GEORGIEV (Sofia)

(*Présenté par* O. VARGA)

Le but de cette note est la démonstration du théorème suivant:

Si la transformation topologique $h(x)$ de la surface fermée orientable Φ_p de genre $p>0$ en lui-même, transforme une courbe fermée simple $C\subset\Phi_p$ homotope à zéro également en elle-même, il existe dans ce cas au moins un point fixe de Φ_p.

Soit la transformation $h(x)$ un automorphisme de la surface fermée orientable Φ_p de genre $p=1,2,\ldots,$[1] ainsi que de la courbe fermée simple $C\subset\Phi_p$ homotope à zéro, [loc. cit., p. 57],

$$h(\Phi_p)=\Phi_p,\quad h(C)=C.$$

Il est évident que l'ensemble Φ_p-C se compose de deux composantes, dont la frontière commune est la courbe C elle-même. L'un de ces deux domaines, désignons le par U, est homéomorphe au plan, mais le second ne possède pas cette propriété pour $p>0$. On aura évidemment

$$\bar{U}=U+C,\quad F(U)=C,$$

$\bar{U}$ et $F(U)$ étant respectivement la fermeture et la frontière du domaine U. De la relation $h(\bar{U})=h(U)+C$ et de $h(\bar{U})=\overline{h(U)}$ on prouve que $\overline{h(U)}-h(U)=C$, c'est-à-dire

$$F[h(U)]=C.$$

La frontière du domaine $h(U)$ étant la courbe C nous aurions

$$h(U)=U,$$

parce que le domaine unique, dont la frontière est la courbe C et qui est homéomorphe au plan, est le domaine U. De cette manière on aura $h(\bar{U})=$ $=h(U)+C=U+C=\bar{U}$, c'est-à-dire

$$(\alpha)\qquad h(\bar{U})=\bar{U}.$$

[1] П. С. Александров, Комбинаторная топология (Москва—Ленинград, 1947), p. 137, § 7.

La relation (α) montre, que la transformation $h(x)$ est un automorphisme d'élément $\bar{U}$. Mais dans ce cas, d'après le théorème, [loc. cit., p. 201, § 4], de L. BROUWER pour les points fixes, la transformation continue $h(x)$ d'élément $\bar{U}$ en lui-même possède au moins un point fixe $x_0 = h(x_0) \in \bar{U} \subset \Phi_p$.

Signalons, que le théorème n'est pas valable pour le sphère Φ_0. En effet, l'automorphisme $h(x)$ de Φ_0 qui fait correspondre au point $x \in \Phi_0$ le point x' diamétralement opposé à x ne possède pas des points fixes, quoique $h(x)$ transforme chaque grande circonférence $C \subset \Phi_0$ en elle-même.

Notons aussi que la condition, que la courbe C soit homotope à zéro, est essentielle. En effet, l'automorphisme $h(x)$ du tore Φ_1, qui fait correspondre au point $x \in \Phi_1$ le point x' diamétralement opposé à x et situé sur le méridien passant par x, ne possède pas des points fixes, bien que $h(x)$ transforme chaque méridien $C \subset \Phi_1$ en lui-même.

(Reçu le 1. mars 1952.)

О НЕКОТОРЫХ АВТОМОРФИЗМАХ С НЕПОДВИЖНЫМИ ТОЧКАМИ НОРМАЛЬНЫХ ЗАМКНУТЫХ ОРИЕНТИРУЕМЫХ ПОВЕРХНОСТЕЙ

Г. ГЕОРГИЕВ (София)

(Р е з ю м е)

В настоящем сообщении устанавливается следующее свойство:

Если топологическое отображение $h(x)$ нормальной замкнутой ориентируемой поверхности Φ_p рода $p \geqq 1$ на себя отображает простую замкнутую линию $C \subset \Phi_p$ гомотопной точки также на себя, то при этом отображении имеется, по крайней мере, одна неподвижная точка $x_0 = h(x_0) \in \Phi_p$.

ÜBER DIE CESÀROSCHE SUMMIERBARKEIT DER ORTHOGONALEN POLYNOMREIHEN

Von

KÁROLY TANDORI (Budapest)

(*Vorgelegt von* G. ALEXITS)

1. Einleitung

Es sei $w(x)$ eine im Intervall $[a, b]$ nicht-negative, L-integrierbare Funktion. Nehmen wir an, daß auf einer Menge von positivem Maß $w(x) > 0$ ist. Betrachten wir das vollständige System $\{p_n(x)\}$ der in $[a, b]$ zur Gewichtsfunktion $w(x)$ gehörenden orthonormierten Polynome; für diese Polynome besteht also die Relation

$$\int_a^b w(x)p_n(x)p_m(x)\,dx = \begin{cases} 0 & (n \neq m), \\ 1 & (n = m). \end{cases}$$

Bezeichnen wir mit L_w^2 die Klasse der in $[a, b]$ meßbaren Funktionen $f(x)$, für welche

$$\int_a^b w(x)f^2(x)\,dx < \infty$$

gilt. Ist z. B. $f(x)$ in $[a, b]$ stetig, so gehört $f(x)$ zu L_w^2.

Für jede Funktion $f(x) \in L_w^2$ existiert die Entwicklung

$$f(x) \sim \sum_{k=0}^{\infty} a_k p_k(x),$$

wo

$$a_k = \int_a^b w(x)f(x)p_k(x)\,dx \qquad (k = 0, 1, 2, \ldots)$$

die verallgemeinerten Fourierkoeffizienten der Funktion $f(x)$ bedeuten. Bezeichne $s_\nu(f; x)$ die ν-te Partialsumme dieser Entwicklung, d. h.

$$s_\nu(f; x) = \sum_{k=0}^{\nu} a_k p_k(x).$$

Betreffs der Konvergenz fast überall der orthogonalen Polynomentwicklung einer Funktion $f(x) \in L_w^2$ hat Prof. G. ALEXITS[1] ein gewisses Analogon

[1] G. ALEXITS, Sur la convergence des séries de polynomes orthogonaux, *Commentarii Math. Helvetici,* **16** (1943—44), S. 200—208.

des für Fourierreihen bekannten Satzes von KOLMOGOROFF—SELIVERSTOFF und PLESSNER [2] bewiesen. Prof. BÉLA SZ.-NAGY hat dieses Ergebnis unter viel schwächeren Annahmen verallgemeinert.[3] Bei ihm wird als einzige Bedingung gefordert, daß die Polynome in jedem, ganz im Inneren von $[a, b]$ liegenden abgeschlossenen Intervall gleichmäßig beschränkt seien.

Prof. ALEXITS hat mir die Vermutung mitgeteilt, daß die Entwicklung jeder zur Klasse L^2_w gehörenden Funktion unter dieser Bedingung fast überall $(C, 1)$-summierbar ist. In dieser Arbeit werde ich diesbezüglich zwei Sätze beweisen. Diese Sätze und ihre Beweisführungen sind den auf die „starke Summierbarkeit" der Fourierreihen bezüglichen bekannten Sätzen von HARDY und LITTLEWOOD [4] ähnlich.

SATZ 1. *Ist das Polynomsystem* $\{p_n(x)\}$ *in einem Intervall* $[c, d]$ $(a \leqq c < d \leqq b)$ *gleichmäßig beschränkt, so besteht für jede Funktion* $f(x) \in L^2_w$ *die Beziehung*

$$\frac{1}{n+1} \sum_{\nu=0}^{n} [s_\nu(f; x) - f(x)]^2 \to 0 \tag{1}$$

in $[c, d]$ *fast überall.*

SATZ 2. *Ist das Polynomsystem* $\{p_n(x)\}$ *in einem Intervall* $[c, d]$ $(a \leqq c < d \leqq b)$ *gleichmäßig beschränkt und die Gewichtsfunktion* $w(x)$ *in diesem Intervall von oben beschränkt, so gilt* (1) *in jedem Punkte* $x \in (c, d)$, *in welchem* $f(x)$ *stetig ist. Wenn aber* $f(x)$ *im ganzen offenen Intervall* (c, d) *stetig ist, so besteht* (1) *in jedem Intervall* $[c+\delta, d-\delta]$ *für jedes* $\delta > 0$ *gleichmäßig.*

Nach der Cauchyschen Ungleichung folgt aus (1)

$$\frac{1}{n+1} \sum_{\nu=0}^{n} |s_\nu(f; x) - f(x)| \to 0,$$

also erhalten wir, wenn

$$\sigma_n(f; x) = \frac{s_0(f; x) + \ldots + s_n(f; x)}{n+1}$$

bezeichnet, aus unseren Sätzen die folgenden Ergebnisse:

Ist das Polynomsystem $\{p_n(x)\}$ *in einem Teilintervall* $[c, d]$ *von* $[a, b]$ *gleichmäßig beschränkt, so gilt*

$$\sigma_n(f; x) \to f(x) \tag{2}$$

für jede Funktion $f(x) \in L^2_w$ *in* $[c, d]$ *fast überall.*

Bleibt auch die Gewichtsfunktion $w(x)$ *in* $[c, d]$ *beschränkt, so besteht* (2) *in jedem Punkte von* (c, d), *in welchem* $f(x)$ *stetig ist. Wenn* $f(x)$ *in* (c, d)

[2] Siehe z. B. A. ZYGMUND, *Trigonometrical series* (Warszawa—Lwów, 1935), S. 253—254.

[3] BÉLA SZ.-NAGY, Über die Konvergenz von Reihen orthogonaler Polynome, *Math. Nachrichten,* 4 (1950—51), S. 50—55. Verfasser beschäftigt sich in dieser Arbeit mit den allgemeinsten, zu einer Gewichtsfunktion der Form $d\alpha(x)$ gehörenden orthogonalen Polynomsystemen.

[4] Siehe z. B. loc. cit.[2], S. 237—239.

überall stetig ist, so gilt (2) *in jedem Intervall* $[c+\delta, d-\delta]$ $(\delta>0)$ *gleichmäßig.*

Sind die Bedingungen des Satzes 1, *bzw.* 2 *in jedem ganz im Inneren von* $[a, b]$ *liegenden, abgeschlossenen Teilintervall* $[c, d]$ *erfüllt, so ist die Entwicklung jeder Funktion* $f(x)\in L_w^2$ *in* $[a, b]$ *fast überall* $(C, 1)$*-summierbar, bzw. ist die Entwicklung jeder stetigen Funktion in jedem* $[c, d]$ *gleichmäßig* $(C, 1)$*-summierbar.*

Aus der letzten Bemerkung ergibt sich auf bekannte Weise[5] die

FOLGERUNG 1. *Wenn in jedem ganz im Inneren von* $[a, b]$ *liegenden, abgeschlossenen Teilintervall das Polynomsystem* $\{p_n(x)\}$ *gleichmäßig beschränkt ist und die Gewichtsfunktion* $w(x)$ *beschränkt bleibt, so bleiben die zu der* $(C, 1)$*-Summation gehörenden Lebesgueschen Funktionen*

$$L_n^{(1)}(x)=\int_a^b\left|\sum_{k=0}^{n}\left(1-\frac{k}{n+1}\right)p_k(x)p_k(t)\right|dt \qquad (n=0, 1, 2, \ldots)$$

in jedem ganz im Inneren von $[a, b]$ *liegenden, abgeschlossenen Teilintervall gleichmäßig beschränkt.*

Prof. BÉLA SZ.-NAGY hat mir die Bemerkung mitgeteilt, daß man aus den obigen Ergebnissen folgende merkwürdige Behauptung erhält:

FOLGERUNG 2. *Gilt für die Gewichtsfunktion* $w(x)=0$ *auf einer Menge* $M\subset[c, d]$ *von positivem Maß, so bleibt das Polynomsystem* $\{p_n(x)\}$ *in* $[c, d]$ *nicht gleichmäßig beschränkt.*

Betrachten wir nämlich zwei Funktionen $f_1(x)\in L_w^2$, $f_2(x)\in L_w^2$, die auf der Menge M verschieden, während auf $[a, b]-M$ einander gleich sind. Wäre das Polynomsystem in $[c, d]$ gleichmäßig beschränkt, so müßte nach unseren obigen Ergebnissen in $[c, d]$ fast überall sowohl $\lim_{n\to\infty}\sigma_n(f_1; x)=f_1(x)$ als auch $\lim_{n\to\infty}\sigma_n(f_2; x)=f_2(x)$ sein. Da aber bei der Auswertung der Entwicklungskoeffizienten a_k nur diejenigen Stellen in Betracht kommen, wo die Gewichtsfunktion $w(x)$ von 0 verschieden ist, gilt in $[c, d]$ fast überall $\lim_{n\to\infty}\sigma_n(f_1, x)=$ $=\lim_{n\to\infty}\sigma_n(f_2; x)$, im Gegensatz zu unserer Annahme $f_1(x)\neq f_2(x)$ für $x\in M$.

In Verbindung mit der Bemerkung von BÉLA SZ.-NAGY teilte mir Prof. P. TURÁN mit, daß man ein schärferes Resultat erhält, wenn die Menge M ein *ganzes* Intervall ausfüllt. Dann wächst nämlich die Folge $\{p_n(x)\}$ in einem Teilintervall exponentiell; genauer:

Verschwindet $w(x)$ *in einem echten Teilintervall* $[c, d]$ *von* $[a, b]$, *so existiert zu jedem Intervall* $c<c'\leq x\leq d'<d$ *eine Konstante* $\gamma=\gamma(c', d')>1$

[5] Siehe z. B. S. KACZMARZ und H. STEINHAUS, *Theorie der Orthogonalreihen* (Warszawa Lwów, 1935), S. 157.

derart, daß für jedes $x \in [c', d']$ die Beziehung

$$\overline{\lim_{n\to\infty}} \frac{p_n(x)}{\gamma^n} > 0$$

besteht.

Sei in der Tat $P_n(x)$ ein Polynom n-ten Grades, welches in x_0 die Bedingung

$$P_n(x_0) = 1 \tag{3}$$

erfüllt. Stellen wir das Polynom $P_n(x)$ in der Form

$$\sum_{\nu=0}^{n} c_\nu p_\nu(x)$$

dar, so ergibt sich die Gleichung

$$\int_a^b P_n^2(x) w(x)\, dx = \sum_{\nu=0}^{n} c_\nu^2,$$

und die Bedingung (3) nimmt die Form

$$\sum_{\nu=0}^{n} c_\nu p_\nu(x_0) = 1$$

an. Nach der Schwarzschen Ungleichung gilt also

$$1 = \left(\sum_{\nu=0}^{n} c_\nu p_\nu(x_0)\right)^2 \leq \left(\sum_{\nu=0}^{n} c_\nu^2\right)\left(\sum_{\nu=0}^{n} p_\nu^2(x_0)\right),$$

woraus die Abschätzung

$$\int_a^b P_n^2(x) w(x)\, dx \geq \frac{1}{\sum_{\nu=0}^{n} p_\nu^2(x_0)} \tag{4}$$

folgt. Sei alsdann $x_0 \in [c', d']$ und betrachten wir die Funktion

$$g(x) = \left\{1 - \left(\frac{x - x_0}{b - a}\right)^2\right\}^{\left[\frac{n}{2}\right]};$$

$g(x)$ ist ein Polynom höchstens n-ten Grades und erfüllt die Bedingung (3). Wegen unserer Annahme betreffs $w(x)$ gilt

$$\int_a^b g^2(x) w(x)\, dx = \left(\int_a^c + \int_d^b\right) g^2(x) w(x)\, dx. \tag{5}$$

Da aber eine positive Konstante $\vartheta = \vartheta(x_0) < 1$ existiert derart, daß außerhalb des Intervalls $[c, d]$ $0 < g(x) \leq \vartheta^n$ ist, ergibt sich aus (5)

$$\int_a^b g^2(x) w(x)\, dx \leq \vartheta^{2n} \int_a^b w(x)\, dx.$$

Nach (4) ergibt sich somit für jedes n die Ungleichung

$$\frac{1}{\sum_{\nu=0}^{n} p_\nu^2(x_0)} \leqq \vartheta^{2n} \int_a^b w(x)\, dx.$$

Setzt man hier $\vartheta^{-1} = \gamma$, so folgt daraus wegen $\sum_{\nu=0}^{n} p_\nu^2(x) \geqq \text{Konst.}\ \gamma^{2n}$ die Behauptung.

2. Beweis des Satzes 1

Bekanntlich gilt

$$s_\nu(f;\, x) - f(x) = \int_a^b w(t) K_\nu(x, t)\, [f(t) - f(x)]\, dt, \tag{6}$$

wo

$$K_\nu(x, t) = \sum_{k=0}^{\nu} p_k(x) p_k(t)$$

ist. Nach der Christoffel—Darbouxschen Formel läßt sich $K_\nu(x, t)$ in der Form

$$K_\nu(x, t) = \frac{\alpha_\nu}{\alpha_{\nu+1}} \frac{p_{\nu+1}(x) p_\nu(t) - p_\nu(x) p_{\nu+1}(t)}{t - x}$$

schreiben, wo α_ν den Koeffizienten der Potenz x^ν in dem Polynom $p_\nu(x)$ bedeutet.[6] Für diese Konstanten gilt die Beziehung[7]

$$\frac{\alpha_\nu}{\alpha_{\nu+1}} = O(1). \tag{7}$$

Sei x ein innerer Punkt des Intervalls $[c, d]$ und n so groß, daß $c < x - \frac{1}{n} < x + \frac{1}{n} < d$ gelte. Zerlegen wir das Integral (6) im Falle $1 \leqq \nu \leqq n$ folgenderweise:

$$s_\nu(f; x) - f(x) = \Bigl(\int_a^{x-\frac{1}{n}} + \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} + \int_{x+\frac{1}{n}}^{b} \Bigr) w(t) K_\nu(x, t)\, [f(t) - f(x)]\, dt = \sum_{i=1}^{3} I_{i\nu}^{(n)}(x).$$

Mit Hilfe der Ungleichung von Minkowski erhalten wir:

$$\Bigl\{ \frac{1}{n+1} \sum_{\nu=0}^{n} (s_\nu(f; x) - f(x))^2 \Bigr\}^{1/2} \leqq \sum_{i=1}^{3} \Bigl\{ \frac{1}{n+1} \sum_{\nu=0}^{n} |I_{i\nu}^{(n)}(x)|^2 \Bigr\}^{1/2} =$$
$$= S_1^{(n)}(x) + S_2^{(n)}(x) + S_3^{(n)}(x).$$

Zum Beweise des Satzes 1 werden wir zeigen, daß jedes Glied der rechten Seite in $[c, d]$ f. ü. a. nach 0 konvergiert.

[6] Siehe z. B. G. Szegő, *Orthogonal polynomials* (New York, 1939), S. 42.
[7] Siehe z. B. loc. cit. [3], S. 53.

Betrachten wir zuerst $S_2^{(n)}(x)$. Da das Polynomsystem in $[c, d]$ gleichmäßig beschränkt ist, ergibt sich die Abschätzung

$$|I_{2\nu}^{(n)}(x)| = \left| \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} w(t) K_\nu(x, t) [f(t)-f(x)]\, dt \right| \leq O(\nu) \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} w(t)|f(t)-f(x)|\, dt =$$

$$= O(\nu) \left\{ \int_0^{1/n} w(x+v)\, |f(x+v)-f(x)|\, dv + \int_0^{1/n} w(x-v)\, |f(x-v)-f(x)|\, dv \right\}. \tag{8}$$

Es ist klar, daß

$$w(x \pm v)\, |f(x \pm v)-f(x)| \leq$$
$$\leq |w(x \pm v) f(x \pm v) - w(x) f(x)| + |f(x)|\, |w(x \pm v) - w(x)|$$

gilt. Da $w(t)f(t)$ und $w(t)$ integrierbare Funktionen sind,[8] besteht[9]

$$\int_0^{1/n} |w(x \pm v) f(x \pm v) - w(x) f(x)|\, dv = o\left(\frac{1}{n}\right), \quad \int_0^{1/n} |w(x \pm v) - w(x)|\, dv = o\left(\frac{1}{n}\right)$$

f. ü. a. Aus (6) folgt $I_{2\nu}^{(n)}(x) = o\left(\frac{\nu}{n}\right)$, voraus sich $S_2^{(n)}(x) \to 0$ in $[c, d]$ f. ü. a. ergibt.

Wir zeigen nun, daß auch $S_1^{(n)}(x)$ und $S_2^{(n)}(x)$ in $[c, d]$ f. ü. a. nach 0 konvergieren. Den Beweis führen wir nur im Falle $S_1^{(n)}(x)$ durch, der Beweis für $S_3^{(n)}(x)$ ist ähnlich.

Mit Hilfe der Formel von CHRISTOFFEL—DARBOUX erhalten wir die Gleichung

$$I_{1\nu}^{(n)}(x) = \int_a^{x-\frac{1}{n}} w(t)[f(t)-f(x)] K_\nu(x, t)\, dt =$$

$$= \frac{\alpha_\nu}{\alpha_{\nu+1}} p_{\nu+1}(x) \int_a^{x-\frac{1}{n}} w(t) \frac{f(t)-f(x)}{t-x} p_\nu(t)\, dt - \frac{\alpha_\nu}{\alpha_{\nu+1}} p_\nu(x) \int_a^{x-\frac{1}{n}} w(t) \frac{f(t)-f(x)}{t-x} p_{\nu+1}(t)\, dt.$$

Da das Polynomsystem $\{p_n(x)\}$ nach Annahme im Intervall $[c, d]$ gleichmäßig beschränkt ist, besteht wegen (7) $I_{1\nu}^{(n)}(x) = O(1) c_\nu(x) + O(1) c_{\nu+1}(x)$, wo $c_\nu(x)$ den ν-ten verallgemeinerten Fourierkoeffizienten der Funktion

$$g_x(t) = \begin{cases} \dfrac{f(t)-f(x)}{t-x} & \left(a \leq t \leq x - \dfrac{1}{n}\right), \\ 0 & \left(x - \dfrac{1}{n} < t \leq b\right) \end{cases}$$

[8] Nach der Schwarzschen Ungleichung folgt aus $f \in L_w^2$, daß wf im Lebesgueschen Sinne integrierbar ist.

[9] Siehe z. B. loc. cit. [2], S. 30.

bedeutet. Es is klar, daß $g_x(t) \in L^2_w$ gilt, da $\dfrac{1}{t-x}$ im Intervall $\left[a, x-\dfrac{1}{n}\right]$ beschränkt bleibt.

Auf Grund der Ungleichung $|\alpha+\beta|^2 \leq 2(|\alpha|^2+|\beta|^2)$ erhalten wir somit die Abschätzung

$$(9)\qquad \left\{\frac{1}{n+1}\sum_{\nu=0}^{n}|I_{2\nu}^{(n)}(x)|^2\right\}^{1/2} \leq \left\{\frac{O(1)}{n+1}\sum_{\nu=0}^{n}c_\nu^2(x)+\frac{O(1)}{n+1}\sum_{\nu=0}^{n}c_{\nu+1}^2(x)\right\}^{1/2} \leq$$
$$\leq \frac{O(1)}{\sqrt{n+1}}\left\{\sum_{\nu=0}^{\infty}c_\nu^2(x)\right\}^{1/2}.$$

Nach der Parsevalschen Formel ist

$$(10)\qquad \sum_{\nu=0}^{\infty}c_\nu^2(x)=\int_a^{1-\frac{1}{n}} w(t)\left[\frac{f(t)-f(x)}{t-x}\right]^2 dt=\int_{1/n}^{x-a} w(x-v)\left[\frac{f(x-v)-f(x)}{v}\right]^2 dv.$$

Wir wollen nun die Funktion

$$\Phi_x(h)=\int_0^h w(x-v)\,[f(x-v)-f(x)]^2\,dv$$

abschätzen. Es ist klar, daß die Ungleichung

$$w(x-v)\,[f(x-v)-f(x)]^2=[\sqrt{w(x-v)}f(x-v)-\sqrt{w(x-v)}f(x)]^2\leq$$
$$\leq 2\,\{[\sqrt{w(x-v)}f(x-v)-\sqrt{w(x)}f(x)]^2+f^2(x)[\sqrt{w(x-v)}-\sqrt{w(x)}]^2\}$$

gilt. Da $\sqrt{w(t)}f(t)$ und $\sqrt{w(t)}$ L-integrierbar sind,[10] folgt

$$\int_0^{1/n}[\sqrt{w(x-v)}f(x-v)-\sqrt{w(x)}f(x)]^2\,dv=o\left(\frac{1}{n}\right),$$

$$\int_0^{1/n}[\sqrt{w(x-v)}-\sqrt{w(x)}]^2\,dv=o\left(\frac{1}{n}\right)$$

f. ü. a.,[11] mithin gilt f. ü. a. auch $\Phi_x\left(\dfrac{1}{n}\right)=o\left(\dfrac{1}{n}\right)$. Partielle Integration führt also zur Abschätzung

$$\int_{\frac{1}{n}}^{x-a} w(x-v)\left[\frac{f(x-v)-f(x)}{v}\right]^2 dv=\left[\frac{\Phi_x(v)}{v^2}\right]_{\frac{1}{n}}^{x-a}+2\int_{\frac{1}{n}}^{x-a}\frac{\Phi_x(v)}{v^3}\,dv=$$
$$=o(n)+\int_{\frac{1}{n}}^{x-a}o(v^{-2})\,dv=o(n).$$

[10] Nach der Schwarzschen Ungleichung folgt nämlich aus $f\in L^2_w$ und $w\in L$, daß $\sqrt{w}f\in L$ und $\sqrt{w}\in L$ ist.

[11] Siehe z. B. loc. cit. [2], S. 237—238.

Nach (9) und (10) folgt daher, daß $S_1^{(n)}(x)$ in $[c, d]$ f. ü. a. nach 0 konvergiert, womit der Satz 1 bewiesen ist.

3. Beweis des Satzes 2

Satz 2 läßt sich ähnlich beweisen, wie der Satz 1. Wir haben nur statt (8) zu beachten, daß die Abschätzung

$$|I_{2\nu}^{(n)}(x)| = \left| \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} w(t) K_\nu(x,t) [f(t)-f(x)]\, dt \right| \leq O(1) \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} |f(t)-f(x)|\, dt$$

wegen der angenommenen Beschränktheit von $w(x)$ in $[c, d]$ gilt. Wenn $f(x)$ in dem Punkt x_0 stetig ist, so ist

$$\int_{x_0-\frac{1}{n}}^{x_0+\frac{1}{n}} |f(t)-f(x_0)|\, dt = o\left(\frac{1}{n}\right).$$

Ist $f(x)$ sogar im ganzen Intervall stetig, so besteht diese Relation in jedem ganz im Inneren von $[c, d]$ liegenden, abgeschlossenen Intervall gleichmäßig.

Der Beweis des Satzes 1 muß noch an einer anderen Stelle abgeändert werden: Wir zerlegen das Integral (10) folgenderweise:

$$\sum_{\nu=0}^{\infty} c_\nu^2(x) = \int_a^{x-\frac{1}{n}} w(t) \left[\frac{f(t)-f(x)}{t-x}\right]^2 dt = \left(\int_a^c + \int_c^{x-\frac{1}{n}}\right) w(t) \left[\frac{f(t)-f(x)}{t-x}\right]^2 dt.$$

Da $w(x)$ nach Annahme in $[c, d]$ beschränkt bleibt, folgt daraus

$$\sum_{\nu=0}^{\infty} c_\nu^2(x) \leq O(1) + O(1) \int_c^{x-\frac{1}{n}} \left[\frac{f(t)-f(x)}{t-x}\right]^2 dt.$$

Wenn aber $f(x)$ in dem Punkt x_0 stetig ist, so besteht

$$\int_c^{x_0-\frac{1}{n}} \left[\frac{f(t)-f(x_0)}{t-x_0}\right]^2 dt = o(n).$$

Daraus ergibt sich, wie im Beweis des Satzes 1, durch partielle Integration $\sum_{\nu=0}^{\infty} c_\nu^2(x_0) = o(n)$. Es ist leicht zu sehen, daß $\sum_{\nu=0}^{\infty} c_\nu^2(x_0) = o(n)$ in jedem ganz im Inneren von $[c, d]$ liegenden Intervall gleichmäßig erfüllt ist, wenn $f(x)$ in (c, d) stetig ist. Aus (7) folgt also, daß $S_1^{(n)}(x)$ an jeder Stetigkeitsstelle $x \in (c, d)$ nach 0 konvergiert, und wenn $f(x)$ in (c, d) überall stetig ist, so ist die Konvergenz im jedem ganz im Inneren von $[c, d]$ liegenden, abgeschlossenen Intervall gleichmäßig.

BEMERKUNG. Aus dem Gang des ganzen Beweises ist ersichtlich, daß der volle Hardy—Littlewoodsche Satz über starke Summierbarkeit unter zusätzlichen Annahmen über das System $\{p_n(x)\}$ auch für orthogonale Polynomentwicklungen gilt.

III. MATHEMATISCHES INSTITUT
DER TECHNISCHEN UNIVERSITÄT VON BUDAPEST

(Eingegangen am 8. März 1952.)

О СУММАБИЛЬНОСТИ РЯДОВ ОРТОГОНАЛЬНЫХ МНОГОЧЛЕНОВ В СМЫСЛЕ ЧЕЗАРО

К. ТАНДОРИ (Будапешт)

(Резюме)

Пусть $\{p_n(x)\}$ ортогональная относительно функции $w(x)$ и нормированная система многочленов, определённая на отрезке $[a, b]$. Автор доказывает, что если система многочленов равномерно ограниченна на отрезке $[c, d]$ $(a \leqq c < d \leqq b)$, то для любой функции $f(x)$ на отрезке $[c, d]$ справедливо почти везде (1), если только $\int_a^b f^2(x)\, w(x)\, dx < \infty$ (теорема 1); если ещё и $w(x)$ ограниченна на отрезке (c, d), то (1) справедливо для любой точки непрерывности $f(x)$, а если $f(x)$ непрерывна на (c, d), то (1) равномерно справедлива на любом замкнутом отрезке, лежащем внутри (c, d) (теорема 2).

Из теоремы 1 согласно замечанию Б. С.-Надь следует, что если на некотором множестве положительной меры, являющемся частью $[c, d]$, $w(x)$ равняется нулю, то система многочленов не может быть равномерно ограниченна на $[c, d]$.

Работа содержит результат П. Турана, согласно которому, если $w(x) = 0$ на $[c, d]$, то для любого отрезка $c < c' \leqq x \leqq d' < d$ можно найти такую $\gamma = \gamma(c', d')$, что для любой точки $x \in [c', d']$

$$\overline{\lim_{n \to \infty}}\, |p_n(x)|/\gamma^n > 0.$$

ÜBER DIE STARKE $(C, 1)$-SUMMIERBARKEIT VON ORTHOGONALEN POLYNOMREIHEN

Von

GÉZA FREUD (Budapest)

(*Vorgelegt von* G. ALEXITS)

1. Einleitung

K. TANDORI [7] hat in der vorstehenden Arbeit bewiesen, daß der bekannte Satz von HARDY und LITTLEWOOD [3] über die starke $(C, 1)$-Summierbarkeit der Fourierreihe für eine recht allgemeine Klasse von orthogonalen Polynomreihen verallgemeinert werden kann. Genauer formuliert, zeigt TANDORI folgendes: Sei $p_0(x), p_1(x), \ldots$ die Folge der im Intervall (a, b) durch die L-integrierbare Gewichtsfunktion $w(x) \geqq 0$ eindeutig bestimmten orthogonalen normierten Polynome, wo $p_n(x)$ das Polynom n-ten Grades bedeutet;[1] gilt in einem inneren Teilintervall (α, β) von (a, b)

$$(1) \qquad |p_n(x)| \leqq K \qquad (n = 0, 1, \ldots)$$

(mit K unabhängig von n), ist ferner $f(x)$ meßbar und $w(x)[f(x)]^2$ L-integrierbar (im Folgenden kurz: $f \in L^2(w)$), so ist die Entwicklung

$$(2) \qquad f(x) \sim \sum_{\nu=0}^{\infty} a_\nu p_\nu(x), \quad a_\nu = \int_a^b f(x) p_\nu(x) w(x)\, dx$$

von $f(x)$ in (α, β) fast überall (kurz: f. ü.) stark $(C, 1)$-summierbar.

Wir wollen zeigen, daß die Sätze von TANDORI wesentlich verallgemeinert werden können, da wir beweisen, daß statt (1) schon die schwächere Forderung

$$(3) \qquad \sum_{\nu=0}^{n} [p_\nu(x)]^2 = O(n)$$

genügt.

Wir zeigen ferner, daß es genügt das Erfülltsein von (3) nur für ein festes x anzunehmen, die $(C, 1)$-Summierbarkeit folgt dann aus der lokalen Bedingung (4).

[1] Wie üblich, sei der Koeffizient von x^n in $p_n(x)$ positiv.

Die Bedeutung dieser Verallgemeinerung liegt u. a. auch darin, daß (3) *schon dann erfüllt ist, wenn $w(t)$ in einer Umgebung von x f. ü. eine positive untere Schranke besitzt.*

Obzwar (1) für die klassischen Jacobischen Polynome erfüllt ist, scheint keine allgemeine Eigenschaft von $w(x)$ bekannt zu sein, durch welche das Bestehen von (1) gesichert wäre. Die in der Literatur vorhandenen hinreichenden Bedingungen enthalten starke Voraussetzungen über $w(x)$ (BERNSTEIN [1], GERONIMUS [2], KOROUS [4],[2] SZEGŐ [6][3]), ergeben aber auch mehr als es in (1) gefordert wird, nämlich Asymptotik bzw. das Bestehen von (1) im Inneren des ganzen Orthogonalitätsintervalls.

2. Sätze und Folgerungen

SATZ I. *Ist* (3) *an einer Stelle x $(a<x<b)$ gültig, $f(x)\in L^2(w)$ und*

$$\varphi_x(h)=\int_x^{x+h} w(t)[f(t)-f(x)]^2\,dt=o(|h|), \tag{4}$$

so folgt

$$\lim_{n=\infty}\frac{1}{n+1}\sum_{\nu=0}^{n}|s_\nu(x)-f(x)|=0. \tag{5}$$

Sind (3) *und* (4) *in einem Intervall $(\alpha,\beta$, $a\leqq\alpha<\beta\leqq b$, gleichmäßig erfüllt, so gilt* (5) *in jedem inneren Teilintervall von (α,β) gleichmäßig.*

SATZ II. *Sei*

$$w(x)>m>0 \tag{6}$$

in einem Teilintervall (α,β) von (a,b); dann ist (3) *in jedem inneren Teilintervall von (α,β) gleichmäßig erfüllt.*

Ist $f\in L^2(w)$, so ist die Bedingung (4) f. ü. in (a,b) erfüllt (TANDORI [7]), also ist die $(C,1)$-Summierbarkeit von (2) in (α,β) f. ü. gesichert. Weiter folgt nach einem Satz von ZYGMUND [8] der

SATZ III. *Ist* (3) *in jedem Punkt x einer Teilmenge E von (a,b) erfüllt, so ist* (2) *in E f. ü. stark $(C,1)$-summierbar.*[4]

Ist

$$M>w(x)>m>0 \quad \text{für } \alpha\leqq x\leqq\beta, \tag{7}$$

so besteht (3) nach Satz II gleichmäßig in jedem inneren Teilintervall von

[2] S. auch SZEGŐ, loc. cit., Theorem 7.1.3, S. 157.

[3] Theorem 12.1.6, S. 291.

[4] Die starke $(C,1)$-Summierbarkeit bedeutet bekanntlich, daß

$$\lim_{n=\infty}\frac{1}{n+1}\sum_{\nu=0}^{n}[s_\nu(x)-f(x)]^2=0$$

gültig ist.

(α, β); ist außerdem $f(x)$ in (α, β) stetig, so gilt (4) gleichmäßig in jedem inneren Teilintervall von (α, β). Hieraus folgt nach Satz I, daß auch (5) gleichmäßig in jedem inneren Teilintervall von (α, β) erfüllt ist, so daß auch der folgende Satz richtig ist:

SATZ IV. *Es sei* (7) *erfüllt und* $f(x)$ *stetig in* (α, β); *dann konvergieren die arithmetischen Mittel der Teilsummen von* (2) *gleichmäßig in jedem inneren Teilintervall von* (α, β) *gegen* $f(x)$. (Verallgemeinerung des Approximationssatzes von FEJÉR.)

3. Beweis des Satzes I

Es seien n und $\nu \leqq n$ nichtnegative ganze Zahlen. Nach (2) ist

$$s_\nu(x) = \sum_{r=0}^{\nu} a_r p_r(x) = \int_a^b K_\nu(x, t) f(t) w(t)\, dt \tag{8}$$

mit

$$K_\nu(x, t) = \sum_{r=0}^{\nu} p_r(x) p_r(t) = \gamma_\nu \frac{p_{\nu+1}(x) p_\nu(t) - p_\nu(x) p_{\nu+1}(t)}{x - t}, \tag{9}$$

wobei nach der Ungleichung von SCHWARZ

$$0 < \gamma_\nu = \int_a^b p_\nu(t) p_{\nu+1}(t) \cdot t w(t)\, dt \leqq \operatorname{Max}(|a|, |b|) \tag{10}$$

ist. Aus (8) und (9) erhalten wir

$$s_\nu(x) - f(x) = A_\nu - \gamma_\nu p_{\nu+1}(x) B_\nu + \gamma_\nu p_\nu(x) B_{\nu+1} \tag{11}$$

mit

$$A_\nu = \int_{x-\frac{c}{n+1}}^{x+\frac{c}{n+1}} K_\nu(x, t) [f(t) - f(x)]\, w(t)\, dt, \tag{12}$$

$$B_\nu = \int_a^{x-\frac{c}{n+1}} \frac{f(t)-f(x)}{t-x} p_\nu(t) w(t)\, dt + \int_{x+\frac{c}{n+1}}^{b} \frac{f(t)-f(x)}{t-x} p_\nu(t) w(t)\, dt. \tag{13}$$

Wählen wir die Konstante $c > 0$ so, daß $a < x - c < x + c < b$ sei. Aus (12), (9), (3) und (4) folgt dann wegen $\nu \leqq n$:

$$\begin{aligned} |A_\nu| &\leqq \Bigg(\int_{x-\frac{c}{n+1}}^{x+\frac{c}{n+1}} [f(t)-f(x)]^2 w(t)\, dt \Bigg)^{1/2} \Bigg(\int_a^b [K_\nu(x, t)]^2 w(t)\, dt \Bigg)^{1/2} = \\ &= o(n^{-1/2}) \Bigg(\sum_{r=0}^{\nu} [p_r(x)]^2 \Bigg)^{1/2} = o(1). \end{aligned} \tag{14}$$

Nach (13) ist B_ν der ν-te Koeffizient der Orthogonalentwicklung von

$$F(t)=\begin{cases} 0 \quad \text{für } x-\dfrac{c}{n+1}<t<x+\dfrac{c}{n+1}, \\ \dfrac{f(t)-f(x)}{t-x} \quad \text{für } t<x-\dfrac{c}{n+1} \text{ oder } t>x+\dfrac{c}{n+1}, \end{cases}$$

also folgt aus der Besselschen Ungleichung und (4):

$$\sum_{\nu=0}^{n+1} B_\nu^2 \leq \int_a^{x-\frac{c}{n+1}} [F(t)]^2 w(t)\,dt + \int_{x+\frac{c}{n+1}}^{b} [F(t)]^2 w(t)\,dt =$$

$$(15) \qquad = \left[\frac{\varphi_x(t-x)}{(t-x)^2}\right]_a^{x-\frac{c}{n+1}} + \left[\frac{\varphi_x(t-x)}{(t-x)^2}\right]_{x+\frac{c}{n+1}}^{b} + 2\int_a^{x-\frac{c}{n+1}} \frac{\varphi_x(t-x)}{(t-x)^3}\,dt +$$

$$+2\int_{x+\frac{c}{n+1}}^{b} \frac{\varphi_x(t-x)}{(t-x)^3}\,dt = o(n).$$

Da nach Annahme $\sum_{\nu=0}^{n} [p_\nu(x)]^2 = O(n)$ ist, ergibt sich somit

$$(16) \qquad \sum_{\nu=0}^{n} |p_\nu(x)|\,|B_{\nu+1}| \leq \left(\sum_{\nu=0}^{n} [p_\nu(x)]^2\right)^{1/2} \left(\sum_{\nu=1}^{n+1} B_\nu^2\right)^{1/2} = O(n^{1/2})\,o(n^{1/2}) = o(n)$$

und ähnlicherweise

$$(17) \qquad \sum_{\nu=0}^{n} |p_{\nu+1}(x)|\,|B_\nu| = o(n).$$

Aus (11), (10), (14), (16) und (17) folgt

$$(18) \qquad \sum_{\nu=0}^{n} |s_\nu(x)-f(x)| = o(n).$$

Sind (3) und (4) in (α, β) gleichmäßig erfüllt, und ist (α_0, β_0) ein inneres Teilintervall von (α, β), so kann für jeden Punkt in (α_0, β_0) dieselbe Konstante c gewählt werden, (18) gilt also in (α_0, β_0) gleichmäßig.

4. Beweis des Satzes II

Bekanntlich is $K_n(x,x)$ die genaue obere Schranke von $|\pi_n(x)|^2$, wenn $\pi_n(x)$ die Polynome höchstens n-ten Grades durchläuft, welche der Bedingung

$$(19) \qquad \int_a^b |\pi_n(t)|^2 w(t)\,dt \leq 1$$

genügen.[5] Im Weiteren bezeichne $\pi_n(t)$ dasjenige der (19) genügenden Polynome, welches den Wert der oberen Schranke $|\pi_n(x)|^2 = K_n(x, x)$ tatsächlich annimmt. Aus (19) und der Forderung $w(x) \geqq m > 0$ folgt, daß

$$m \int_\alpha^\beta |\pi_n(t)|^2 \, dt \leqq 1$$

ist, also besteht nach dem schon benutzten Szegőschen Hilfssatz

$$m|\pi_n(x)|^2 = m K_n(x, x) \leqq \sum_{\nu=0}^{n} [p_\nu^*(x)]^2, \tag{20}$$

wobei $\{p_\nu^*(x)\}$ die dem Grade nach ansteigende Folge der in (α, β) zur Gewichtsfunktion $w^*(x) \equiv 1$ orthogonalen und normierten Polynome bedeutet. Es gilt, wie leicht ersichtlich,

$$p_\nu^*(x) = \sqrt{\frac{(2\nu+1)}{\beta-\alpha}}\, P_\nu\left(-1 + 2\frac{x-\alpha}{\beta-\alpha}\right), \tag{21}$$

wo $P_\nu(x)$ das Legendresche Polynom vom Grade ν mit der üblichen Normierung $P_\nu(1) = 1$ bedeutet. Nach einem klassischen Satz von LAPLACE (vgl. z. B. LENSE [5][6]) gilt in u gleichmäßig

$$P_\nu(u) = O(\nu^{-1/2}) \tag{22}$$

in jedem ganz im Inneren von $(-1, +1)$ liegenden Teilintervall. Aus (20), (21) und (22) ergibt sich mithin[7]

$$K_n(x, x) = \sum_{\nu=0}^{n} [p_\nu(x)]^2 = O(n), \tag{23}$$

w. z. b. w.

(Eingegangen am 5. Mai 1952.)

Literaturverzeichnis

[1] S. BERNSTEIN, Sur les polynomes orthogonaux relatifs à un segment fini, *Journal de Math.* (9), **9** (1930), S. 127—177 und **10** (1931), S. 219—286.

[2] Я. Л. Геронимус, Об ортогональных полиномах В. А. Стеклова, Доклады Ак. Наук СССР, **83** (1952), S. 5—8.

[3] G. H. HARDY und J. E. LITTLEWOOD, Sur la série de Fourier d'une fonction à carré sommable, *C. R. Acad. Sci. Paris,* **156** (1913), S. 1307—1309.

[4] J. KOROUS, O rozwoji funkcí jedné reálne proměnné v řadu jistích orthogonálních polinomu. *Rozpravy Česke Akademie* (2), **48** (1938), S. 12.

[5] S.: SZEGŐ, loc. cit. [6], Theorem 3.1.3, S. 38.

[6] Formel 4, S. 36.

[7] Diese Schlußweise wurde für andere Zwecke zuerst von ERDŐS und TURÁN angewandt. (Vgl P. ERDŐS—P. TURÁN, On interpolation, III, *Annals of Math.,* **41** (1940), Lemma II. S. 524—525).

[5] J. LENSE, *Kugelfunktionen* (Leipzig, 1950).

[6] G. SZEGŐ, *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ., Vol. XXIII (1939).

[7] K. TANDORI, Über die Cesàrosche Summierbarkeit der orthogonalen Polynomreihen, *Acta Math. Acad. Sci. Hung.*, **3** (1952), S. 73—82.

[8] A. ZYGMUND, Un theorème sur les séries orthogonales, *Studia Math.*, **2** (1930), S. 181—182.

О СИЛЬНОЙ СУММАБИЛЬНОСТИ (C, 1) ОРТОГОНАЛЬНЫХ РЯДОВ МНОГОЧЛЕНОВ

Г. ФРАЙД (Будапешт)

(Резюме)

Пусть $p_0(x), p_1(x), \ldots, p_n(x), \ldots$ есть последовательность ортогональных многочленов на отрезке (a, b) относительно весовой функции $w(x) \geqq 0$ (степень $p_n(x)$ есть n), $f(x) \in L^2(w)$,

$$f(x) \sim \sum_{\nu=0}^{\infty} a_\nu p_\nu(x) \tag{I}$$

и

$$s_r(x) = \sum_{\nu=0}^{r} a_\nu p_\nu(x). \tag{II}$$

Доказываем следующие теоремы:

Теорема 1. Пусть при некотором x, где $a < x < b$,

$$\sum_{\nu=0}^{n} (p_\nu(x))^2 = O(n) \tag{III}$$

и

$$\int_x^{x+h} w(t)\,[f(t)-f(x)]^2\,dt = o(|h|); \tag{IV}$$

тогда

$$\lim_{n=\infty} \frac{1}{n+1} \sum_{r=0}^{n} |s_r(x)-f(x)| = 0. \tag{V}$$

Теорема 2. Пусть $a < \alpha < \beta < b$ и $w(x) > m > 0$, если $\alpha \leqq x \leqq \beta$, где m не зависит от x. Тогда (III) равномерно выполняется на любом отрезке внутри (α, β), откуда, использую теорему 1, следует, что (V) справедливо почти везде на отрезке (α, β).

Теорема 3. Если (III) справедливо во всех точках множества $E \in (a, b)$, то почти в каждой точке множества E

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{\nu=0}^{n} [s_\nu(x)-f(x)]^2 = 0.$$

Теорема 4. Если $a < \alpha < \beta < b$ и $M > w(x) > m > 0$, если $\alpha \leqq x \leqq \beta$, M и m не зависит от x, $f(x)$ непрерывна на отрезке (α, β), то арифметические среднии ряда (I) равномерно сходятся к $f(x)$ на любом отрезке лежащем внутри (α, β).

ÜBER DIE KONVERGENZ ORTHOGONALER POLYNOMREIHEN

Von

GÉZA FREUD (Budapest)

(*Vorgelegt von* G. ALEXITS)

Einleitung

Wir wollen zwei bekannte Sätze aus der Theorie der Fourierreihen, uzw. das Konvergenzkriterium von DIRICHLET—JORDAN und das Konvergenzkriterium von HARDY—LITTLEWOOD auf eine möglichst allgemeine Klasse von orthogonalen Polynomentwicklungen ausdehnen.

Es sollen $c_1, c_2, \ldots$ positive absolute Konstanten bedeuten. Es sei $\varrho(x)$ eine im Intervall $[-1, +1]$ definierte, L-integrierbare, nicht-negative Funktion, welche der Bedingung

$$\varrho(x) \leq c_1(1-x^2)^{-c_2} \tag{1}$$

genügt. Es sei $P_0(x), P_1(x), \ldots, P_n(x), \ldots$ die Folge der normierten orthogonalen Polynome,[1] die zur Gewichtsfunktion $\varrho(x)$ gehören:

$$\int_{-1}^{+1} \varrho(t) P_n(t) P_m(t)\, dt = \delta_{mn}. \tag{2}$$

Wir beschränken uns auf solche $\varrho(x)$, für welche die Orthogonalpolynome $P_n(x)$ im ganzen Intervall $[-1, +1]$, in n gleichmäßig der Ungleichung

$$|P_n(x)| < c_3(1-x^2)^{-c_4} \tag{3}$$

genügen.[2] *Sei $f(x)$ eine Funktion, die in $[-1, +1]$ mit $\varrho(x)$ multipliziert L-integrierbar ist; in einer kleinen einseitigen Umgebung der Stellen -1 und $+1$ sei sogar $\varrho(x)[f(x)]^2$ L-integrierbar.*[3] *Die Entwicklung dieser Funktion nach den $P_n(x)$ lautet:*

$$f(x) \sim \sum_{n=0}^{\infty} a_n P_n(x), \quad \text{wo} \quad a_n = \int_{-1}^{+1} \varrho(t) P_n(t) f(t)\, dt. \tag{4}$$

[1] $P_n(x)$ ist vom n-ten Grade, der Koeffizient von x^n in $P_n(x)$ ist positiv.

[2] Für die Folge der normierten Jacobischen Polynome $\{P_n^{(\alpha,\beta)}(x)\}$ mit $\varrho(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha > -1, \beta > -1$ sind (1) und (3) erfüllt. (Vgl. G. SZEGŐ [5].)

[3] Vgl. auch die Bemerkung am Ende dieser Arbeit.

Wir werden folgende Sätze beweisen:

SATZ A. *Vorausgesetzt wird* (1) *und* (3). *Bezeichne* (ξ_1, ξ_2) *ein inneres Teilintervall von* $[-1, +1]$; $f(x)$ *sei von beschränkter Variation in* (ξ_1, ξ_2) *und stetig an der Stelle* ξ, *wo* $\xi_1 < \xi < \xi_2$. *Dann konvergiert die Reihe* (4) *für* $x = \xi$ *gegen den Grenzwert* $f(\xi)$. *Ist ferner* $f(x)$ *im ganzen Intervall* (ξ_1, ξ_2) *stetig, so ist die Konvergenz in jedem inneren Teilintervall von* (ξ_1, ξ_2) *gleichmäßig.* (Verallgemeinerung des Kriteriums von DIRICHLET—JORDAN.)

SATZ B. *Vorausgesetzt wird* (1) *und* (3), *ferner sei* $\gamma > 0$, $a_n = O(n^{-\gamma})$ *und sei*

$$f(\xi + h) - f(\xi) = o\left(\frac{1}{\log 1/|h|}\right) \tag{5}$$

an der Stelle ξ $(-1 < \xi < +1)$ *für* $h \to 0$. *Dann konvergiert* (4) *an der Stelle* $x = \xi$ *gegen* $f(\xi)$. (Verallgemeinerung des Kriteriums von HARDY—LITTLEWOOD.)

Zum Beweis dieser Sätze werden Abschätzungen für Integrale der Form $\int_{x_1}^{x_2} \varrho(x) P_n(x) P_m(x)\, dx$ und eine Ausdehnung des Lokalisationssatzes von RIEMANN auf Systeme von Orthogonalpolynomen benutzt.

Hilfssätze

LEMMA I. *Aus* (1) *folgt, daß man zu jedem* ξ $(-1 < \xi < +1)$ *Polynome höchstens* $2k$*-ten Grades* $q_k(x, \xi)$, $Q_k(x, \xi)$ (*kurz*: $Q_k(x)$ *und* $q_k(x)$) *angeben kann, für welche in jedem inneren Teilintervall von* $(-1, +1)$ *in* ξ *gleichmäßig*

$$q_k(x) \leq \varphi(x, \xi) \leq Q_k(x); \quad \varphi(x, \xi) = \begin{cases} 1 & \text{für } -1 \leqq x < \xi, \\ 0 & \text{für } \xi < x \leqq +1 \end{cases} \tag{6}$$

und

$$\int_{-1}^{+1} \varrho(x) [Q_k(x) - q_k(x)]\, dx = O\left(\frac{1}{k}\right) \tag{7}$$

gilt.

BEWEIS. Einer bekannten Polynomkonstruktion von MARKOV [4], [3] folgend, gibt es Polynome $Q_k(x)$ und $q_k(x)$ von niedrigerem Grade als $2k$, welche (6) erfüllen und der Ungleichung

$$\int_{-1}^{+1} \varrho(x) [Q_k(x) - q_k(x)]\, dx < \lambda_{k\nu} + \lambda_{k,\nu+1}$$

genügen. Dabei bedeuten $\lambda_{k\nu}$ und $\lambda_{k,\nu+1}$ die in den zu ξ nächstliegenden Grundpunkten gebildeten Cotes'schen Zahlen der Gaussischen mechanischen Quadratur k-ter Ordnung mit der Gewichtsfunktion $\varrho(x)$. Wenden wir die im

Anhange der Arbeit [3] stehende Erdős—Turánsche Schlußweise mit

$$\pi_k(x)=\frac{1}{(k-l-1)^2}\zeta_{k-l-1}(x)\left(\frac{1-x^2}{1-x_{k\nu}^2}\right)^l,\quad l=\left[\frac{c_2}{2}\right]+1$$

an (vgl. [1]), so folgt

$$\lambda_{k\nu}=O\left(\frac{1}{k}\right)=\lambda_{k,\nu+1}.$$

LEMMA II. *Vorausgesetzt* (1) *und* (3), *gilt in jedem inneren Teilintervall* $[\alpha,\beta]$ *von* $(-1,+1)$, *falls* $\alpha\leqq\xi_1\leqq\xi_2\leqq\beta$, *die Abschätzung*

$$\int_{\xi_1}^{\xi_2}\varrho(x)P_n(x)P_m(x)\,dx=O\left(\frac{1}{|n-m|}\right) \tag{8}$$

in ξ_1 *und* ξ_2 *gleichmäßig.*

BEWEIS. Es sei

$$\varphi(x;\xi_1,\xi_2)=\begin{cases}1 \text{ für } \xi_1<x<\xi_2,\\ 0 \text{ für } x\leqq\xi_1 \text{ oder } x\geqq\xi_2,\end{cases} \tag{9}$$

ferner

$$R_k(x)=Q_k(x,\xi_2)-q_k(x,\xi_1),\quad r_k(x)=q_k(x,\xi_2)-Q_k(x,\xi_1). \tag{10}$$

Der Grad der Polynome $R_k(x)$ und $r_k(x)$ ist höchstens $2k$, und wegen (6) und (7) genügen sie den Ungleichungen

$$r_k(x;\xi_1,\xi_2)\leqq\varphi(x;\xi_1,\xi_2)\leqq R_k(x;\xi_1,\xi_2) \tag{11}$$

und

$$\int_{-1}^{+1}\varrho(x)\,[R_k(x)-r_k(x)]\,dx=O\left(\frac{1}{k}\right) \tag{12}$$

gleichmäßig in ξ_1 und ξ_2, wenn $\alpha\leqq\xi_1\leqq\xi_2\leqq\beta$. Wir können ohne Beschränkung der Allgemeinheit voraussetzen, daß $|n-m|>4c_2+4$ ist. Sei sodann k ganz mit

$$2k=|n-m|-2[2c_4]-3,\quad \text{oder}\quad 2k=|n-m|-2[2c_4]-4 \tag{13}$$

und wir bilden das Integral

$$\int_{\xi'}^{\xi''}\varrho(x)P_n(x)P_m(x)\,(1-x^2)^{[2c_4]+1}\,dx=$$

$$=\int_{-1}^{+1}\varrho(x)P_n(x)P_m(x)\,(1-x^2)^{[2c_4]+1}\,[\varphi(x;\xi',\xi'')-r_k(x;\xi',\xi'')]\,dx+$$

$$+\int_{-1}^{+1}\varrho(x)P_n(x)P_m(x)\,(1-x^2)^{[2c_4]+1}\,r_k(x;\xi',\xi'')\,dx,$$

$$(\alpha\leqq\xi'\leqq\xi''\leqq\beta).$$

Das zweite Integral an der rechten Seite verschwindet wegen der Orthogonalität der $P_n(x)$. Ist nämlich z. B. $n>m$, dann ist wegen (13) der Grad des

Polynoms $P_m(x)(1-x^2)^{[2c_4]+1} r_k(x;\xi',\xi'')$ niedriger als n. Das erste Integral auf der rechten Seite ist dem absoluten Werte nach — wegen (3), (6) und (7) — kleiner als

$$c_3^2 \int_{-1}^{+1} \varrho(x)\,[R_k(x;\xi',\xi'')-r_k(x;\xi',\xi'')]\,dx = O\left(\frac{1}{k}\right),$$

also ist

$$\int_{\xi'}^{\xi''} \varrho(x) P_n(x) P_m(x)\,(1-x^2)^{[2c_4]+1}\,dx = O\left(\frac{1}{k}\right) = O\left(\frac{1}{|n-m|}\right),$$

letzteres wegen (13).

Durch Anwendung des zweiten Mittelwertsatzes der Integralrechnung folgt endlich

$$\int_{\xi_1}^{\xi_2} \varrho(x)\,P_n(x)\,P_m(x)\,dx = (1-\xi_1^2)^{-[2c_4]-1} \int_{\xi_1}^{\xi_1'} \varrho(x)\,P_n(x) P_m(x)\,(1-x^2)^{[2c_4]+1}\,dx +$$

$$+(1-\xi_2^2)^{-[2c_4]-1} \int_{\xi_2'}^{\xi_2} \varrho(x)\,P_n(x)\,P_m(x)\,(1-x^2)^{[2c_4]+1}\,dx = O\left(\frac{1}{|n-m|}\right);$$

w. z. b. w.

Alle Abschätzungen gelten für $\alpha \leq \xi_1 \leq \xi_2 \leq \beta$, falls $-1<\alpha<\beta<+1$, in ξ_1 und ξ_2 gleichmäßig.

LEMMA III. *Für den Kern*

$$D_n(x,\xi) = \sum_{k=0}^{n-1} P_k(x) P_k(\xi) \tag{14}$$

des orthogonalen Polynomsystems $\{P_n(x)\}$ *gelten für* $-1<\alpha \leq x, \xi \leq \beta < +1$ *die Abschätzungen*

$$D_n(x,\xi) = O\left(\frac{1}{|x-\xi|}\right) \tag{15}$$

und

$$D_n(x,\xi) = O(n) \tag{16}$$

in x *und* ξ *gleichmäßig.*

BEWEIS. Nach der Christoffelschen Summenformel (vgl. [5], S. 42) ist

$$D_n(x,\xi) = \gamma_n \frac{P_n(x) P_{n-1}(\xi) - P_{n-1}(x) P_n(\xi)}{x-\xi}, \tag{17}$$

wobei [4]

$$0 < \gamma_n = \int_{-1}^{+1} \varrho(x)\,x\,P_n(x) P_{n-1}(x)\,dx \leq 1. \tag{18}$$

[4] Vgl. G. ALEXITS, *Commentarii Math. Helv.*, **16** (1944), S. 200—208.

Die Abschätzung (15) folgt also aus (3) und (17), während (16) aus (3) und (14) folgt.

Aus Lemma III ergibt sich

$$\int_{-1}^{+1} \varrho(x)|D_n(x,\xi)|\,dx = O(\log n). \tag{19}$$

Wir bemerken, daß (19) schon aus schwächeren Voraussetzungen folgt [6].

LEMMA IV. *Vorausgesetzt* (1) *und* (3), *gelten mit festen* x_1 *und* x_2 $(-1 < x_1 < \xi < x_2 < +1)$ *für ein beliebiges* $\eta > 0$ *die Abschätzungen*

$$\left| \int_{\xi+\frac{\eta}{n}}^{x_2} \varrho(x) D_n(x,\xi)\,dx \right| < \frac{c_5}{\eta} \quad \textit{für} \quad \xi + \frac{\eta}{n} \leqq x_2 \tag{20}$$

und

$$\left| \int_{x_1}^{\xi-\frac{\eta}{n}} \varrho(x) D_n(x,\xi)\,dx \right| < \frac{c_6}{\eta} \quad \textit{für} \quad x_1 \leqq \xi - \frac{\eta}{n}. \tag{21}$$

Ist $[\alpha, \beta]$ *ein inneres Teilintervall von* $(-1, +1)$ *und* $\alpha < x_1 < x_2 < \beta$, *so ist* c_5, c_6 *nur von* α *und* β *abhängig.*

BEWEIS. Nach (17) und dem zweiten Mittelwertsatz der Integralrechnung gilt

$$\int_{\xi+\frac{\eta}{n}}^{x_2} \varrho(x) D_n(x,\xi)\,dx = \gamma_n \frac{n}{\eta} \int_{\xi+\frac{\eta}{n}}^{x_2'} \varrho(x)[P_{n-1}(\xi)P_n(x) - P_n(\xi)P_{n-1}(x)]\,dx.$$

Somit folgt (20) aus (3), (18) und Lemma II mit $m=0$; analog beweist man (21).

Beweis des Lokalisationssatzes

Die Funktion $f(x)$ soll die Forderungen der Einleitung erfüllen, ferner sei im Intervall (ξ_1, ξ_2) fast überall $f(x) = 0$. Über $\varrho(x)$ soll nur vorausgesetzt werden, daß es in jedem inneren Teilintervall von $(-1, +1)$ beschränkt ist, und daß die Polynome $P_n(x)$ in jedem inneren Teilintervall von $(-1, +1)$ dem absoluten Wert nach unter einer von n unabhängigen Schranke bleiben. Dann konvergiert die Reihe (4) in jedem inneren Teilintervall von (ξ_1, ξ_2) gleichmäßig gegen den Grenzwert Null.

ε sei eine feste positive Zahl. Wir wählen h so klein, daß für ein inneres Teilintervall (ξ_1', ξ_2') von (ξ_1, ξ_2) für $\xi_1' \leqq \xi \leqq \xi_2'$

$$\int_{-1}^{-1+h} \varrho(x)\left[\frac{f(x)}{x-\xi}\right]^2 dx < \varepsilon^2, \quad \int_{1-h}^{1} \varrho(x)\left[\frac{f(x)}{x-\xi}\right]^2 dx < \varepsilon^2 \tag{22}$$

gleichmäßig in ξ gilt. Nach (3) ist die Polynomenfolge $P_n(x)$ in $[-1+h, 1-h]$ gleichmäßig beschränkt, und ihre Schranke sei $M(h)$. Nun können wir ein Polynom $\pi_\xi(x)$ finden, für welches

$$\int_{-1+h}^{1-h} \varrho(x) \left| \frac{f(x)}{x-\xi} - \pi_\xi(x) \right| dx < \frac{\varepsilon}{M(h)} \tag{23}$$

und

$$\int_{-1}^{-1+h} \varrho(x)[\pi_\xi(x)]^2\, dx < \varepsilon^2, \qquad \int_{1-h}^{1} \varrho(x)[\pi_\xi(x)]^2\, dx < \varepsilon^2 \tag{24}$$

gilt. Die Integrale an den linken Seiten von (22) und (24) sind in ξ stetig, also bleiben diese Abschätzungen mit demselben $\pi_\xi(x)$ gültig, falls ξ um seinen ursprünglichen Wert in einem kleinen Intervall variiert. Nach dem Heine—Borelschen Überdeckungssatze kann man also $\pi_\xi(x)$ für jedes innere abgeschlossene Teilintervall von (ξ_1, ξ_2) aus einer endlichen Folge von Polynomen $\pi_1(x), \pi_2(x), \ldots, \pi_\lambda(x)$ wählen. N sei eine positive ganze Zahl, die größer als der Grad der einzelnen Polynome $\pi_i(x)$ ist, und es sei die ganze Zahl $n > N$. Ist $s_n(\xi)$ die n-te Teilsumme der Reihe (4) für $x = \xi$, dann ergibt sich

$$s_n(\xi) = \int_{-1}^{+1} \varrho(x) D_n(x, \xi)\, [f(x) - (x-\xi)\pi_\xi(x)]\, dx =$$

$$= \gamma_n \int_{-1}^{+1} \varrho(x)\, [P_n(x) P_{n-1}(\xi) - P_{n-1}(x) P_n(\xi)] \left[\frac{f(x)}{x-\xi} - \pi_\xi(x) \right] dx = \tag{25}$$

$$= \int_{-1}^{-1+h} \ldots + \int_{-1+h}^{1-h} \ldots + \int_{1-h}^{1} \ldots$$

Der Integrand ist, ungeachtet beschränkter konstanten Faktoren, eine Summe von zwei Gliedern der Form $\varrho(x) P_n(x) \left[\frac{f(x)}{x-\xi} - \pi_N(x) \right]$. In $[1-h, 1]$ wird wegen (22) und (24)

$$\int_{1-h}^{1} \varrho(x) |P_n(x)| \left| \frac{f(x)}{x-\xi} - \pi_\xi(x) \right| dx \leqq$$

$$\leqq \left(\int_{1-h}^{1} \varrho(x) [P_n(x)]^2\, dx \right)^{1/2} \left(\int_{1-h}^{1} \varrho(x) \left[\frac{f(x)}{x-\xi} - \pi_\xi(x) \right]^2 dx \right)^{1/2} \leqq \tag{26}$$

$$\leqq \left(2 \int_{1-h}^{1} \varrho(x) \left[\frac{f(x)}{x-\xi} \right]^2 dx + 2 \int_{1-h}^{1} \varrho(x)\, [\pi_\xi(x)]^2\, dx \right)^{1/2} < 2\varepsilon$$

und eine ähnliche Abschätzung gilt für $\int_{-1}^{-1+h} \varrho(x)|P_n(x)|\left|\frac{f(x)}{x-\xi} - \pi_\xi(x)\right| dx$. Im Intervall $(-1+h, 1-h)$ gilt wegen (23) und $|P_n(x)| < M(h)$

$$\int_{-1+h}^{1-h} \varrho(x)\,|P_n(x)|\left|\frac{f(x)}{x-\xi} - \pi_\xi(x)\right| dx < \varepsilon. \tag{27}$$

Aus (25), (26), (27) folgt

$$\lim_{n=\infty} s_n(\xi) = 0$$

gleichmäßig in jedem inneren Teilintervall (ξ_1', ξ_2') von (ξ_1, ξ_2), w. z. b. w.

Beweis des Satzes A

Aus (1) und (16) ergibt sich, falls $\xi_1 < \xi - \frac{\eta}{n} < \xi + \frac{\eta}{n} < \xi_2$,

$$\left|\int_{\xi-\frac{\eta}{n}}^{\xi+\frac{\eta}{n}} \varrho(x)\,D_n(x,\xi)\,dx\right| < c_7\eta. \tag{28}$$

Nach dem Lokalisationssatze wird der wesentliche Teil des Restgliedes $s_n(\xi)-f(\xi)$ gleich

$$\int_{\xi_1}^{\xi_2} \varrho(x) D_n(x,\xi)\,[f(x)-f(\xi)]\,dx = \int_{\xi_1}^{\xi-\frac{\eta}{n}} \ldots + \int_{\xi-\frac{\eta}{n}}^{\xi+\frac{\eta}{n}} \ldots + \int_{\xi+\frac{\eta}{n}}^{\xi_2} \ldots \tag{29}$$

Bezeichnen wir mit $\varphi\left(\frac{\eta}{n}\right)$ das Maximum von $|f(x)-f(\xi)|$ für $|x-\xi| < \frac{\eta}{n}$ und wählen wir η als Funktion von n derart, daß mit $n \to \infty$ auch $\eta \to \infty$, aber $\eta/n \to 0$ und $\eta\varphi\left(\frac{\eta}{n}\right) \to 0$.[5] Wegen (28) gilt dann

$$\left|\int_{\xi-\frac{\eta}{n}}^{\xi+\frac{\eta}{n}} \varrho(x)\,D_n(x,\xi)\,[f(x)-f(\xi)]\,dx\right| < c_7\,\eta\varphi\left(\frac{\eta}{n}\right) \to 0.$$

Ist die Variation von $f(x)$ in (ξ_1, ξ_2) gleich V, und das Maximum von $|f(x)-f(\xi)|$ in (ξ_1, ξ_2) gleich M, so folgt aus dem zweiten Mittelwertsatze der

[5] Es sei etwa $\eta_n = \mathrm{Min}\,(n^{1/2}, \{\varphi(n^{-1/2})\}^{-1/2})$.

Integralrechnung und Lemma IV:

$$\left|\int_{\xi_1}^{\xi-\frac{\eta}{n}} \varrho(x) D_n(x,\xi)[f(x)-f(\xi)]\,dx\right| < (M+V)\frac{c_5}{\eta} \to 0,$$

und eine ähnliche Abschätzung gilt für $\int\limits_{\xi+\frac{\eta}{n}}^{\xi_2} \ldots$ in (29).

Beweis des Satzes B

Wir zerlegen den wesentlichen Teil des Restgliedes folgenderweise:

$$(30)\qquad \int_{\xi_1}^{\xi_2} \varrho(x) D_n(x,\xi)[f(x)-f(\xi)]\,dx = \int_{\xi_1}^{\xi-n^{-\gamma/2}} \ldots + \int_{\xi-n^{-\gamma/2}}^{\xi+n^{-\gamma/2}} \ldots + \int_{\xi+n^{-\gamma/2}}^{\xi_2} \ldots$$

Wegen (5) gilt in $(\xi-n^{-\gamma/2}, \xi+n^{-\gamma/2})$

$$f(x)-f(\xi) = o\left(\frac{1}{\log n}\right),$$

also strebt wegen (19) das mittlere Glied von (30) gegen Null. Wegen (3), (17) und (18) ist das erste Glied in (30), abgesehen von einem beschränkten Faktor, die Summe von zwei Gliedern der Form

$$(31)\qquad \int_{\xi_1}^{\xi-n^{-\gamma/2}} \varrho(x) P_n(x)[f(x)-f(\xi)]\frac{dx}{x-\xi} = n^{\gamma/2} \int_{x_1}^{\xi-n^{-\gamma/2}} \varrho(x) P_n(x)[f(x)-f(\xi)]\,dx.$$

Nach Satz A ist die Orthogonalpolynomreihe

$$(32)\qquad \varphi_n(x) = \sum_{k=0}^{\infty} b_{kn} P_k(x)$$

der Funktion

$$\varphi_n(x) = \begin{cases} P_n(x) & \text{für } x_1 < x < \xi - n^{-\gamma/2}, \\ 0 & \text{für } x \leqq x_1 \text{ oder } x \geqq \xi - n^{-\gamma/2} \end{cases}$$

gleichmäßig konvergent in jedem inneren Teilintervall von $(-1, +1)$, welches die Stellen x_1 und $\xi - n^{-\gamma/2}$ nicht enthält. Ferner folgt aus Lemma II:

$$(33)\qquad b_{kn} = \int_{\xi_1}^{\xi-n^{-\gamma/2}} \varrho(x) P_n(x) P_k(x)\,dx = O\left(\frac{1}{|k-n|}\right).$$

Aus dem Beweisgang des Satzes A ergibt sich, daß die Teilsummen der Reihe (5) für eine Funktion $F(x)$ beschränkter Variation (angewendet mit $F(x) = \varphi_n(x)$) in jedem inneren Teilintervall von (ξ_1, ξ_2) gleichmäßig beschränkt bleiben. Daher ist die mit $f(x)-f(\xi)$ multiplizierte Reihe (32) nach dem

Lebesgueschen Satze gliedweise integrierbar:

$$\int_{x_1}^{\xi-n^{-\gamma/2}} \varrho(x) P_n(x)[f(x)-f(\xi)]\,dx = \sum_{k=0}^{\infty} a_k b_{kn} = \tag{34}$$
$$= O\left(\frac{1}{n^\gamma} + \sum_{k\neq n}{}' \frac{1}{k^\gamma |n-k|}\right) = O\left(\frac{\log n}{n^\gamma}\right).$$

Wegen (31) und (34) hat das erste Glied auf der rechten Seite von (30) den Grenzwert Null, und das dritte Glied kann ebenso behandelt werden.

BEMERKUNG. Die Bedingung, daß in der Ungebung von $+1$ und -1 auch $\varrho(x)[f(x)]^2$ integrierbar sei, wurde nur im Beweise des Lokalisationssatzes benutzt. Es kann durch jede Forderung ersetzt werden, welche die Lokalisation sichert.

Der Lokalisationssatz gilt auch unter folgenden Bedingungen: Es sei

$$|P_n(x)| < K(x), \quad -1 \leq x \leq +1, \quad n = 0, 1, 2, \ldots \tag{35}$$

und zusammen mit $\varrho(x)$ sei auch $\varrho(x)K(x)$ in $(-1, +1)$ integrierbar. Endlich genüge die Funktion $f(x)$, die in (ξ_1, ξ_2) fast überall verschwindet, der Bedingung, daß $\varrho(x)K(x)f(x)$ in $(-1, +1)$ L-integrierbar ist.

Um den Lokalisationssatz zu beweisen, suchen wir ein Polynom $\pi_\xi(x)$, welches der Ungleichung

$$\int_{-1}^{+1} \varrho(x) K(x) \left| \frac{f(x)}{x-\xi} - \pi_\xi(x) \right| dx < \varepsilon \tag{36}$$

genügt. Hieraus folgt die Gültigkeit von (27) auch für $h=0$, woraus $\lim_{n=\infty} s_n(\xi) = 0$ folgt. Die Begründung der Gleichmäßigkeit erfolgt, wie vorher.

Herrn Prof. G. ALEXITS bin ich für wertvolle Ratschläge und sein förderndes Interesse zu Dank verpflichtet.

(*Eingegangen am 21. März 1952.*)

Literaturverzeichnis

1. P. ERDŐS and P. TURÁN, On interpolation, III. *Annals of Math.*, **41** (1940), S. 510—553.

2. L. FEJÉR, Mechanische Quadraturen mit positiven Cotesschen Zahlen, *Math. Zeitschrift*, **37** (1933). S. 287—309.

3. G. FREUD, Restglied eines Tauberschen Satzes, I: *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 299—308.

4. А. А. Марков, Доказательство некоторых неравенств П. Л. Чебышева, Избранные Труды (Москва, 1948), S. 20-21.

5. G. SZEGŐ, *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ., Vol. XXIII.

6. B. SZ.-NAGY, Über die Konvergenz von Reihen orthogonaler Polynome, *Math. Nachrichten*, **4** (1950—51), S. 50—55.

О СХОДИМОСТИ РЯДОВ ОРТОГОНАЛЬНЫХ МНОГОЧЛЕНОВ

Г. ФРАЙД (Будапешт)

(Резюме)

Пусть $\varrho(x) \geqq 0$ L-интегрируемая функция на отрезке $[-1, +1]$,

$$\varrho(x) \leqq c_1(1-x^2)^{-c_2}$$

и, если $P_0(x), P_1(x), \ldots, P_n(x), \ldots$ последовательность ортогональных многочленов относительно весовой функции $\varrho(x)$, то пусть

$$|P_n(x)| < c_3(1-x)^{-c_4'}(1+x)^{-c_4''}.$$

Пусть измеримая функция $f(x)$ удовлетворяет одному из следующих условий:

А) На отрезке $[-1, +1]$, $f(x) \in L_1(\varrho)$ и в небольшой окрестности точек -1 и $+1$ $f \in L_2(\varrho)$.

Б) На отрезке $[-1, +1]$, $f(x) \in L_1(\varrho(x)(1-x)^{-c_4'}(1+x)^{-c_4''})$.

Тогда для разложения $f(x)$ справедлива локализационная теорема Р и м а н н а.

Если $f(x)$ функция с ограниченным изменением, то её ряд ортогоналных многочленов сходится во всех точках непрерывности ξ функции $f(x)$ к $f(\xi)$ (критерий Д и р и х л е—Ж о р д а н а).

Наконец, справедлив и критерий Х а р д и—Л и т т л ь в у д а, т. е. если

$$a_n = O(n^{-\alpha}) \quad \text{и} \quad f(\xi+h)-f(\xi) = o\left(\frac{1}{\log 1/|h|}\right),$$

то ряд ортогональных многочленов функции $f(x)$ в точке ξ сходится к $f(\xi)$.

EINE ELEMENTARE AUFLÖSUNG DER DIOPHANTISCHEN GLEICHUNG $x^3+1=2y^2$

Von

WILHELM LJUNGGREN (Bergen, Norwegen)

(*Vorgelegt von* A. Rényi)

In einer früheren Arbeit[1] habe ich den folgenden Satz bewiesen: Es sei D eine natürliche und quadratfreie Zahl und weiter $D \not\equiv 0 \pmod 3$ und ohne Primteiler der Form $6t+1$. Dann haben die Gleichungen

(1) $x^3 \pm 1 = Dy^2, \quad x^3 \pm 8 = Dy^2, \quad x^3 \pm 1 = 3Dy^2$ und $x^3 \pm 8 = 3Dy^2$

zusammen höchstens eine Lösung in ganzen, positiven und teilerfremden Zahlen x und y, von den Fällen $D=1$ und $D=2$ abgesehen. In den Ausnahmefällen sind die folgenden Gleichungen lösbar:

$$x^3+1=y^2 \quad \text{mit} \quad x=2,\ y=3,$$
$$x^3+8=y^2 \quad \text{mit} \quad x=1,\ y=3,$$
$$x^3-8=3y^2 \quad \text{mit} \quad x=11,\ y=21,$$
$$x^3+1=2y^2 \quad \text{mit} \quad x=1,\ y=1 \quad \text{und} \quad x=23,\ y=78.$$

Der Beweis ist aber nicht elementar. Der Zweck dieses kleinen Aufsatzes ist eine *elementare* Auflösung der Gleichung

$$x^3+1=2y^2 \tag{2}$$

zu geben.[2] In seiner Arbeit über die ähnliche Gleichung $x^3-1=2y^2$ hebt R. Obláth hervor,[3] daß man mit seiner Methode die Gleichung (2) nicht behandeln kann. Die Gleichung (2) kann auch so geschrieben werden:

$$(x+1)(x^2-x+1)=2y^2.$$

Nun ist $x^2-x+1 \equiv (x+1)^2-3(x+1)+3$, darum kann der größte gemein-

[1] Einige Sätze über unbestimmte Gleichungen von der Form $Ax^4+Bx^2+C=Dy^2$, *Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo,* I. Math. Naturv. Klasse, **1942**, No. 9, S. 49.

[2] Diese Gleichung ist erwähnt in L. E. Dickson, *History of the theory of numbers,* Bd. II (New York, 1934), S. 630. In meiner zitierten Arbeit [1] gab ich die erste Lösung dieser Gleichung.

[3] R. Obláth, Über die diophantische Gleichung $x^3-1=2y^2$, *Acta Math. Acad. Sci. Hung.,* **1** (1950), S. 113—117 und 321—322.

same Teiler von x^2-x+1 und $x+1$ nur gleich 1 oder 3 sein. Im ersten Falle bekommen wir

$$x+1=2u^2;\quad x^2-x+1=v^2.$$

Dies gibt die triviale Lösung $x=1,\ y=1$. Im zweiten Falle erhalten wir

$$x+1=6u^2;\quad x^2-x+1=3v^2.$$

Wird x aus diesen beiden Gleichungen eliminiert, ergibt sich

$$4v^2-3(4u^2-1)^2=1. \tag{3}$$

Der Wert $u=0$ gibt die triviale Lösung $x=-1,\ y=0$. Für $u\neq 0$ sind die Lösungen der Pellschen Gleichung (3) durch die folgende Formel gegeben:

$$2v+(4u^2-1)\sqrt{3}=(2+\sqrt{3})^n \qquad (n=1,3,5,7,\ldots). \tag{4}$$

Zuerst beweisen wir, daß $n\equiv 3 \pmod 4$ sein muß. Aus (4) folgt

$$4u^2-1=\sum_{i=0}^{\frac{n-1}{2}}\binom{n}{2i+1}2^{n-2i-1}3^i\equiv(-1)^{\frac{n-1}{2}} \pmod 4,$$

d. h. $n=4t+3$. Mit den Bezeichnungen $\varepsilon=2+\sqrt{3}$, $\varepsilon'=2-\sqrt{3}$, wo also $\varepsilon\varepsilon'=1$ ist, ergibt sich

$$4u^2=\frac{\varepsilon^n-\varepsilon'^n}{\varepsilon-\varepsilon'}+1=\frac{\varepsilon^{\frac{n+1}{2}}-\varepsilon'^{\frac{n+1}{2}}}{\varepsilon-\varepsilon'}\left(\varepsilon^{\frac{n-1}{2}}+\varepsilon'^{\frac{n-1}{2}}\right),$$

oder

$$4u^2=\frac{\varepsilon^{2t+2}-\varepsilon'^{2t+2}}{\varepsilon^2-\varepsilon'^2}\cdot\frac{\varepsilon^{2t+1}+\varepsilon'^{2t+1}}{\varepsilon+\varepsilon'}\cdot(\varepsilon+\varepsilon')^2,$$

oder endlich

$$\left(\frac{u}{2}\right)^2=\frac{\varepsilon^{2t+2}-\varepsilon'^{2t+2}}{\varepsilon^2-\varepsilon'^2}\cdot\frac{\varepsilon^{2t+1}+\varepsilon'^{2t+1}}{\varepsilon+\varepsilon'}. \tag{5}$$

Die beiden Faktoren der rechten Seite in (5) sind prim zueinander. Dies ergibt sich aus den leicht zu bestätigenden Identitäten

$$2\cdot\frac{\varepsilon^{2t+2}-\varepsilon'^{2t+2}}{\varepsilon^2-\varepsilon'^2}-\frac{\varepsilon^{2t+1}+\varepsilon'^{2t+1}}{\varepsilon+\varepsilon'}=\frac{\varepsilon^{2t+1}-\varepsilon'^{2t+1}}{\varepsilon-\varepsilon'}, \tag{6}$$

$$\frac{(\varepsilon+\varepsilon')^2}{2}\cdot\left(\frac{\varepsilon^{2t+1}+\varepsilon'^{2t+1}}{\varepsilon+\varepsilon'}\right)^2-\frac{(\varepsilon-\varepsilon')^2}{2}\left(\frac{\varepsilon^{2t+1}-\varepsilon'^{2t+1}}{\varepsilon-\varepsilon'}\right)^2=1. \tag{6'}$$

Aus (5) erhalten wir somit

$$\frac{\varepsilon^{2t+2}-\varepsilon'^{2t+2}}{\varepsilon^2-\varepsilon'^2}=p^2,\qquad \frac{\varepsilon^{2t+1}+\varepsilon'^{2t+1}}{\varepsilon+\varepsilon'}=q^2.$$

Wird dies in (6') eingesetzt, bekommen wir

$$4q^4-3r^2=1. \tag{7}$$

Die einzige Lösung dieser Gleichung in ganzen, positiven Zahlen ist $q=r=1$. Aus (7) folgern wir nämlich $2q^2+1=3s^2$, $2q^2-1=t^2$, woraus $4q^2-t^2=3s^2$, d. h. $2q\pm t=3h^2$, $2q\mp t=k^2$. Hieraus ergibt sich weiter

$4q = 3h^2 + k^2$, $s = hk$. Werden diese Werte von q und s in $2q^2 + 1 = 3s^2$ eingesetzt, bekommen wir

$$9h^4 - 18h^2k^2 + k^4 = -8.$$

Diese Gleichung kann auch so geschrieben werden:

$$k^4 - 2\left(\frac{3h^2 - 3k^2}{2}\right)^2 = 1.$$

Hieraus schließen wir $h^2 = k^2 = 1$. Die Gleichung $x^4 - 2y^2 = 1$, die mit dem folgenden System äquivalent ist: $x^2 + 1 = 2a^2$, $x^2 - 1 = 16b^2$, besitzt augenscheinlich nur die triviale Lösung $x^2 = 1$, $y = 0$. Damit haben wir bewiesen, daß $t = 0$, d. h. $n = 3$ ist. Wir erhalten so die einzige Lösung $x = 23$, $y = \pm 78$.

Also gilt der folgende Satz:

Die diophantische Gleichung $x^3 + 1 = 2y^2$ besitzt nur die folgenden Lösungen in ganzen rationalen Zahlen: $x = -1$, $y = 0$; $x = 1$, $y = \pm 1$ und $x = 23$, $y = \pm 78$.

(Eingegangen am 13. April 1952.)

ЭЛЕМЕНТАРНОЕ РЕШЕНИЕ ДИОФАНТОВОГО УРАВНЕНИЯ $x^3 + 1 = 2y^2$

В. ЛЮНГРЕН (Берген)

(Резюме)

Автор элементарными методами доказывает следующую теорему: диофантовое уравнение $x^3 + 1 = 2y^2$ имеет лишь следующие решения среди рациональных целых чисел: $x = -1$, $y = 0$; $x = 1$, $y = \pm 1$ и $x = 23$, $y = \pm 78$.

ON SUBDIRECT UNIONS, I

By

L. FUCHS (Budapest)

(*Presented by* G. HAJÓS)

§ 1. Introduction

During the last decade the notion of subdirect union [1] has proved to be one of the most important and fruitful notions in abstract algebra. Indeed, a great number of significant problems from quite different fields of abstract algebra have found their solutions in terms of the notion of subdirect union. It turned out that in several cases without any hope of getting structure theorems by direct decompositions, subdirect unions may be used with great success. For example, let us mention N. JACOBSON's structure theorem [2] generalizing the classical Wedderburn—Artin structure theorems from rings with minimal condition and no nilpotent ideals different from 0 to arbitrary rings with zero Jacobson radical, or the well-known result in the theory of infinite Abelian groups, according to which a torsion free Abelian group of finite rank n is a subdirect sum of n groups of rank one, each of which is isomorphic to some subgroup of the additive group of the rational numbers,[3] etc. The power of this concept may be judged even from the fact, proved by G. BIRKHOFF,[4] which states that any algebraic structure [5] may be represented as a subdirect union of subdirectly irreducible algebraic structures.

However, there is a great difference between the direct decompositions on the one hand, and the representations as subdirect unions on the other hand.

If an algebraic structure G can be represented as the direct union of its substructures:[6] $G = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_n$, then the structure of G may be

[1] For the definition see § 3 below.

[2] JACOBSON [8]; cf. McCOY [9], and BROWN and McCOY [5]. (The numbers in brackets refer to the Bibliography given at the end of the paper.)

[3] See e. g. BAER [1].

[4] See BIRKHOFF [2], and also [3], pp. 91—92.

[5] By an *algebraic structure* we mean an arbitrary system with binary operations satisfying certain laws and possibly having operator domains. (Cf. BOURBAKI's "structure algèbrique".)

[6] For the sake of simplicity we restrict ourselves only to the case when the number of the components in the direct union is finite. We denote direct unions by the sign $\dotplus$, while $+$ is reserved for subdirect unions.

considered to be reduced to those of $A_1, A_2, \ldots, A_n$. This consideration may be justified by the fact that if we are given any structures $A_1^*, A_2^*, \ldots, A_n^*$ with A_i^* isomorphic to A_i, then by an "outer" construction of a direct union we reobtain G; in fact, G is determined, uniquely up to an isomorphism, as the totality of all vectors $(a_1^*, a_2^*, \ldots, a_n^*)$ $(a_i^* \in A_i^*)$ with componentwise operations.

Quite different is the situation in the case of subdirect unions. BIRKHOFF's structure theorem on subdirect unions[4] asserts that if $\theta_1, \theta_2, \ldots, \theta_n$ are congruence relations[7] in a given algebraic structure G such that $\theta_1 \cap \theta_2 \cap \cdots \cap \theta_n = 0$, then taking for the components of a subdirect union representation of G the structures A_i of the residue classes modulo θ_i, we obtain an isomorphism $a \longleftrightarrow (a_1, a_2, \ldots, a_n)$ between G and a subdirect union of the A_i's, where a_i denotes the residue class containing a. This theorem shows that if G is *given* and congruence relations θ_i may be found in G with the mentioned property, then a subdirect union representation can be formed for G. However, it should be observed that BIRKHOFF's result is based on an "inner" construction of the subdirect union: he works in terms of relations *in* G whose structure is not *a priori* known, and therefore, it does not tell us anything how this subdirect union can be obtained by an "outer" construction from the A_i's *without knowledge of* G, i. e. by using only relations in the A_i's (and not in G). This is the reason why subdirect unions can not be regarded as satisfactory as direct unions.

The problem is now obvious: *find a method completing* BIRKHOFF*'s subdirect union construction (which may be characterized by the attribute "analytical") to a "synthetical" one: to a construction based solely on the components* A_i.

In the literature I have failed to find the statement of this problem, although it seems to be fairly natural. It can be considered as a parallel one to that raised and solved by O. SCHREIER:[8] to construct a group G if its normal subgroup N and the factor group $G/N \simeq F$ are given. SCHREIER's solution is synthetical in the sense that he succeeds in describing the structure of G in terms of relations in N and F only. (The corresponding analytical construction would be the description of the coset expansion of G modulo N.)

The present paper is devoted to an approach to the solution of the stated problem. The main difficulty in discussing subdirect unions lies in the fact that the components are no longer substructures as in the case of direct unions: they are quotient structures whose knowledge does not imply immediately a construction for the subdirect union.

In what follows we shall confine ourselves to subdirect unions with two components only. In § 3 we shall give a "synthetical" method for constructing

[7] A congruence relation is an equivalence relation $x \equiv y$ (θ) with the substitution property for all operations defined in the algebraic structure.

[8] See e. g. ZASSENHAUS [17], p. 94.

subdirect unions and in § 4 we show that it is the most general one for certain algebraic structures. These structures are the most important ones:[9]

I) *groups*, written additively (nevertheless, with this notation we do not intend to exclude the non-commutative groups), possibly with an operator domain Ω;

II) *rings*, or *algebras* over arbitrary fields;

III) *Boolean algebras*, i. e. relatively complemented distributive lattices with 0 and 1.

For brevity, we shall quote these structures simply as *structures of type* I, II *and* III, respectively.

That our construction method for subdirect unions given in § 3 is not of general validity for arbitrary structures, will be shown by an instance from lattice theory.

As a consequence of our discussions we shall establish some isomorphism statements connected with subdirect unions, and in § 7 an interesting characterization of the subgroups of the direct sum of Abelian groups will be obtained. § 8 is devoted to illustrative examples from the theory of groups. The paper closes with an observation concerning the isomorphism of the subdirect union of two structures.

Finally, let us remark the following. Our main theorem states that each subdirect union of two structures A and B of type I—III may be obtained by "identifying" certain isomorphic quotient structures of A and B. Therefore, it seems to me that the present question has a deep analogy with O. SCHREIER's problem of free products of groups with amalgamated subgroups.[10] Indeed, he studied free products of groups such that certain isomorphic subgroups of them are identified. In view of this analogy, we are allowed to say that subdirect unions are nothing else than *direct unions with amalgamated quotient structures.*

§ 2. The *W*-homomorphism

M. WEDDERBURN[11] has introduced a new notion of homomorphism as a many-many correspondence between the elements of two groups A and B. We shall use this homomorphism in what follows. It has the advantage over the usual one that it is symmetrical in A and B. The property of this homomorphism which will be used here for its definition, was stated as a theorem

[9] We laid no stress on xtending the domain of validity of our results as far as possible. — Let us mention that the algebraic structures for which the proof of our results presents no difficulties and to which we are restricting ourselves, are all such that their homomorphisms are uniquely determined by the set of elements mapped upon 0. (Of course, we might consider Boolean algebras as a special type of ring.)

[10] SCHREIER [13], and in a generalized form, NEUMANN [10].

[11] See WEDDERBURN [16]. Cf. also DUBREIL—JACOTIN [6] and DUBREIL [7].

by WEDDERBURN in his cited paper.[12] Of course, it is easy to extend this notion to more general structures than groups.

Let A and B be two algebraic structures of the same type, i. e. both are additive or multiplicative groups, rings, lattices etc. with the additional condition that in case they have operator domains Ω_A and Ω_B, then these coincide: $\Omega_A = \Omega_B = \Omega$.

If the structures A and B of the same type may be mapped (operator-) homomorphically onto a third structure F of the same type, then we say that A and B are *W-homomorphic with respect to F,* in sign

$$A \overset{F}{\sim} B.$$

It is immediate that

$$A \overset{B}{\sim} B$$

means the usual homomorphism $A \sim B$ of A onto B. Therefore, the W-homomorphism may be considered as a proper extension of the common notion of homomorphism.

We also note that any two structures A and B of the same type may be regarded as W-homomorphic, viz. with respect to the zero structure consisting of a single element 0:

$$A \overset{0}{\sim} B.$$

We shall call this W-homomorphism *trivial.*

If one of the two structures A and B, say B, is simple in the sense that it has no homomorphisms other than isomorphisms and a homomorphism onto the zero structure, then between A and B only one kind of non-trivial W-homomorphisms may exist, viz. $A \overset{B}{\sim} B$.

§ 3. Subdirect unions

Let $G = A \dotplus B$ be a *subdirect union* of the structures A and B of the same type; G may be defined by the following axioms:

(i) *G consists of certain pairs (a, b), $a \in A$ and $b \in B$, with the following definition of equality: $(a, b) = (a', b')$ if and only if $a = a'$ and $b = b'$.*

(ii) *For each $a \in A$ there exists at least one $b \in B$ such that $(a, b) \in G$ and for each $b \in B$ there is at least one $a \in A$ with the same property.*

If A and B are regarded merely as sets and G satisfies (i) and (ii), we may say that the set G is *a subdirect union of the sets A and B.* In order to make G into an algebraic structure with the same operations as those given in A and B, we have to impose on G the following conditions too:

(iii) *If $(a, b) \in G$ and $(a', b') \in G$, and if $\circ$ denotes an arbitrary operation in A and B of the same type* (addition, multiplication, meet or join, or taking

[12] The present definition is simpler than that used by WEDDERBURN, but has the fault of beauty that it is based on the usual concept of homomorphism. For the present purpose we prefer the definition below. (T. SZELE has proposed to the author to do this.)

an inverse operation if it exists, etc.), [13] *then let*

$$(a \circ a', b \circ b') \in G$$

and

$$(a, b) \circ (a', b') = (a \circ a', b \circ b').$$

In other words, the operations in G should be performed componentwise just as in the case of direct unions.

(iv) *If θ is any operator ($\in \Omega$) acting on A and B, then together with $(a, b) \in G$ let*

$$(a^\theta, b^\theta) \in G$$

and

$$(a, b)^\theta = (a^\theta, b^\theta).$$

Of course, two structures A and B of the same type have, in general, several non-isomorphic subdirect unions. Evidently, each of them is a substructure of the direct union $A \dotplus B$. Therefore, for the operations in G the same axioms hold as those in the components.

§ 4. Construction of subdirect unions

Here we shall give a method for constructing subdirect unions of two structures A and B of the same type. In the cases I—III this method will prove to be the general one.

Let A and B be two structures of the same type and suppose that A and B are W-homomorphic with respect to a third structure F of the same type:

$$A \overset{F}{\sim} B.$$

We shall prove the following

THEOREM 1. *Let G be the set of all pairs (a, b), $a \in A$ and $b \in B$, with the property that under certain fixed homomorphisms $A \sim F$ and $B \sim F$ the elements $a \in A$ and $b \in B$ are mapped upon the same element $f \in F$. Then G is a subdirect union of A and B with the same operations which are performed according to the rule*

$$(a, b) \circ (a', b') = (a \circ a', b \circ b')$$

where $\circ$ denotes an arbitrary operation in A and B of the same type.

Indeed, (ii) is fulfilled, for each $a \in A$ is mapped upon some $f \in F$ and each $f \in F$ is an image of some $b \in B$, and dually. The operations can always be performed in G, for if $a \in A$, $b \in B$ are mapped upon $f \in F$, and $a' \in A$, $b' \in B$ are mapped upon $f' \in F$, then by the fundamental property of homomorphisms we obtain that both $a \circ a' \in A$ and $b \circ b' \in B$ are mapped upon $f \circ f' \in F$ for all

[13] There is no loss of generality in denoting corresponding operations by the same sign.

$\circ$ which are meaningful in A and B; therefore, $(a \circ a', b \circ b') \in G$. The same reasoning applies for proving axiom (iv).

We call the reader's attention to the fact that the operation $\circ$ may be also an inverse operation and therefore the possibility of the inverse operations in G needs no special verification.

The properties of the operations in G, such as associativity, distributivity etc., are immediate consequences of those in A and B. This completes the proof of the theorem.

It is obvious that if $F = 0$, then the construction described in Theorem 1 yields the direct union of A and B.

We point out the interesting fact that in certain cases the subdirect union is not a proper one: it may be isomorphic to one of the components. In fact, in case F is isomorphic, say, to B, i. e. $A \overset{B}{\simeq} B$, then in the above construction, with each $a \in A$ there is associated a unique $b \in B$, viz. such that $(a, b) \in G$, and therefore the resulting subdirect union is isomorphic to A under the mapping $(a, b) \leftrightarrow a$.

It is worth while noticing already here that the subdirect union is not uniquely determined by the structures A, B, F: there may exist several non-isomorphic subdirect unions to the same given structures A, B, F with $A \overset{F}{\simeq} B$. Indeed, G depends also on how A and B are mapped homomorphically onto F. For examples we refer to § 9.

§ 5. The structure theorem on subdirect unions

We shall now prove that the method given in the preceding section is the most general one, i. e. all subdirect unions arise in the way described in Theorem 1, provided that the structures are of one of the types I—III.

THEOREM 2. *If G is a subdirect union of the structures A and B of type* I *or* II *or* III, *then there is a structure F of the same type such that A and B are W-homomorphic with respect to F, and G consists of all pairs (a, b) where $a \in A$ and $b \in B$ have the same image under the homomorphisms $A \sim F$ and $B \sim F$.*

For the proof we define congruence relations α in A and β in B as follows.

Put

$$a \equiv a' \ (\alpha)$$

if and only if there is a $b \in B$ such that $(a, b) \in G$ and $(a', b) \in G$; and

$$b \equiv b' \ (\beta)$$

if and only if there is an $a \in A$ with $(a, b) \in G$ and $(a, b') \in G$.

By symmetry, it is enough to consider the relation α only.

The relation α is by definition symmetric and is by (ii) reflexive too. Transitivity will follow from the important observation:

LEMMA. *If* $(a, b) \in G$ *and* $a \equiv a'$ (α), *then* $(a', b) \in G$, *i. e. congruent elements are paired with the same elements.*

For the proof, by hypothesis, we assume that there is a $b' \in B$ with $(a, b') \in G$ as well as $(a', b') \in G$. We have to separate the cases I, II from III.

For groups and rings, the element[14]

$$(a, b) - (a, b') + (a', b') = (a', b)$$

must belong to G.

For Boolean algebras, the element[15]

$$[(a, b) \cap (\overline{a, b'})] \cup (a', b') = ((a \cap \bar{a}) \cup a', (b \cap \bar{b}') \cup b') = (a', b \cup b')$$

must belong to G. Similarly we obtain the inclusion relation $(a', b \cap b') \in G$. Hence on account of

$$(a', b \cap b') \leqq (a', b') \leqq (a', b \cup b') \qquad (\text{all in } G)$$

(a', b) is the relative complement[16] of (a', b') with respect to $(a', b \cap b')$ and $(a', b \cup b')$. Since relative complement has to exist in G, we conclude that $(a', b) \in G$, in fact.

Our lemma is thus established.

Before proceeding we observe that our crucial lemma in general fails to hold in case the hypothesis of A and B being of type I or II or III is omitted. Indeed, if e. g. A and B are chains (i. e. simply ordered lattices), each with three elements:

$$A:\ 1 < 2 < 3 \qquad \text{and} \qquad B:\ a < b < c,$$

then for the following subdirect union our lemma as well as Theorem 2 are no more valid:

$$G:\ (1, a) < (2, a) < (2, b) < (3, c),$$

since $(2, b) \in G$ and $1 \equiv 2$ (α), but $(1, b) \notin G$.

Resuming the proof of the theorem, next we show that α is a transitive relation. For, if $a \equiv a'$ (α) and $a' \equiv a''$ (α), then by a repeated application of the lemma we infer that $a \equiv a''$ (α) also holds.

For α the substitution property holds, i. e. $a_1 \equiv a_1'$ (α) and $a_2 \equiv a_2'$ (α) imply $a_1 \circ a_2 \equiv a_1' \circ a_2'$ (α) for all operations $\circ$ in A. Hypotheses imply the existence of elements $b_1, b_2 \in B$ such that $(a_1, b_1), (a_1', b_1), (a_2, b_2)$ and (a_2', b_2) all belong to G. By (iii) we may hence conclude

$$(a_1 \circ a_2, b_1 \circ b_2) \in G \quad \text{and} \quad (a_1' \circ a_2', b_1 \circ b_2) \in G,$$

consequently,

$$a_1 \circ a_2 \equiv a_1' \circ a_2' \ (\alpha),$$

q. e. d.

14 Note that we make use of associativity.

15 $\bar{a}$ means the (unique) complement of a.

16 See BIRKHOFF [3].

Finally we verify the operator-admissibility of α, i. e. the fact that $a \equiv a'$ (α) implies $a^{\theta} \equiv a'^{\theta}$ (α) for all $\theta \in \Omega$ where Ω is the common operator domain of A, B and G. If $a \equiv a'$ (α), then there exists a $b \in B$ with $(a, b) \in G$ and $(a', b) \in G$. Hence $(a, b)^{\theta} = (a^{\theta}, b^{\theta})$ as well as $(a', b)^{\theta} = (a'^{\theta}, b^{\theta})$ belong to G, therefore $a^{\theta} \equiv a'^{\theta}$ (α), in fact.

In view of what has been proved for α it is seen that α defines a subdivision of A into disjoint classes, further, by the substitution property and the operator-admissibility, the set A_{α} of these classes may be made into a structure with the same operations as those originally defined in A and with the same operator domain Ω. Clearly, A may be mapped Ω-homomorphically onto A_{α}, this homomorphism being realized by $a \rightarrow [a]$, where $[a]$ denotes the class containing a.

The same result may be stated of the relation β in B.

Lemma implies that between the residue classes modulo α and β, in other words, between the elements of the quotient structures A_{α} and B_{β}, there is a one-to-one correspondence:

$$[a] \leftrightarrow [b] \textit{ if and only if } (a, b) \in G.$$

Moreover, the proof above shows that this correspondence is an isomorphism. If we denote by F a structure isomorphic to A_{α} and B_{β}, we arrive at the result enunciated in Theorem 2.

Theorem 2 may be stated in an other form too. It is only to be noted that a homomorphism of a group or a ring or a Boolean algebra is completely determined by the zero class, i. e. by the set of all elements congruent to 0. These classes A_0 resp. B_0 are normal substructures[17] of A and B, respectively, and F is obviously isomorphic to both quotient structures A/A_0 and B/B_0. Hence we infer:

THEOREM 3. *All subdirect unions of the structures A and B of type either* I *or* II *or* III *may be constructed as follows: take a pair of normal substructures A_0 and B_0 of A and B, respectively, such that $A/A_0 \simeq B/B_0$ with a fixed isomorphism; then form all pairs (a, b) where $a \in A$ and $b \in B$ belong to residue classes corresponding to each other under the prescribed isomorphism.*[18]

A_0 and B_0 will be called the *kernels* of the subdirect union $G = A \dotplus B$ in question. Plainly, $A_0 = A$ and $B_0 = B$ correspond to the case of the direct union $A \dotplus B$.

[17] By a *normal* substructure we mean one which defines a homomorphism of the whole structure (normal subgroup, twosided ideal etc.).

[18] Observe the interesting fact that, in case of groups, in a double module expansion (see SPEISER [14], p. 62) of the subdirect sum, modulis the subgroups A^* and B^* (see § 6), the cosets may be described in a simple manner. (A proof on these lines was suggested by G. HAJÓS.)

If A and B are both finite, containing n and m elements, respectively, and if F consists of r elements, then a subdirect union of A and B corresponding to a W-homomorphism $A \overset{F}{\sim} B$ contains $\frac{nm}{r}$ elements.

It is evident that if B is a simple structure, then a subdirect union of B with any other structure A (of the same type) is either the direct union $A \dotplus B$ or is an improper subdirect union existing if and only if A and B are trivially W-homomorphic.

§6. Some isomorphism statements

The structure theorem of subdirect unions (Theorem 2 or 3) enables us to obtain some isomorphism propositions correlated with subdirect unions.

Let $G = A + B$ be a subdirect union with the kernels A_0 and B_0. Of course, G contains normal substructures A^* and B^* isomorphic to A_0 and B_0, respectively. A^* consists of all $(a, 0)$ with $a \in A_0$ and B^* consists of all $(0, b)$ with $b \in B_0$. Clearly, A^* and B^* have no element in common except $(0, 0)$:

$$A^* \cap B^* = (0, 0).$$

Hence G contains the direct union $A^* \dotplus B^* = A_0 \dotplus B_0$.

THEOREM 4. *The following isomorphisms hold:*

$$G/(A_0 \dotplus B_0) \simeq A/A_0; \qquad G/(A_0 \dotplus B_0) \simeq B/B_0$$

where G is a subdirect union of A and B with the kernels A_0 and B_0.

It is clearly sufficient to verify only the first isomorphism. We map $(a, b) \in G$ upon the element $[a]$ of A/A_0, i. e. the residue class of A (modulo A_0) containing a. It is evident that this mapping is a homomorphism and its kernel consists of all elements of G which are mapped upon $[0]$. These elements are just those having a first component belonging to A_0, i. e. the kernel is $A_0 \dotplus B_0$. Hence the first indicated isomorphism is established.

Again, let A^* and B^* have the same meaning as above.

THEOREM 5. *The following isomorphisms hold:*

$$G/B^* \simeq A \quad \textit{and} \quad G/A^* \simeq B.$$

Let us map G onto A. The correspondence under which $a \in A$ is the image of $(a, b) \in G$ is clearly a homomorphism whose kernel consists of all elements of G which are mapped upon $0 \in A$. Thus the kernel is equal to B^*. Therefore, indeed, $G/B^* \simeq A$, and similarly $G/A^* \simeq B$.

Now it is easy to prove G. BIRKHOFF's theorem[19] on the subdirect representability (for two components).

[19] See BIRKHOFF [2] and [3], for groups BOURBAKI [4], p. 9', example 6.

THEOREM 6. *If an algebraic structure G has two normal substructures A^* and B^* such that $A^* \cap B^* = 0 \in G$, then G may be represented as a subdirect union of the quotient structures $A \simeq G/B^*$ and $B \simeq G/A^*$ with kernels A_0, B_0 isomorphic to A^* and B^*, respectively.*

Let $a \in A$ and $b \in B$ correspond to the element $x \in G$ under the homomorphisms $G \overset{\sim}{\to} A$ and $G \overset{\sim}{\to} B$, respectively. The mapping $x \to (a, b)$ is a homomorphism of G onto a structure H which is a subdirect union of A and B. In order to show that G is isomorphic to H, we prove that $y \to (0, 0)$ implies $y = 0 \in G$. Indeed, since $y \to 0$ of A under $G \overset{\sim}{\to} G/B^* \simeq A$ and $y \to 0$ of B under $G \overset{\sim}{\to} G/A^* \simeq B$, we obtain $y \in B^*$ and $y \in A^*$, whence $y \in A^* \cap B^*$ and $y = 0$. Thus $G \simeq A + B$. Here the kernels are isomorphic to A^* and B^*, respectively, for the kernels must be isomorphic to the structure of all $(a, 0) \in G$ and to that of all $(0, b) \in G$, respectively, and these sets may be characterized by the property of being mapped upon 0 of B and 0 of A under $G \overset{\sim}{\to} B$ and $G \overset{\sim}{\to} A$, respectively.

§ 7. On the substructures of the direct union

As an application of Theorem 3, let us consider a subdirect union $G = A + B$ with kernels A_0 and B_0. We have seen that G contains the direct union $A_0 \dotplus B_0$, and by Theorem 3, this is the greatest direct union contained in G with the property of having substructures of A and B for its components.

This observation admits of getting a survey over all substructures of the direct union of two structures of type I—III.

THEOREM 7. *If G is a substructure of the direct union $A \dotplus B$ of structures of type* I—III, *then there are two direct unions $\bar{A} \dotplus \bar{B}$ and $A_0 \dotplus B_0$ with*

$$A \supseteq \bar{A} \supseteq A_0 \quad \text{and} \quad B \supseteq \bar{B} \supseteq B_0,$$

determined uniquely, such that $\bar{A} \dotplus \bar{B}$ is the minimal direct union containing G and $A_0 \dotplus B_0$ is the maximal direct union contained in G (with components in A and B, respectively). Further, A_0 and B_0 are normal substructures of $\bar{A}$ and $\bar{B}$, respectively, and

$$\bar{A}/A_0 \simeq \bar{B}/B_0.$$

For the proof first take into account that $\bar{A}$ is the structure of all elements $a \in A$ figuring at least in one element (a, b) of G. The structure $\bar{B}$ is analogously defined, while A_0 and B_0 have the same meaning as before. Considering that G is a subdirect union of the structures $\bar{A}$ and $\bar{B}$ with the kernels A_0 and B_0, our statement follows at once from Theorem 3.

In case A and B are Abelian groups (possibly with operator domains), we have the following theorem.

THEOREM 8. *If A and B are Abelian groups, then with the notations of the preceding theorem:*

$$(\bar{A} \dotplus \bar{B})/G \simeq G/(A_0 \dotplus B_0).$$

By Theorem 4, it is sufficient to prove the isomorphism

$$(\bar{A} \dotplus \bar{B})/G \simeq \bar{A}/A_0 \simeq \bar{B}/B_0.$$

The so-called second isomorphism theorem [20] implies

$$(\bar{A} \dotplus \bar{B})/G \simeq [(\bar{A} \dotplus \bar{B})/(A_0 \dotplus B_0)]/[G/(A_0 \dotplus B_0)]$$

and hence we obtain

$$(\bar{A} \dotplus \bar{B})/G \simeq (\bar{A}/A_0 \dotplus \bar{B}/B_0)/(\bar{A}/A_0 + \bar{B}/B_0)$$

where $\bar{A}/A_0 + \bar{B}/B_0$ is a subdirect sum with zero kernels (the components are isomorphic). What we have to verify is that the factor group on the right hand side is isomorphic to $\bar{A}/A_0 \simeq \bar{B}/B_0$.

For the sake of simplicity, we write C and D for $\bar{A}/A_0$ and $\bar{B}/B_0$, respectively; $C \simeq D$. Then the isomorphism to be proved reads as follows:

$$(C \dotplus D)/(C + D) \simeq D \qquad (\text{or} \simeq C)$$

where $C+D$ is a subdirect sum with zero kernels.

The direct sum $C \dotplus D$ consists of all pairs $(c, d)\,(c \in C, d \in D)$, while the subdirect sum in question consists of all pairs (c', d') with $c' \in C$ taken arbitrarily and then taking $d' \in D$ corresponding to this c' under a certain fixed isomorphism $\varphi(c') = d'$ of C onto D determined by $C+D$. Now the mapping $(c, d) \to d - \varphi(c)$ of $C \dotplus D$ onto D is a homomorphism:

$$(c_1, d_1) + (c_2, d_2) = (c_1 + c_2, d_1 + d_2) \to$$
$$\to (d_1 + d_2) - \varphi(c_1 + c_2) = (d_1 - \varphi(c_1)) + (d_2 - \varphi(c_2)).$$

The kernel of this homomorphism is the set of all (c', d') of $C \dotplus D$ which are mapped upon $d' - \varphi(c') = 0$ of D, i. e. the set $(c', \varphi(c'))$ which is equal to the subdirect sum $C+D$ under consideration. This completes the proof.

Theorems 7 and 8 give an interesting information about the subgroups G of the direct sum of two Abelian groups A and B. Theorem 7 tells us that G is an intermediate group between $A_0 \dotplus B_0$ and $\bar{A} \dotplus \bar{B}$ such that — roughly speaking — the "distance" between A_0 and $\bar{A}$, and that between B_0 and $\bar{B}$ are the same, while Theorem 8 states that G occupies, in a certain sense, just the "medial" position between these direct sums.

§ 8. Examples

I. **Cyclic p-groups.** Here we discuss the subdirect sums of two cyclic p-groups where p is a prime number. (We denote the group operation by addition.)

[20] See e. g. VAN DER WAERDEN [15], p. 149, or ZASSENHAUS [17], p. 35.

Let A and B be two finite cyclic groups of prime power order p^m and q^n, respectively. If $p \neq q$, A and B are W-homomorphic only with respect to the zero group 0, therefore, in this case they have only one subdirect sum, namely, the direct sum $A \dotplus B$. If $p = q$, that is,[21] $A = \mathfrak{Z}(p^m)$ and $B = \mathfrak{Z}(p^n)$, suppose for the sake of definiteness e. g. $m \geqq n$. Let $A+B$ be a subdirect sum with the kernels $A_0 = \mathfrak{Z}(p^{m-k})$ and $B_0 = \mathfrak{Z}(p^{n-k})$ where $0 \leq k \leq n$. Denote by a' a generating element of A and by b' an element of B such that $(a', b') \in A+B$. It is clear that (a', b') generates a cyclic subgroup G' of $A+B$, isomorphic to A. Now consider the set of all elements of $A+B$ which are of the form $(0, b)$ with $b \in B$; these elements form a subgroup G'' of $A+B$ isomorphic to B_0. The groups G' and G'' have no element in common except, of course, $(0, 0)$, for if (a, b) belongs to $G' \cap G''$, then it has the form (ra', rb') with some natural number r, and also the form $(0, b)$. Hence we get $ra' = 0$, r is divisible by p^m, and so a fortiori by p^n, i. e. $rb' = 0$ is obtained, actually. Since each element (a, b) of $A+B$ has a decomposition

$$(a, b) = (ra', b) = (ra', rb') + (0, b - rb')$$

with $(ra', rb') = r(a', b') \in G'$ and $(0, b - rb') \in G''$, where r denotes a natural number, we see that

$$A+B = G' \dotplus G'' \simeq A \dotplus B_0.$$

We have thus proved

THEOREM 9. *Any subdirect sum of the groups* $\mathfrak{Z}(p^m)$ *and* $\mathfrak{Z}(p^n)$ *with the kernels* $\mathfrak{Z}(p^{m-k})$ *and* $\mathfrak{Z}(p^{n-k})$, *is isomorphic to*

$$\mathfrak{Z}(p^m) \dotplus \mathfrak{Z}(p^{n-k}) \quad \text{or} \quad \mathfrak{Z}(p^n) \dotplus \mathfrak{Z}(p^{m-k})$$

according as $m \geqq n$ *or* $m < n$.

It should be noted that our reasoning also shows that if A is an infinite cyclic group and B is any Abelian group (finite or infinite), then a subdirect sum $A+B$ with kernels A_0 and B_0 is isomorphic to the direct sum $A \dotplus B_0$. Hence a subdirect sum of two Abelian groups, one of which is an infinite cyclic group, is either itself an infinite cyclic group or is directly decomposable,[22] i. e. representable as the direct sum of two of its proper subgroups.

II. **Groups of type** $\mathfrak{Z}(p^\infty)$. Let us consider two groups of Prüferian type[23] $\mathfrak{Z}(p^\infty)$, $A = \mathfrak{Z}(p^\infty)$ and $B = \mathfrak{Z}(q^\infty)$ where p and q are primes. A group of this type possesses the well-known property that each proper factor group of it is isomorphic to itself. Therefore, if $p \neq q$, A and B have no subdirect sum other than the direct one.[24]

[21] By $\mathfrak{Z}(r)$ we denote the cyclic group of order r.

[22] According as B_0 is the zero group or not.

[23] See PRÜFER [12].

[24] It is readily seen that a group of type $\mathfrak{Z}(p^\infty)$ has no non-direct subdirect sum with a finite group.

Consider now the case $p = q$. Then to each pair of finite subgroups $A_0 = \mathfrak{Z}(p^m)$ and $B_0 = \mathfrak{Z}(p^n)$ of A and B, respectively, there exists, by Theorem 1, a subdirect sum $G = A + B$ with the kernels A_0 and B_0. We are going to prove:

THEOREM 10. *A subdirect sum* $G = A + B$ *of the groups* $A = \mathfrak{Z}(p^\infty)$ *and* $B = \mathfrak{Z}(p^\infty)$ *with the kernels* $A_0 = \mathfrak{Z}(p^m)$ *and* $B_0 = \mathfrak{Z}(p^n)$ *is isomorphic to the direct sum* $A \dot{+} B_0$ *provided* $m \geqq n$.

Assume $m \geqq n$ $(n < \infty)$ and consider a generator system $a_0, a_1, a_2, \ldots$ of the group A; the a_i's are connected by the usual defining relations

$$a_0 \neq 0, \quad pa_0 = 0, \quad pa_1 = a_0, \quad pa_2 = a_1, \ldots$$

G must contain elements of the form (a_r, b) for $r = 0, 1, 2, \ldots$; for a fixed r, let the (finite) set of all (a_r, b) be denoted by A_r. Then we have obviously

$$p^k A_r \subseteq A_{r-k} \qquad (k = 0, 1, 2, \ldots, r).$$

Since A_0 consists of but a finite number of elements and contains all the sets

$$(*) \qquad A_0,\ pA_1,\ p^2A_2,\ \ldots,\ p^sA_s,\ \ldots,$$

we conclude that there exists an element $c_0 = (a_0, b_0) \neq 0$ in G which occurs in an infinity of the sets $(*)$.

Now let A'_r denote the set of those (a_r, b) of A_r for which $p^r(a_r, b) = (a_0, b_0)$ and let $c_1 = (a_1, b_1)$ be an element common to an infinity of the sets

$$A'_1,\ pA'_2,\ p^2A'_3,\ \ldots,\ p^{s-1}A'_s,\ \ldots$$

The choice of c_0 guarantees the existence of such a c_1. Thus proceeding, we obtain elements $c_0, c_1, c_2, \ldots$ of G satisfying the following set of relations:

$$c_0 \neq 0, \quad pc_0 = 0, \quad pc_1 = c_0, \quad pc_2 = c_1, \ldots$$

Consequently, G contains a subgroup C of type $\mathfrak{Z}(p^\infty)$, generated by the elements $c_0, c_1, c_2, \ldots$

Next we observe that if[25] $O(a) = p^{m+s}$ $(s = 1, 2, \ldots)$, then $(a, b) \in G$ implies $O(b) = p^{n+s}$. In fact, this is an immediate consequence of Theorem 3, for in the factor groups $A/A_0 \cong \mathfrak{Z}(p^\infty)$ and $B/B_0 \cong \mathfrak{Z}(p^\infty)$ only cosets of the same order p^s may correspond to each other under an isomorphism. Hence, in the mentioned case, we have the divisibility relation $O(b) | O(a)$. This conclusion holds for all the c_ν's constructed above (not only for $\nu > m$) in view of the relations $c_\nu = p^{m-\nu+1} c_{m+1}$ $(\nu = 0, 1, 2, \ldots, m)$.

We are now in a position to show that C has no element $\neq (0, 0)$ in common with the subgroup D of G which consists of all $(0, b)$ with $b \in B_0$ and is therefore isomorphic to B_0. To prove this, assume $g = (a, b) \in C \cap D$. Then g has the form $g = kc_r = (ka_r, kb_r)$ and also the form $g = (0, b')$ for some r, k, and $b' \in B_0$. Hence $ka_r = 0$, $O(a_r) | k$ and therefore, by the previous remark, $O(b_r) | k$, whence $kb_r = 0$ follows. This establishes $g = (0, 0)$, in fact.

[25] The order of an element x will be denoted by $O(x)$.

Each element (a, b) of G may be decomposed in the form

$$(a, b) = (ka_r, b) = (ka_r, kb_r) + (0, b - kb_r)$$

where

$$(ka_r, kb_r) = kc_r \in C \quad \text{and} \quad (0, b - kb_r) \in D.$$

Therefore, we actually arrive at a direct sum representation

$$G = A + B = C \dotplus D \cong A \dotplus B_0.$$

III. **Example for a directly indecomposable subdirect sum.** The examples of (non-trivial) subdirect sums considered so far were all directly decomposable. This is always the case if the components are finite Abelian groups, as it follows at once from the fundamental theorem on finitely generated Abelian groups. But in case the groups are non-commutative or infinite, the subdirect sums are, in general, directly indecomposable.[26] For non-commutative groups this may be illustrated by the following simple example.

Let both A and B be isomorphic to the symmetric group $\mathfrak{S}_3$ of order 6, and let A_0 and B_0 denote those subgroups of A and B, respectively, which are of order 3; these are normal subgroups and are already commutative. The subdirect sum $G = A + B$ with the kernels A_0 and B_0 is directly indecomposable. For, a direct decomposition would have components of order 2 and 9, or 3 and 6; hence one of the direct components would lie in the center, but — as may be readily checked — the center of G consists of the group identity alone.

§ 9. On the isomorphism of subdirect unions with the same kernels

First of all we intend to show what we have already mentioned in § 4: there may exist non-isomorphic subdirect sums of two groups with the same kernels.

We need the following simple fact.

Let $A = A' \dotplus A''$ and $B = B' \dotplus B''$ have two normal subgroups $A_0 = A_0' \dotplus A_0''$ $(A_0' \subseteq A', A_0'' \subseteq A'')$ and $B_0 = B_0' \dotplus B_0''$ $(B_0' \subseteq B', B_0'' \subseteq B'')$, respectively, such that $A'/A_0' \cong B'/B_0'$ and $A''/A_0'' \cong B''/B_0''$. Then a subdirect sum $A + B$ (with the kernels A_0 and B_0) which is obtained by letting the cosets of A'/A_0' and A''/A_0'' correspond to the cosets of B'/B_0' and B''/B_0'', respectively, is the direct sum of the subdirect sums $A' + B'$ (with the kernels A_0' and B_0') and $A'' + B''$ (with the kernels A_0'' and B_0''). Indeed, a subdirect sum $A + B$ in question consists of pairs of the form

$$(a, b) = (a' + a'', b' + b'') = (a', b') + (a'', b'') \qquad (a \in A,\ b \in B)$$

[26] For a directly indecomposable infinite Abelian group (of rank 2) see Pontrjagin [11].

where the (uniquely determined) elements $a' \in A'$ and $b' \in B'$, $a'' \in A''$ and $b'' \in B''$ belong to corresponding cosets of A_0' and B_0', resp. A_0'' and B_0''.

In order to illustrate the mentioned possibility, let

$$A = A' \dotplus A'' \simeq \mathfrak{Z}(p^5) \dotplus \mathfrak{Z}(p^3)$$

and

$$B = B' \dotplus B'' \simeq \mathfrak{Z}(p^4) \dotplus \mathfrak{Z}(p^2).$$

We form two non-isomorphic subdirect sums of A and B with the same kernels

$$A_0 = A_0' \dotplus A_0'' \simeq \mathfrak{Z}(p^4) \dotplus \mathfrak{Z}(p^2) \qquad (A_0' \subseteq A',\ A_0'' \subseteq A'')$$

and

$$B_0 = B_0' \dotplus B_0'' \simeq \mathfrak{Z}(p^3) \dotplus \mathfrak{Z}(p) \qquad (B_0' \subseteq B',\ B_0'' \subseteq B'').$$

Since the factor groups A'/A_0', A''/A_0'', B'/B_0' and B''/B_0'' are all isomorphic to $\mathfrak{Z}(p)$, therefore, by virtue of the previous remark, the following subdirect sums exist:

(1) $G = A + B = (A' + B') \dotplus (A'' + B'')$ where the kernels of the subdirect sums on the right hand side are A_0' and B_0', resp. A_0'' and B_0''. The discussions in § 8 imply

$$G \simeq (A' \dotplus B_0') \dotplus (A'' \dotplus B_0'') \simeq \mathfrak{Z}(p^5) \dotplus \mathfrak{Z}(p^3) \dotplus \mathfrak{Z}(p^3) \dotplus \mathfrak{Z}(p).$$

(2) $H = A + B = (A' + B'') \dotplus (A'' + B')$ where the kernels of the subdirect sums on the right hand side are A_0' and B_0'', resp. A_0'' and B_0'. Hence we obtain

$$H \simeq (A' \dotplus B_0'') \dotplus (B' \dotplus A_0'') \simeq \mathfrak{Z}(p^5) \dotplus \mathfrak{Z}(p) \dotplus \mathfrak{Z}(p^4) \dotplus \mathfrak{Z}(p^2).$$

Nevertheless, G and H are not isomorphic.

A similar instance may be given by the aid of groups of type $\mathfrak{Z}(p^\infty)$ by making use of Theorem 10.

It seems natural to raise the problem of finding necessary and sufficient conditions for deciding whether or not two given subdirect unions of two structures A and B of type I—III with the same kernels A_0 and B_0, are isomorphic. I did not succeed in solving completely this problem. Now I am going to give an almost trivial sufficient condition for the isomorphism of two subdirect unions of the considered type.

In the formulation of this condition we shall make use of the following concept.

Let A be an arbitrary algebraic structure, A_0 one of its normal substructures and F the quotient structure A/A_0. It is evident that each automorphism Γ of A which leaves A_0 invariant: $A_0^\Gamma = A_0$, induces an automorphism γ of F such that, under γ, the image of the residue class containing a is the residue class containing a^Γ. If to an arbitrarily given automorphism γ of F there exists an automorphism Γ of A such that γ may be obtained in the way just described, we say that the automorphism γ of F is *extensible* *to A*. Plainly, in general, not all the automorphisms of F are extensible to A, but in case of groups the inner automorphisms are always extensible, as it can readily be proved.

For convenience, we shall identify the quotient structures A/A_0 and B/B_0 under two arbitrary, but fixed isomorphisms with an abstract structure F. Then each subdirect union G of A and B with the kernels A_0 and B_0 induces an automorphism ξ of F, and conversely. This automorphism ξ may be described as follows: if $f \in F$, choose an element a of A which belongs to the residue class modulo A_0 corresponding to f, then take a $b \in B$ such that $(a, b) \in G$ and find $f' \in F$ corresponding to this b; in this case $f^{\xi} = f'$. The construction of $A \dotplus B$ for a given ξ is obvious. We conclude that there are as many *formally* different subdirect unions of A and B with the kernels A_0 and B_0, as many automorphisms of F exist. Now we prove:

THEOREM 11. *Two automorphisms ξ and η of F define isomorphic subdirect unions, if there are automorphisms γ and δ of F so that γ is extensible to A and δ is extensible to B, further the equation $\gamma\xi = \eta\delta$ holds.*

Let Γ denote an automorphism of A which is an extension of γ, and Δ an extension of δ to B. Note that the correspondence $(a, b) \longleftrightarrow (a^{\Gamma^{-1}}, b)$ is an isomorphism between the subdirect unions belonging to ξ and $\gamma\xi$, respectively, and $(a', b') \longleftrightarrow (a', b'^{\Delta})$ is the same for η and $\eta\delta$. Considering that $\gamma\xi = \eta\delta$, we have our assertion.

(Received 15 April 1952)

Bibliography

1. R. BAER, Abelian groups without elements of finite order, *Duke Math. Journal*, **3** (937), pp. 68—122.
2. G. BIRKHOFF, Subdirect unions in universal algebra, *Bull. Amer. Math. Soc.*, **50** (1944), pp. 764—768.
3. G. BIRKHOFF, *Lattice theory* (New York, 1948), 2nd ed.
4. N. BOURBAKI, *Algèbre*, I: Structures algèbriques (Paris, 1951), 2nd ed.
5. B. BROWN and N. H. MCCOY, Radicals and subdirect sums, *Amer. Journal Math.*, **69** (1947), pp. 44—58.
6. M.-L. DUBREIL—JACOTIN, Quelques propriétés des applications multiformes, *Comptes Rendus Acad. Sci. Paris*, **230** (1950), pp. 806—808.
7. P. DUBREIL, Relations binaires et applications, *Comptes Rendus Acad. Sci. Paris*, **230** (1950), pp. 1028—1030.
8. N. JACOBSON, The radical and semi-simplicity for arbitrary rings, *Amer. Journal Math.*, **67** (1945), pp. 300—320.
9. N. H. MCCOY, Subdirect sums of rings, *Bull. Amer. Math. Soc.*, **53** (1947), pp. 856—877.
10. H. NEUMANN, Generalized free products with amalgamated subgroups, *Amer. Journal Math.*, **70** (1948), pp. 590—625.

11. L. Pontrjagin, The theory of topological commutative groups, *Annals of Math.*, **35** (1934), pp. 361—388.
12. H. Prüfer, Untersuchungen über die Zerlegbarkeit der abzählbaren primären Abelschen Gruppen, *Math. Zeitschrift*, **17** (1923), pp. 35—61.
13. O. Schreier, Die Untergruppen der freien Gruppen, *Abhandlungen math. Sem. Hamburg. Univ.*, **5** (1927), pp. 161—183.
14. A. Speiser, *Theorie der Gruppen von endlicher Ordnung* (Berlin, 1927), 2nd ed.
15. B. L. van der Waerden, *Moderne Algebra*, I (Berlin, 1937), 2nd ed
16. J. H. M. Wedderburn, Homomorphisms of groups, *Annals of Math.*, **42** (1941), pp. 486—487.
17. H. Zassenhaus, *The theory of groups* (New York, 1949).

О ПОДПРЯМЫХ СУММАХ, I

Л. ФУКС (Будапешт)

(Р е з ю м е)

В ряде структурных теорем абстрактной алгебры встречаются подпрямые суммы, которые однако не дают полного обозрения данной алгебраической системы, как прямые суммы. Причина этого заключается в том, что пока ещё не известен такой метод, при помощи которого было бы возможно „внешнее" построение всех подпрямых сумм. (Под „внешним" построением подразумевается то, что о производимой группе мы используем лишь такие факты, которые могут быть сформулированы при помощи данных компонент.) Настоящая работа даёт общий метод построения подпрямой суммы двух групп, или колец, или алгебр Буля. Этот метод заключается в следующем.

Пусть группы (кольца, алгебры Буля) A и B при помощи гомоморфии переходят в группу (кольцо, алгебру Буля) F. Те пары элементов (a, b) $(a \in A, b \in B)$, которые при помощи вышеуказанных гомоморфий переходят в тот же элемент $f \in F$, дают подпрямую сумму A и B, если действия производить покомпонентно. Автор доказывает, что этот метод является общей конструкцией для указанных систем.

В дальнейшей части работы автор доказывает несколько утверждений об изоморфии, показывает несколько примеров и делает несколько замечаний об изоморфизме подпрямых сумм, представленных из тождественных компонент.

ON THE DECOMPOSIBILITY OF ABELIAN p-GROUPS INTO THE DIRECT SUM OF CYCLIC GROUPS

By

A. KERTÉSZ (Debrecen)

(*Presented by* L. RÉDEI)

§ 1. Introduction

According to an important theorem of PRÜFER a countable abelian p-group is the direct sum of cyclic groups if and only if it does not contain elements of infinite height [3].[1, 2] It is well known that this theorem cannot be extended to non-countable groups. KULIKOV succeeded in giving a sufficient and necessary condition for an abelian p-group of arbitrary power to be the direct sum of cyclic groups [2]. In what follows I shall give a criterion of a different kind for the same fact; this criterion can easily be proved and the above mentioned results of PRÜFER and KULIKOV follow without difficulty from it. Other immediate corollaries of our theorem are the fundamental theorem of finite abelian groups, and a result of BAER according to which an abelian p-group containing an element of maximal order is always the direct sum of cyclic groups [3]. As an abelian torsion group is the direct sum of uniquely determined p-groups, our result can be extended in an obvious way to torsion groups of arbitrary power. I remark that in the following proofs there is made no use of ordinal numbers and of the method of transfinite induction; merely we need ZORN's lemma, but even this is involved only in the proof of BAER's and KULIKOV's theorem.

The notations and symbols used are the following. The letters $x, a, b, \ldots, g$ denote elements of groups and the other small Latin letters ordinary integers, in particular p a prime number. Groups are written additively. We denote by $O(a)$ the order of the element a of a group. In what follows only abelian p-groups are considered, i. e. abelian groups in which the order of every element is a power of a fixed prime p. In investigating such groups, an important rôle is played by the concept of the *height* of an element. An element $a \neq 0$ of the p-group G is said to have the height $h = H(a)$, if the

[1] The numbers in brackets refer to the Bibliography at the end of this paper.

[2] For notation and terminology see the next paragraph.

equation $p^n x = a$ is solvable in G for $n \leq h$, but not for $n > h$. We define $H(a) = \infty$ if $p^n x = a$ has a solution $x \in G$ for each natural number n. We emphasize that $H(a)$ is defined only for $a \neq 0$.[3] If G is the direct sum of its subgroups $B_1, B_2, \ldots$, and $g = b_1 + b_2 + \ldots (b_\nu \in B_\nu)$, then evidently $H(g) \leq \leq H(b_\nu)$ $(\nu = 1, 2, \ldots)$ holds. The elements $a_1, \ldots, a_n$ of the group G are called independent if $r_1 a_1 + \ldots + r_n a_n = 0$ implies $r_1 a_1 = \ldots = r_n a_n = 0$. An arbitrary set S of elements of G is independent, if every finite subset of S is independent. The independence so defined is therefore a property of finite character, consequently, by virtue of ZORN's lemma (or the equivalent lemma of TUKEY), every subset R of G contains a *maximal independent system* S. If $R = G$, we call S a maximal independent system of G.

§ 2. The criterion

In formulating the theorem below, we shall have to make use of the following

DEFINITION. A maximal independent system P of the abelian p-group G will be called a *principal system of* G, if no element of P can be exchanged for an element of a greater height of G without violating the independence of the system.[3]

REMARK. Each element of finite height of a principal system is obviously of order p, for an element $a \in P$ of order p^k $(k > 1)$ would be exchangeable for the element $p^{k-1}a$ of a greater height.

Our principal aim is to prove the following

THEOREM. *An abelian p-group G containing no element of infinite height is the direct sum of cyclic groups if and only if there exists a principal system of elements in G.*

PROOF. The necessity of the condition of the Theorem is almost evident. In fact, let us suppose that G is the direct sum of the cyclic groups $\{c_\nu\}$ with $O(c_\nu) = p^{m_\nu}$. In this case the set of all elements $p^{m_\nu - 1} c_\nu = a_\nu$ is a principal system P of G. For P is a maximal independent system in G; furthermore for an arbitrary element $a' \in G$ of order p^k a relation of the form

$$p^{k-1} a' = r_1 a_1 + \ldots + r_n a_n$$

holds (with a suitable notation of the elements of P), showing that $a_1, \ldots, a_n$ are the only elements of P one of which may be replaced by a' without violating the independence of the system P. Since $H(a') \leq H(p^{k-1} a') \leq H(a_i)$ $(i = 1, \ldots, n)$, no element of P can be exchanged for an element of a greater height of G.

[3] By the *height* of an element g of G we always mean the number $H(g)$ defined above, i. e. the height refers always to the whole group G, even when an element is, for the moment, considered as an element of some subgroup of G.

In order to prove the sufficiency of the condition in the Theorem, we consider an arbitrary abelian p-group without elements of infinite height and containing a principal system S. We determine for any $a_\nu \in S$ an element $c_\nu \in G$ such that

(1) $$p^{h_\nu} c_\nu = a_\nu \qquad \text{with} \quad h_\nu = H(a_\nu).$$

We state that G is the direct sum of the cyclic groups $\{c_\nu\}$. Indeed, the independence of the system of the c_ν's follows immediately from the fact that a relation among the c_ν's would imply, by multiplication by a suitable power of p, a relation among some of the a_ν's. Therefore only the fact remains to be proved that the direct sum G' of the cyclic groups $\{c_\nu\}$ cannot be a proper subgroup of G.

Before proving this, we make the following remark: *A relation*

(2) $$pg = r_1 c_1 + \ldots + r_m c_m \qquad (g \in G)$$

implies $p \mid r_i \quad (i = 1, \ldots, m)$. For let us suppose the contrary. Then there exists a relation (2) for which $p \nmid r_i \quad (i = 1, \ldots, m)$. Among the elements $a_1, \ldots, a_m$ corresponding to the elements $c_1, \ldots, c_m$ in (2), let a_1 be one of maximal height:

$$h_1 = H(a_1) \geqq H(a_i) \qquad (i = 1, \ldots, m).$$

Thus, by virtue of (1), (2) implies

$$a' = p^{h_1+1} g = r_1 a_1 + \ldots .$$

This means, however that the element a_1 of S can be exchanged for an element a' with $H(a') > H(a_1)$, without violating the independence of the system. This contradicts the principal property of the system S.

Now suppose that G' is a proper subgroup of G. Then there exists an element $g \in G$ such that

(3) $$g \notin G', \quad pg \in G',$$

i. e. a relation (2) holds. Thus $r_i = p r_i' \quad (i = 1, \ldots, m)$ and so

$$p(g - r_1' c_1 - \ldots - r_m' c_m) = 0.$$

Therefore

(4) $$g' = g - r_1' c_1 - \ldots - r_m' c_m$$

— as an element $\neq 0$, see (3) — is an element of order p, consequently, by the maximal independence of the system S, g' is a linear combination of some a_ν's, i. e., by (1), of some c_ν's. Then (4) means that $g \in G'$ which contradicts (3), completing the proof of the Theorem.

§ 3. Applications

From the Theorem of § 2 it follows at once the fundamental theorem of finite abelian groups:

COROLLARY 1. *Every finite abelian group is the direct sum of cyclic groups.*

Indeed, any finite abelian p-group obviously contains a principal system of elements.

As another application we have the following theorem of BAER [1]:

COROLLARY 2. *Every abelian p-group having an element of maximal order is the direct sum of cyclic groups.*

As a matter of fact, if p^u is the maximum of the orders of elements of an abelian p-group G, then choose a maximal independent system of elements of height $u-1$, extend this system to an independent system by adjoining an (eventually empty) maximal system of elements of height $u-2$, and so on. The set of elements thus obtained is obviously a principal system in G.

Now we prove the following theorem of KULIKOV [2]:

COROLLARY 3. *An abelian p-group G is the direct sum of cyclic groups if and only if G is the union of a countable ascending chain*

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots \tag{5}$$

of its subgroups such that each A_n contains an element of maximal height $j_n < \infty$.[3]

PROOF. The necessity of the condition is evident. In order to prove its sufficiency, we construct a principal system in G as follows. Choose a maximal independent system of elements $\in A_1$ of height j_1 ("j_1-layer"), adjoin a maximal system of elements $\in A_1$ of height j_1-1 ("(j_1-1)-layer"), and so on, such that the resulting system S_1 shall be *a maximal independent system* of A_1.[3] In an analogous way we extend S_1 to a maximal independent system S_2 of A_2 by adjoining a maximal layer of height j_2, then a maximal layer of height j_2-1, and so on. We proceed likewise in constructing successively the systems $S_3, S_4, \ldots$, each S_n being *a maximal independent system* of A_n. Now we assert that the union S of all systems $S_1, S_2, \ldots$ is a principal system of G. Clearly, S is a maximal independent system of G, and *the elements of S are all of order p.* Hence for an arbitrary element g of order p^k of G a representation

$$p^{k-1}g = r_1a_1 + \ldots + r_na_n \tag{6}$$

holds, $a_1, \ldots, a_n$ being suitable elements of S, ordered so that

$$a_i \in A_m \quad \text{implies} \quad a_{i-s} \in A_m \qquad (s = 1, \ldots, i-1). \tag{7}$$

Equation (6) means that $a_1, \ldots, a_n$ are the only elements of S one of which may be replaced by g without violating the independence of the system S. Therefore it is sufficient to show that

$$h = H(p^{k-1}g) \leqq H(a_i) \qquad (i = 1, \ldots, n), \tag{8}$$

for, by $H(g) \leqq H(p^{k-1}g)$, (8) implies $H(g) \leqq H(a_i)$ $(i = 1, \ldots, n)$, which proves that S is in fact a principal system of G.

Suppose (8) is not true and let

$$H(a_i) < h \leqq H(a_{i+t}) \qquad (t = 1, \ldots, n-i) \tag{9}$$

Then, by (6) and (7), the element

$$g' = p^{k-1}g - (r_{i+1}a_{i+1} + \ldots + r_n a_n) = r_1 a_1 + \ldots + r_i a_i \in A_m$$

with $H(g') \geqq h > H(a_i)$ (see (9)) may replace the element $a_i \in S_m$ without violating the independence of S_m, which contradicts the maximality of the $H(g')$-layer of S_m.

An immediate consequence of Corollary 3 is the following theorem of PRÜFER [3]:

COROLLARY 4. *A countable abelian p-group is the direct sum of cyclic groups if and only if it contains no element of infinite height.*

For a countable abelian p-group is the union of a countable ascending chain of finite groups each of which contains an element of maximal height.[4]

I am indebted to Professor T. SZELE for his valuable remarks.

REMARK. L. FUCHS has kindly called my attention to the interesting fact that the theorem above may be formulated in the following manner too: *A subset $C = (c_1, c_2, \ldots)$ is a basis of the abelian p-group G if and only if C is a maximal independent system of G and no element of C can be replaced by an element of a greater order of G without violating the independence of C.* [$C = (c_1, c_2 \ldots)$ is called a *basis* of G if G is the direct sum of the cyclic groups $\{c_1\}, \{c_2\}, \ldots$.]

(Received 17 May 1952)

Bibliography

[1] R. BAER, Der Kern, eine charakteristische Untergruppe, *Compositio Math.*, **1** (1935), pp. 254—283.

[2] Л. Я. Куликов, К теории абелевых групп произвольной мощности, Мат. Сборник, Н. С., **16** (1945), (58), pp. 129—162.

[3] H. PRÜFER, Untersuchungen über die Zerlegbarkeit der abzählbaren primären Abelschen Gruppen, *Math. Z.*, **17** (1923), pp. 35—61.

[4] KULIKOV obtains in the same way Corollary 4 from Corollary 3 in [2].

О ПРЯМОМ РАЗЛОЖЕНИИ АБЕЛЕВЫХ p-ГРУПП В ПРЯМУЮ СУММУ ЦИКЛИЧЕСКИХ ГРУПП

А. КЕРТЕС (Дебрецен)

(Резюме)

В настоящей работе доказывается следующая теорема: Абелева p-группа (произвольной мощности), без элементов бесконечной высоты, разлагается в прямую сумму циклических групп тогда и только тогда, если в ней содержится максимальная независимая система элементов, ни один из элементов которой не может быть заменен элементом большей высоты без ущерба для независимости системы.

Из этой теоремы непосредственно вытекают три теоремы открытые Р. Бэром [1], Л. Я. Куликовым [2] и Х. Прюфером [3] соответственно.

ON GROUPS WITH ATOMIC LAYERS

By

T. SZELE (Debrecen)

(*Presented by* G. HAJÓS)

In what follows we shall determine all groups containing no different isomorphic subgroups. It will turn out that these coincide with the locally cyclic torsion groups, the latter being the subgroups of the additive group of all rational numbers mod 1 (or, in a multiplicative realization, the subgroups of the multiplicative group of all complex roots of unity). Here we shall use the additive notation, therefore, we consider the group in question in the form

$$C = C(2^\infty) + C(3^\infty) + \cdots + C(p^\infty) + \cdots \tag{1}$$

where $C(p^\infty)$ denotes the group of type (p^∞) (or quasicyclic group); p denotes throughout a prime number. It is obvious that, for each non-negative integer n, the group $C(p^\infty)$ contains exactly one subgroup of order p^n; these are cyclic groups and exhaust all proper subgroups of $C(p^\infty)$. We denote by $C(p^n)$ the cyclic group of order p^n.

The groups containing no different isomorphic subgroups may be characterized in an other form too. The set of all elements of order m in G is usually called the *layer* $L(m)$ of G; here m is a natural number or infinity. We shall call the layer $L(m)$ of G *atomic*, if $L(m)$ has no proper subset which is a layer of some subgroup of G. Clearly, $L(\infty)$ is an atomic layer only if it is the void set, furthermore, for a natural number m, $L(m)$ is an atomic layer if and only if it is either the void set or it consists of exactly $\varphi(m)$ elements where $\varphi(m)$ is EULER's function.

Now we are going to prove the following

THEOREM. *For an arbitrary group G the following statements are equivalent:*

a) *G contains no pair of distinct isomorphic subgroups;*

b) *every layer of G is atomic;*

c) *G is isomorphic to some subgroup of* (1).

PROOF. a) *implies* b). Assume that G contains no two different isomorphic subgroups. Then G contains no elements of infinite order, for if a were such an element, then the cyclic subgroups of G, generated by the elements a and $2a$, respectively, would be distinct isomorphic subgroups of G. Let $a \in G$ be an element of finite order m. Then all the elements of G which are of order m must occur in the sequence $a, 2a, 3a, \ldots$, for, in the contrary, G would have two different isomorphic subgroups. Consequently, we have shown that the elements of G of order m are in number $\varphi(m)$, i. e., each layer of G is atomic.

b) *implies* c). Let G be an arbitrary group every layer of which is atomic. Then G contains no element of infinite order. Next we show that G is a locally cyclic group. Indeed, let a and b be two arbitrary elements of G. Since the order of $-b+a+b$ is equal to that of a, a relation

$$-b+a+b=ka$$

holds for a suitable natural number k. Therefore the sum of two elements of the form $ra+sb$ may be written in the same form, consequently, the set of all elements $\in G$ of the form $ra+sb$ constitutes a *finite* subgroup H of G. Let us denote by n the order of H and by $\psi(d)$ the number of the elements of order d in H. Then

$$\sum_{d \mid n} \psi(d) = n = \sum_{d \mid n} \varphi(d)$$

holds, and from $0 \leq \psi(d) \leq \varphi(d)$ we obtain $\psi(d)=\varphi(d)$. In particular, $\psi(n)=\varphi(n)>0$, consequently, H is a cyclic group. (One can recognize in this argument the method by which the existence of a primitive root mod p is usually established.) Thus we have shown that G is locally cyclic, hence an abelian group.

Now G, as an abelian group without elements of infinite order, decomposes into the direct sum of p-groups. In view of (1), it remains to be proved that the p-component G_p of G is a group of type $C(p^n)$ where $0 \leq n \leq \infty$. This is actually a consequence of the fact that every layer of G is atomic. For if G_p contains an element c of a maximal order p^n, then the group $\{c\}=C(p^n)$ obviously exhausts all the elements of G_p. In case G_p does not contain an element of a maximal order, then the layer $L(p^n)$ of G consists, for each natural number n, of exactly $\varphi(p^n)$ elements, and, obviously, the same can be stated of the layer $L(p^n)$ of the subgroup pG_p. Hence we conclude $pG_p=G_p$. This implies the existence of an infinite sequence $c_1, c_2, c_3, \ldots$ of elements in G_p such that

$$c_1 \neq 0, \quad pc_1=0, \quad pc_2=c_1, \quad pc_3=c_2, \ldots .$$

In this case the subgroup generated by the elements $c_1, c_2, c_3, \ldots$ is a group of type $C(p^\infty)$. Of course, this can not be a proper subgroup of G_p, so that $G_p=C(p^\infty)$.

c) *implies* a). This is obvious, for if D and D' are distinct subgroups of (1), then the p-components of D and D' must differ at least for one prime p. But two different subgroups of the p-component of (1), i. e. of $C(p^\infty)$, are never isomorphic (already their orders are not equal).

(Received 30 May 1952)

О ГРУППАХ С АТОМНЫМИ СЛОЯМИ

Т. СЕЛЕ (Дебрецен)

(Р е з ю м е)

В настоящей работе доказывается, что произвольная группа G тогда и только тогда не содержит две различные, изоморфные между собой, подгруппы, если G изоморфна подгруппе (мултипликативной) группы всех корней из единицы.

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 du Commerce Extérieur des Livres et Journaux „Kultúra" (Budapest, VI., Sztálin-út 2. 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 2. Account No. 45-790-057-50-032) or with representatives abroad.

Die Acta Mathematica veröffentlichen Abhandlungen aus dem Bereiche der mathematischen Wissenschaften in russischer, französicher, 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 2. Bankkonto Nr. 45-790-057-50-032) oder bei seinen Auslandsvertretungen und Kommissionären.

32.— Ft.

INDEX

Kalmár, L., Another proof of the Markov—Post theorem.
Л. Кальмар, Новое доказательство теоремы Маркова—Post'a 1
Alexits, G., Sur l'ordre de grandeur de l'approximation d'une fonction périodique par les sommes de Fejér.
Г. Алексич, О порядке величины аппроксимации периодических функций Фейеровскими средними 29
Alexits, G., Über den Annäherungsgrad der Orthogonalpolynomentwicklungen.
Г. Алексич, О порядке величины аппроксимации рядами ортогональных многочленов 43
Sz.-Nagy, B., On the stability of the index of unbounded linear transformations.
Б. С.-Надь, Об устойчивости индекса неограниченных линейных преобразований 49
Atkinson, F. V., A spectral problem for completely continuous operators.
Ф. В. Аткинсон, Спектральная проблема с вполне непрерывным оператором 53
Sz.-Nagy, B., On a spectral problem of Atkinson.
Б. С.-Надь, Об одной спектральной проблеме Аткинсона 61
Molnár, J., Inhaltsabschätzung eines sphärischen Polygons.
И. Мольнар, Об оценке поверхностей сферических многоугольников 67
Georgiev, G., Sur certains automorphismes à points fixes des surfaces fermées orientables.
Г. Георгиев, О некоторых автоморфизмах с неподвижными точками нормальных замкнутых ориентируемых поверхностей 71
Tandori, K., Über die Cesàrosche Summierbarkeit der orthogonalen Polynomreihen.
К. Тандори, О суммабильности рядов ортогональных многочленов в смысле Чезаро 73
Freud, G., Über die starke $(C, 1)$-Summierbarkeit von orthogonalen Polynomreihen.
Г. Фрайд, О сильной суммабильности $(C, 1)$ ортогональных рядов многочленов 83
Freud, G., Über die Konvergenz orthogonaler Polynomreihen.
Г. Фрайд, О сходимости рядов ортогональных многочленов 89
Ljunggren, W., Eine elementare Auflösung der diophantischen Gleichung $x^3+1=2y^2$.
В. Люнгрен, Элементарное решение диофантового уравнения $x^3+1=2y^2$ . . 99
Fuchs, L., On subdirect unions, I.
Л. Фукс, О подпрямых суммах, I. 103
Kertész, A., On the decomposibility of abelian p-groups into the direct sum of cyclic groups.
А. Кертес, О прямом разложении абелевых p-групп в прямую сумму циклических групп 121
Szele, T., On groups with atomic layers.
Т. Селе, О группах с атомными слоями 127

Technikai szerkesztő: Fuchs László
Akadémiai Kiadó (Budapest, V., Alkotmány-u. 21.) Felelős: Mestyán János

Csongrádmegyei Nyomdaipari Vállalat, Szeged 522127 Felelős vezető: Vincze György

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS III. FASCICULUS 3.

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1952

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 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), külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin-út 21. Bankszámla 45-790-057-50-032 sz.) vagy külföldi képviseleteinél 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.

ON PROJECTIONS OF PROBABILITY DISTRIBUTIONS

By

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

J. RADON, in a paper [5] published in the year 1917, solved the following problem: It is to determine a continuous function $f(x, y)$, defined in a bounded domain K of the (x, y) plane, if given the values of the integral of this function along every chord of the domain K. From his results it follows, in particular, the following

THEOREM R. *If K is a bounded domain of the (x, y) plane, and the integral of the continuous function $f(x, y)$ vanishes along every chord of the domain K, then $f(x, y)$ is identically equal to zero.*

Since that this theorem has been independently re-discovered by many authors. Thus, it has been proved by H. STEINHAUS in 1941 in a lecture held at the conference of the University of Lwow; [1] at that time, Prof. H. STEINHAUS was unaware of the results of J. RADON. Recently he found this paper and kindly called my attention to it. My attention was called to the present problem by G. HAJÓS who raised the same problem in connection with a conjecture of S. TARSKI [7], which has been proved in the meantime by TH. BANG [1], [2]. The theorem of BANG reads as follows: If a *convex* domain K is covered by n parallel strips $S_1, S_2, \ldots, S_n$, the breadths of which are $d_1, d_2, \ldots, d_n$, respectively, then the sum $\sum_{k=1}^{n} d_k$ is greater than or equal to the breadth d of the domain K. Before making clear the connection between Theorem R and that of BANG, let us make the following remark: if for some domain K there exists a non-negative integrable function $f(x, y)$ whose integral along every chord of K is equal to 1, the statement of BANG's theorem follows easily for this domain, because if the strips $S_1, S_2, \ldots, S_n$ cover K, we have

$$\sum_{k=1}^{n} d_k = \sum_{k=1}^{n} \iint_{S_k} f(x, y)\, dx\, dy \geqq \iint_{K} f(x, y)\, dx\, dy = d. \tag{1}$$

[1] Cf. [6] where a short summary of the lecture can be found.

Such a function is known only for the circle; if K is the circle $x^2+y^2=1$, the function

$$f(x,y)=\frac{1}{\pi\sqrt{1-x^2-y^2}} \tag{2}$$

has the required properties. To prove this, it suffices, by reasons of symmetry, to consider only chords which are parallel to the y-axis; for the chord $x=a$, for example, we have

$$\frac{1}{\pi}\int\limits_{-\sqrt{1-a^2}}^{+\sqrt{1-a^2}}\frac{dy}{\sqrt{1-a^2-y^2}}=\frac{1}{\pi}\int\limits_{-1}^{+1}\frac{dz}{\sqrt{1-z^2}}=1. \tag{3}$$

This can be proved also without calculations by means of the well known geometrical fact that the surface of a segment of the sphere of radius 1 depends only on the height of the segment.

It is easy to see that such a function can exist only for domains K of constant breadth. As a matter of fact, for every position of the coordinate system we have

$$\iint\limits_{K} f(x,y)\,dx\,dy=\int\limits_{K} dy=\int\limits_{K} dx, \tag{4}$$

i. e. the breadth of the domain in the direction of the coordinate axes is equal to the constant $\iint\limits_{K} f(x,y)\,dx\,dy$ which is independent of the direction of the coordinate axes. It is not known whether there exists actually such a function for domains of constant breadth other than the circle.

Now we turn to considering the connection of Theorem R with TARSKI's conjecture. This consists in that if a function $f(x,y)$ having the property that its integral is constant along every chord of K exists at all for some domain K, one may ask whether it is unique or not. Theorem R shows that (if the continuity of $f(x,y)$ is also required) $f(x,y)$ is unique.[2] Theorem R has been independently re-discovered and generalized by I. SZARSKI and T. WAZEWSKI in their paper [8], in which a very simple proof of this theorem can be found. Further generalizations of Theorem R will be included in a paper of I. MIKUSINSKI and C. RYLL-NARDZEWSKI to be published in the *Studia Mathematica*, and in a paper of W. WOLIBNER under preparation. The purpose of the present paper is also to give a new proof and some generalizations (different from those already mentioned) of Theorem R.

It will be shown that the whole problem belongs essentially to probability theory, and can be attacked by analytical methods of probability theory, namely by the application of the theorem of unicity concerning characteristic functions.

[2] If continuity is not supposed, $f(x,y)$ can be modified on an arbitrary set, the common part of which with every straight line has the linear measure 0.

We show namely that Theorem R is a consequence of a theorem of H. CRAMÉR and H. WOLD [4], which can be formulated as follows:

THEOREM CW. *Every probability distribution on the plane is uniquely determined by the totality of its linear projections.*

Clearly, this theorem can be formulated also for mass distributions instead of probability distributions.

Chapter I contains the proof — which is essentially that of H. CRAMÉR and H. WOLD, and is included in the paper only to make it self-contained — of Theorem CW mentioned just now, as well as of a generalization of this theorem for spaces of 3 or more dimensions (Theorem CW′), further the proof of the theorem (Theorem 1) that, for a broad class of distributions, the knowledge of an arbitrary infinite set of different projections already determines the distribution uniquely. It follows from this theorem that to ensure that the continuous function $f(x, y)$ defined in the bounded domain K vanishes identically, it suffices to suppose that its integral vanishes along every chord parallel to some line belonging to an arbitrary enumerable infinite set of straight lines (Theorem 2). The paper leaves open the question whether or not this is true for every distribution.

Chapter II is devoted to the study of finite distributions, i. e. — using the terminology of mass distributions — of distributions consisting of a finite number of mass points, that is, points in which positive masses are concentrated. It has been conjectured by the author and proved by G. HAJÓS that a distribution consisting of n mass points in the plane is uniquely determined by $n+1$ arbitrary projections, but n projections are not always sufficient to determine the distribution (Theorem 3). The proof of this theorem is included in the paper with the kind permission of G. HAJÓS. We shall show that the same is true for n equal mass points in the space (Theorem 4).

The author expresses his sincere thanks to H. STEINHAUS, T. WAZEWSKI, M. FISZ and G. HAJÓS for their valuable remarks.

Chapter I

The mentioned theorem of J. RADON can be formulated also in the following equivalent form:

THEOREM R′. *A continuous and non-negative function $f(x, y)$ defined in the convex domain K, is uniquely determined if the value of its integral along every chord of K is given.*

Let us show that Theorem R′ follows from Theorem R, and conversely. If the value of the integral of the non-negative and continuous function $f(x, y)$ is the same along every chord as the value of the integral of the continuous and non-negative function $g(x, y)$, then the integral of $f(x, y) - g(x, y)$ vanishes along every chord and thus, according to Theorem R, we have

$$f(x, y) \equiv g(x, y).$$

Thus Theorem R′ follows from Theorem R. On the other hand, if the integral of the continuous function $f(x, y)$ vanishes along every chord of K, let us put

$$f_1(x, y) = \begin{cases} f(x, y) & \text{if } f(x, y) \geqq 0 \\ 0 & \text{if } f(x, y) < 0 \end{cases}$$

and

$$f_2(x, y) = f_1(x, y) - f(x, y);$$

it follows that $f_1(x, y)$ and $f_2(x, y)$ are continuous and non-negative, and the integrals of $f_1(x, y)$ and $f_2(x, y)$ have the same value for every chord of K. Hence from Theorem R′ we conclude that $f_1(x, y) \equiv f_2(x, y)$ and therefore we have $f(x, y) \equiv 0$; that is, Theorem R follows from Theorem R′.

Now instead of supposing that the integral of $f(x, y)$ is known along every straight line, we may suppose that the value of the integral

$$I(H) = \iint_{HK} f(x, y)\,dx\,dy \tag{1.2}$$

is known for every half-plane H where HK denotes the common part of the domain K and the half-plane H. In fact, if $I(H)$ is known for every half-plane, then the value of the integral

$$I(S) = \iint_{SK} f(x, y)\,dx\,dy \tag{1.3}$$

is known for every parallel strip S, and thus the value of the integral of $f(x, y)$ along every chord l can be calculated by means of the limit relation

$$i(l) = \int_l f(x, y)\,ds = \lim_{\Delta \to 0} \frac{1}{\Delta} \iint_{S_\Delta K} f(x, y)\,dx\,dy \tag{1.4}$$

where S_Δ is a parallel strip whose mid-line is l and whose breadth is Δ. Conversely, if $i(l)$ is known for every chord l, $I(H)$ can be calculated for every half-plane as $I(H) = \int_{-\infty}^{+\infty} i(l_x)\,dx$ where l_x denotes a chord which is parallel to the boundary line of H and which cuts the perpendicular to this line through the origin at a point having the abscissa x on this line. Thus we may suppose, instead of that $i(l)$ is known for every chord l, that $I(H)$ is known for every half-plane H.

Now the first step of generalization consists in that we omit the restriction of $f(x, y)$ being defined in a bounded domain, and consider functions $f(x, y)$ defined on the whole plane, but suppose that the integral $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\,dx\,dy$ is finite. Without restricting generality we may suppose that

(1.5) $$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\,dx\,dy = 1.$$

The second step in the generalization consists in that we omit the restriction that the non-negative function $f(x, y)$ should be continuous, and suppose only that it is L-integrable. Thus we may consider $f(x, y)$ as the density function of a probability distribution, and ask whether the values of the integral of this density function for every half-plane determine uniquely the density function, — or what is the same — the corresponding distribution function

(1.6). $$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)\,dv\,du.$$

Let l denote an arbitrary straight line through the origin, and let H_p denote the half-plane whose boundary line is perpendicular to l and cuts l in a point whose coordinate on l is equal to p. Clearly,

(1.7) $$F_l(p) = I(H_p) = \iint_{H_p} f(x, y)\,dx\,dy$$

as a function of p is nothing else than the distribution function of the projection on l of a random point of the plane whose distribution function is defined by (1.6). In what follows we shall call the linear distribution on l with the distribution function $F_l(p)$ the projection on l of the distribution on the plane with the distribution function $F(x, y)$. The last step of generalization consists in that we consider also distributions which have no density function and prove the following

THEOREM CW. *Let $F(x, y)$ denote the distribution function of an arbitrary probability distribution on the plane, and let us suppose that the projection of this distribution is known on every straight line l through the origin, i. e. that*

(1.8) $$F_{l_\varphi}(p) = \iint_{x\cos\varphi + y\sin\varphi \leqq p} dF(x, y)$$

is known as a function of p for every value of φ $(0 \leqq \varphi < \pi)$ where φ denotes the angle between the straight line l_φ and the x-axis. Then $F(x, y)$ is uniquely determined for every value of x and y.[3]

As it has been remarked in the introduction, this theorem is due to H. CRAMÉR and H. WOLD. We reproduce the simple proof of this theorem to make the paper self-contained.

PROOF. In what follows we shall denote by $M(\zeta)$ the mean value of a random variable ζ, and by $\Pr(A)$ the probability of the event A.

[3] The uncertainity of the value of $F(x, y)$ at its points of discontinuity does not come in, since we suppose (as usual) that $F(x, y)$ is continuous to the left as a function of x as well as a function of y.

Let (ξ, η) denote the coordinates of a random point on the plane having the distribution function $F(x, y)$. Let us denote by $\psi(u, v)$ the characteristic function of the point (ξ, η) (or, in other words, of the probability distribution with the distribution function $F(x, y)$), i. e. we put

$$\psi(u, v) = M(e^{i(u\xi+v\eta)}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{i(ux+vy)} dF(x, y). \tag{1.9}$$

The projection of the point (ξ, η) on the line l_φ has the coordinate $\xi\cos\varphi + \eta\sin\varphi = \zeta_\varphi$ and thus $F_{l_\varphi}(p)$ is the distribution function of ζ_φ. If $F_{l_\varphi}(p)$ is known as a function of p, its characteristic function

$$\psi_\varphi(t) = M(e^{it\zeta_\varphi}) = \int_{-\infty}^{+\infty} e^{itp} dF_{l_\varphi}(p) \tag{1.10}$$

is also known. But by (1.9) and (1.10) we have

$$\psi_\varphi(t) = M(e^{it\zeta_\varphi}) = M(e^{it(\xi\cos\varphi+\eta\sin\varphi)}) = \psi(t\cos\varphi, t\sin\varphi). \tag{1.11}$$

Thus it follows that, for every real value of u and v, we have

$$\psi(u, v) = \psi_{\operatorname{arc\,tg}\frac{v}{u}}(\sqrt{u^2+v^2}). \tag{1.12}$$

Hence $\psi(u, v)$ is known for every real value of u and v. As it is well known that a distribution function is uniquely determined by means of its characteristic function,[4] Theorem CW is completely proved.

Using the same method, the following theorem can also be proved.

THEOREM CW'. *A probability distribution in the n dimensional space is uniquely determined by its projections on such a set of subspaces of* $1, 2, \ldots, (n-1)$ *dimensions which together cover the whole space.*

Thus, for instance, a probability distribution in the space of 3 dimensions is uniquely determined by its projections on every straight line passing through the origin, or by its projections on every plane passing through a given line, or else by its projections on any collection of straight lines and planes which cover together the whole space. The proof of Theorem CW gives at the same time a criterion for a set of distribution functions $F_{l_\varphi}(p)$ to be the projections of a plane distribution. Clearly, the necessary and sufficient condition of this consists in that $\psi_{\operatorname{arc\,tg}\frac{v}{u}}(\sqrt{u^2+v^2})$ shall be a characteristic function where

$$\psi_\varphi(t) = \int_{-\infty}^{+\infty} e^{itp} dF_{l_\varphi}(p). \tag{1.13}$$

By the same method we can prove also the following

THEOREM 1. *If the point (ξ, η) is contained with probability* 1 *in a circle* $\xi^2+\eta^2 \leq R^2$ *and the distribution of the projection of the point (ξ, η) is given*

[4] See e. g. [8], p. 101.

on an arbitrary infinite set of straight lines through the origin, i. e. the distribution function $F_{l_\varphi}(p)$ *of the random variable* $\zeta_\varphi = \xi\cos\varphi + \eta\sin\varphi$ *is given for an infinity of* mod π *different values of* φ, *then the distribution function* $F(x, y) = \Pr(\xi < x, \eta < y)$ *of the random point* (ξ, η) *is uniquely determined.*

Before proving Theorem 1, let us formulate as a theorem a corollary of this theorem which is a straightforward generalization of Theorem R.

THEOREM 2. *Let* $f(x, y)$ *denote a continuous function which is equal to* 0 *if* $x^2+y^2 \geq R^2$ *for some* $R > 0$. *If the integral of* $f(x, y)$ *vanishes along every line parallel to some line belonging to an arbitrary infinite set of lines passing through the origin, then* $f(x, y) \equiv 0$.[5]

PROOF of THEOREM 1. Let us denote by $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ the values of φ for which $F_{l_\varphi}(p)$ is known. Let φ_0 denote a limit point of the sequence φ_n. According to (1. 11) $\psi(t\cos\varphi, t\sin\varphi)$ is known for every value of t and for every $\varphi = \varphi_n$. As $\psi(t\cos\varphi, t\sin\varphi)$ is an analytic function of φ for every fixed value of t, it follows that $\psi(t\cos\varphi, t\sin\varphi)$ is known for any fixed value of t for values $\varphi = \varphi_{n_k}$ where $\lim\limits_{k\to\infty} \varphi_{n_k} = \varphi_0$. Hence we conclude that $\psi(t\cos\varphi, t\sin\varphi) = \psi_\varphi(t)$ is known for *every* value of t and φ, and thus Theorem 1 follows in the same way as Theorem CW was proved. The analyticity of $\psi_\varphi(t)$ is clear from the existence of the derivative

$$(1.\ 15) \qquad \frac{\partial \psi_\varphi(t)}{\partial \varphi} = it \iint\limits_{x^2+y^2 \leq R^2} (-x\sin\varphi + y\cos\varphi) e^{it(x\cos\varphi + y\sin\varphi)} dF(x, y)$$

for every (complex) value of φ.

Chapter II

In this chapter we consider discrete distributions. For the sake of simplicity we shall use the terminology of mass distributions. Let us consider a discrete mass distribution on the plane, consisting of n mass points, i. e. a distribution consisting of the masses $m_k > 0$ situated in the points (x_k, y_k) $(k = 1, 2, \ldots, n)$. We shall prove first the following

THEOREM 3. *A discrete mass distribution consisting of n distinct mass points with masses* $m_1, m_2, \ldots, m_n$ *situated in the points* $(x_1, y_1), (x_2, y_2), \ldots,$

[5] It will be seen from the proof of Theorem 1 that Theorem 2 also holds if instead of supposing $f(x, y) = 0$ for $x^2+y^2 \geq R^2$ we suppose only that $f(x, y)$ is sufficiently small for large values of x^2+y^2, for example, if for every $\lambda > 0$ we have

$$(1.\ 14) \qquad |f(x, y)| \leq e^{-\lambda\sqrt{x^2+y^2}}$$

provided that $\sqrt{x^2+y^2} \geq R(\lambda)$ where $R(\lambda)$ is an arbitrary positive function of λ. This will become clear by taking into account that, in the proof of Theorem 1, the condition that $\xi^2+\eta^2 \leq R^2$ is fulfilled with probability 1 serves only to ensure that the characteristic function $\psi(p\cos\varphi, p\sin\varphi)$ should be, for any value of p, an analytic function of φ, and this is ensured by the restriction (1. 14).

(x_n, y_n), respectively, is completely determined if its projections on $n+1$ arbitrary different straight lines through the origin are given.

I have proved Theorem 3 only for the case of equal masses, in which case Theorem 3 is a special case of Theorem 4, and communicated the assertion for unequal masses as a conjecture to G. HAJÓS who succeeded in proving it. I am very thankful to him for his kind permission to publish here his elegant proof.

PROOF OF THEOREM 3. For $n=1$ the theorem is trivial. Let us suppose $n \geqq 2$. Let us mark the two extreme points of each projection and consider the projecting lines (i. e. the perpendiculars to the line on which the projection is considered) through these points; we shall call these lines, for the sake of brevity, *extreme projecting lines*. Thus if $n+1$ projections of the mass distribution are known, we have at least $2n+1$ extreme projecting lines, because at least n projections have two extreme points and only one projection can eventually shrink to a point (and this can happen only in case all points are lying on the same straight line). Each extreme projecting line passes through at least one mass point. As there are n mass points, there must be at least one mass point through which three or more extreme projecting lines are passing. Since all mass points are situated in one of the two closed half-planes determined by every extreme projecting line, we see that if $r \geqq 3$ extreme projecting lines pass through a point P of the plane, these lines divide the plane into $2r$ angular domains, and all mass points must lie in the interior or on the boundary of one of these domains. This domain is bounded by two extreme projecting lines, therefore all other extreme projecting lines, and thus at least one projecting line can have no common point with our mass system other than the point P itself. But since every extreme projecting line passes through at least one mass point, we infer that P itself must be a mass point. Thus we have proved that there is at least one mass point through which three or more projecting lines are passing, and conversely: every point of the plane which is common to three or more extreme projecting lines is a mass point. Consequently, considering only the extreme projecting lines, at least one mass point can be found. The projections of the remaining $n-1$ points on $n+1$ lines being known (by omitting the projection of the point already found), we can apply the same procedure again, and thus find one-by-one all mass points. Therefore Theorem 3 is proved. Clearly, the above proof furnishes an effective method for actually finding all mass points, and the corresponding masses.

It is easy to see that Theorem 3 can not be improved: n projections do not always determine a discrete mass distribution consisting of n points. As a matter of fact, let us consider a regular polygon Π with $2n$ sides, and let the system of n equal mass points, each of mass 1, situated in every second vertex of the polygon Π be called *system A*, and let the system of

n equal mass points, each of mass 1, situated in those n vertices of the polygon Π in which there is no mass point of the system A, be called *system* B. It is easy to see that, denoting by $l_1, l_2, \ldots, l_n$ the perpendiculars to the pairs of opposite sides of the polygon Π, the projection of the system A is the same on each line l_i as the projection of the system B. Analyzing the above proof, it is easy to see that all mass distributions which are not completely determined by n projections are essentially equivalent to that mentioned just now, and can be obtained by replacing the regular polygon of $2n$ sides by some convex polygon of $2n$ sides, with vertices $P_1, P_2, \ldots, P_{2n}$ and having the following property: the lines P_iP_j and P_kP_l are parallel provided that $i+j \equiv 1 \pmod 2$ and $i+j \equiv k+l \pmod{2n}$. Clearly, all polygons obtained from a regular polygon of $2n$ sides by an affine transformation satisfy this condition, but not only these; for instance, the hexagon $P_1P_2P_3P_4P_5P_6$ shown by Fig. 1 has the required property, though it cannot be obtained by an affine transformation from a regular hexagon.

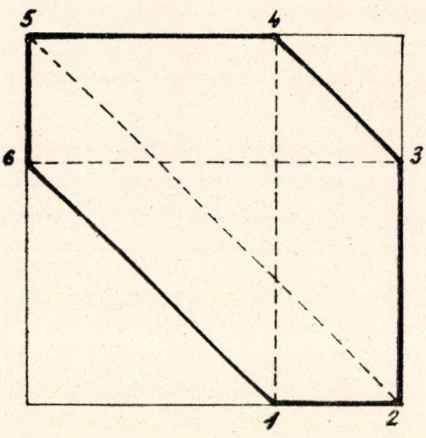


Fig. 1

Instead of speaking of projections *on* some straight line l we can speak of projections *from* the direction perpendicular to l on some fixed straight line L.

Projections from a direction can be considered as projections from some point at infinity of the projective plane. It is easy to see that the proof, and hence the assertion of Theorem 3, remains valid for the more general case when projections from finite points to l are also admitted.

Let us consider now projections of discrete mass systems of the 3-dimensional space. We prove the following

THEOREM 4. *Let us consider a discrete mass distribution M in the 3-dimensional (x, y, z)-space, consisting of n equal masses situated in the points with the rectangular coordinates (x_k, y_k, z_k) $(k = 1, 2, \ldots, n)$. If the (orthogonal)* *projection of the mass distribution M is given on $n+1$ arbitrary planes, no two of which are parallel, then M is completely determined.*

Before proving Theorem 4, let us mention that the theorem can not be improved. As a matter of fact, if the masses are situated in every second vertex of a regular polygon of $2n$ sides in some plane α and the planes on which these masses are projected are all orthogonal to α, we obtain the counter-example, considered in connection with Theorem 3.

PROOF OF THEOREM 4. Let us denote by $A_1, A_2, \ldots, A_{n+1}$ the planes on which the mass distribution is projected. We may suppose that all planes A_k pass through the origin of the rectangular coordinate system (x, y, z) and that they do not pass through the z-axis, and the line of intersection of two of them does not lie in the (y, z)-plane. Let us choose a rectangular coordinate system (u_k, v_k) in each plane A_k such that its origin coincides with the origin of the coordinate system (x, y, z) and let us denote by $\alpha_{1k}, \beta_{1k}, \gamma_{1k}$ the cosines of direction of the straight line $v_k=0$ and by $\alpha_{2k}, \beta_{2k}, \gamma_{2k}$ the cosines of direction of the straight line $u_k=0$. It follows that the projection of the point (x_j, y_j, z_j) on A_k has the coordinates

$$\left.\begin{aligned} u_{jk} &= \alpha_{1k}x_j+\beta_{1k}y_j+\gamma_{1k}z_j \\ v_{jk} &= \alpha_{2k}x_j+\beta_{2k}y_j+\gamma_{2k}z_j \end{aligned}\right\} \quad \begin{aligned} j &= 1, 2, \ldots, n \\ k &= 1, 2, \ldots, n+1 \end{aligned}$$

in the coordinate system (u_k, v_k). Therefore, if these projections are given, the numbers

$$\frac{\alpha_{2k}u_{jk}-\alpha_{1k}v_{jk}}{\beta_{1k}\alpha_{2k}-\beta_{2k}\alpha_{1k}} = y_j+\lambda_k z_j$$

are known where

$$\lambda_k = \frac{\gamma_{1k}\alpha_{2k}-\gamma_{2k}\alpha_{1k}}{\beta_{1k}\alpha_{2k}-\beta_{2k}\alpha_{1k}}.$$

As $\alpha_{2k}\gamma_{1k}-\alpha_{1k}\gamma_{2k}$ and $\alpha_{2k}\beta_{1k}-\alpha_{1k}\beta_{2k}$ are two cosines of direction of the perpendicular to the plane A_k, which does not pass through the z-axis, the second is different from zero, and as the line of intersection of two planes A_k and $A_{k'}$ $(k' \neq k)$ does not lie in the (y, z)-plane, their ratio λ_k is different for different values of k. Thus the numbers $y_j+\lambda z_j$ $(j=1, 2, \ldots, n)$ are known in their totality for $n+1$ different values of λ. Consequently, all elementary symmetric functions of these numbers,

$$S_1(\lambda) = \sum_{j=1}^{n}(y_j+\lambda z_j),$$

$$S_2(\lambda) = \sum_{1\leq j_1<j_2\leq n}(y_{j_1}+\lambda z_{j_1})(y_{j_2}+\lambda z_{j_2}),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$S_n(\lambda) = \prod_{j=1}^{n}(y_j+\lambda z_j),$$

are known for $n+1$ different values of λ. As $S_1(\lambda), S_2(\lambda), \dots, S_n(\lambda)$ are polynomials of degree not greater than n in λ, it follows that these polynomials are completely determined, and therefore their values can be calculated (e. g. by NEWTON's interpolation formula) for $\lambda = i$. Hence we can obtain the values of $S_1(i), S_2(i), \dots, S_n(i)$, i. e. the values of the elementary symmetric functions of the complex numbers $w_j = y_j + iz_j$. We conclude that the complex numbers w_j can be determined as the roots of the equation

$$w^n - S_1(i)w^{n-1} + S_2(i)w^{n-2} - \dots + (-1)^n S_n(i) = 0$$

and hence the pairs of numbers (y_j, z_j) can be obtained. Therefore, starting with the projection of the mass distribution considered onto the planes A_k $(k = 1, 2, \dots, n+1)$, we can determine the projection of the same mass distribution onto the plane (y, z). As the position of the coordinate system (x, y, z) is arbitrary (we have to take care only of that no plane A_k should pass through the z-axis and the line of intersection of two planes A_j, A_k should not lie in the plane (y, z)), it follows that the projection of the considered distribution can be determined for *every* plane except those planes which pass through the intersecting line of two planes A_j, A_k; but the projections on such planes can also be determined by a limiting process and therefore, according to Theorem CW, the distribution itself is completely determined. Theorem 4 is herewith proved.

(Received 30 August 1952)

Literature

[1] TH. BANG, On covering by parallel strips, *Mat. Tidsskrift* B, **1950**, pp. 49—53.

[2] TH. BANG, A solution of the "plank problem", *Proc. Amer. Math. Soc.*, **2** (1951), pp. 990—992.

[3] H. CRAMÉR, *Mathematical methods of statistics*.

[4] H. CRAMÉR and H. WOLD, Some theorems on distribution functions, *Journal London Math. Soc.*, **11** (1936), pp. 290—294.

[5] J. RADON, Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Ber. Math.-Phys. Kl. Sächs. Ges. Wiss. Leipzig*, **59** (1917), pp. 262—277.

[6] H. STEINHAUS, Sur un problème de G. Hajós, *Colloquium Mathematicum*, 1951, p. 161, (Summary).

[7] A. TARSKI, Uwagi o stopniu rownowazności wielokatov, *Parametr*, **2** (1932), pp. 310—314.

[8] T. WAZEWSKI et I. SZARSKI, Sur un problème roentgenographique de M. S. Majerek. *Ann. Soc. Polonaise de Math.*, **20** (1947), pp. 389—390.

О ПРОЕКЦИЯХ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ

А. РЕНЬИ (Будапешт)

(Резюме)

Глава I работы содержит доказательство и обобщение следующей теоремы И. Радона: Если интеграл непрерывной функции $f(x, y)$ ровно 0 на каждом хорде некоторой конечной области K, то $f(x, y) \equiv 0$ на K.

Эта теорема была открыта многими авторами независимо друг от друга. В работе показано, что эта теорема является прямым следствием теоремы Х. Крамера и Х. Волда согласно которому распределение вероятностей в плоскости однозначно определено с помощю совокупности его проекциях на всех прямых плоскости. Дальше доказано, что для определении этого распределении достаточно знание его проекций на бесконечно многих различных прямых, проходящих через данную точку.

В главе II и рассматриваются дискретные распределения вероятностей (или масс) и изложено доказательство данное Г. Гаёшом, теоремы, что если известны проекции дискретного распределения, состоящего из n точечных масс плоскости на $n+1$ различных приямых не параллелных между собой, то это распределения однозначно определено. То же самое имеет место для дискретных распределении в пространстве, если все массы одинаковы; знание n различных проекции для этого вообще не достаточно.

ÜBER DIE DETERMINANTENTEILER

Von

L. RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

Bezeichne R einen kommutativen Ring mit Einselement, das wir ohne Gefahr einer Verwechslung mit 1 bezeichnen, ebenso wie die natürliche Zahl 1. Bezeichne ferner M eine rechteckige Matrix mit Elementen aus R. Wir nennen das durch alle Unterdeterminanten k-ter Ordnung von M erzeugte Ideal in R den k-ten *Determinantenteiler* von M und bezeichnen ihn mit $\mathfrak{D}_k$. Unter einer Determinante 0-ter Ordnung verstehen wir das Einselement 1 von R und dann ist auch $\mathfrak{D}_0 = 1$. Ferner setzen wir bequemlichkeitshalber $\mathfrak{D}_k = 0$, falls k größer ist als die Anzahl der Zeilen oder Spalten von M.

Für gewöhnlich pflegt man $\mathfrak{D}_k$ nur im Fall eines Euklidischen Ringes R in Betracht zu ziehen. Dann ist R bekanntlich ein Hauptidealring und entsprechend läßt sich $\mathfrak{D}_k$ einfach als der größte gemeinsame Teiler der genannten Unterdeterminanten von M definieren. In diesem Spezialfall gelten die wohlbekannten Teilbarkeitsrelationen

$$(1) \qquad \mathfrak{D}_k^2 \mid \mathfrak{D}_{k-1}\,\mathfrak{D}_{k+1} \qquad (k = 1, 2, \ldots).$$

Diese sind ja im wesentlichen identisch mit den Teilbarkeitsrelationen

$$(2) \qquad e_k \mid e_{k+1} \qquad (k = 1, 2, \ldots)$$

der Elementarteiler $e_1, e_2, \ldots$ von M, die mit den $\mathfrak{D}_k$ durch $\mathfrak{D}_k = e_1 \ldots e_k$ zusammenhängen. Nach STEINITZ[1] gilt ähnliches allgemeiner für die Dedekindschen Ringe — so nennt man die R mit eindeutiger Primidealzerlegung — dann sind aber $\mathfrak{D}_k$ und e_k Ideale in R.

Im allgemeinen Fall lassen sich keine Elementarteiler ähnlich definieren, es ist auch nicht zu erwarten, daß die den Teilbarkeitsrelationen (1) idealtheoretisch

[1] E. STEINITZ, Rechteckige Systeme und Moduln in algebraischen Zahlkörpern, Teil I, *Math. Annalen*, **71** (1912), S. 328—354; Teil II, *Math. Annalen*, **72** (1912), S. 297—345. S. auch K. ASANO, Über Moduln und Elementarteilertheorie im Körper, in dem Arithmetik definiert ist, *Japanese Journal Math.*, **20** (1950), S. 55—71, W. FRANZ, Elementarteilertheorie in algebraischen Zahlkörpern, *Journal f. reine u. angewandte Math.*, **171** (1934), S. 149—161, C. CHEVALLEY, *L'arithmétique sur les algèbres de matrices* (Paris, 1936), Actualités Sci. Ind., no. 323, S. 5—33, und H. FITTING, Die Determinantenideale eines Moduls, *Jahresbericht d. D. M. V.*, **46** (1936), S. 195—228.

entsprechenden Enthaltungsrelationen

(3) $$\mathfrak{D}_{k-1}\mathfrak{D}_{k+1} \subseteq \mathfrak{D}_k^2 \qquad (k=1,2,\ldots)$$

gültig bleiben. Später werden wir in der Tat ein Beispiel angeben, wo (3) nicht gilt. Merkwürdigerweise ist aber (3) in etwas schwächerer Form allgemein richtig, und zwar beweisen wir den folgenden

SATZ 1. *Für die Determinantenteiler $\mathfrak{D}_0, \mathfrak{D}_1, \ldots$ einer Matrix M über einem kommutativen Ring R mit Einselement gilt*

(4) $$\bar{k}\,\mathfrak{D}_{k-1}\mathfrak{D}_{k+1} \subseteq \mathfrak{D}_k^2 \qquad (k=1,2,\ldots),$$

wobei $\bar{k}$ das kleinste gemeinsame Vielfache von $1,\ldots,k$ bezeichnet. (Selbstverständlich soll $\bar{k}$ in (4) als das durch das $\bar{k}$-fache Einselement von R erzeugte Hauptideal aufgefaßt werden.)

Vor dem Beweis fügen wir zu diesem Satz einige Bemerkungen hinzu. Für beliebige Ideale $\mathfrak{a}, \mathfrak{b}$ von R bilden die ganzen rationalen Zahlen c mit der Eigenschaft $c\mathfrak{a} \subseteq \mathfrak{b}$ offenbar ein Ideal im Ring der ganzen rationalen Zahlen. (Selbstverständlich kann dieses Ideal auch 0 sein.) Hieraus und aus Satz 1 folgt sofort, daß es zu jedem R eine wohlbestimmte unendliche Folge

(5) $$c_1, c_2, \ldots$$

von positiven ganzen Zahlen gibt, mit denen

(6) $$c_k\mathfrak{D}_{k-1}\mathfrak{D}_{k+1} \subseteq \mathfrak{D}_k^2 \qquad (k=1,2,\ldots)$$

für alle Matrizen M über R gilt und dabei alle $c_1, c_2, \ldots$ möglichst klein gewählt sind, außerdem

(7) $$c_k \mid \bar{k} \qquad (k=1,2,\ldots)$$

ist. Endlich muß auch

(8) $$c_k \mid c_{k+1}$$

gelten, das wir sofort beweisen. Wir betrachten neben M auch die um eine Zeile und Spalte geränderte Matrix

$$M' = \begin{pmatrix} 1 & \\ & M \end{pmatrix},$$

wo an den leeren Stellen lauter Elemente 0 stehen. Für die Determinantenteiler $\mathfrak{D}'_0, \mathfrak{D}'_1, \ldots$ von M' gilt offenbar

$$\mathfrak{D}'_{k+1} = \mathfrak{D}_k \qquad (k=0,1,\ldots).$$

Andererseits gilt nach (6) auch

$$c_{k+1}\mathfrak{D}'_k\mathfrak{D}'_{k+2} \subseteq \mathfrak{D}'^2_{k+1} \qquad (k=1,2,\ldots).$$

Da sich hierfür $c_{k+1}\mathfrak{D}_{k-1}\mathfrak{D}_{k+1} \subseteq \mathfrak{D}_k^2$ schreiben läßt, so folgt hieraus nach Vergleich mit (6) die behauptete Eigenschaft (8).

Die Folge (5) wollen wir etwa den *Defekt* des Ringes R nennen. Dieser Begriff scheint uns hauptsächlich für die Integritätsbereiche von Wichtigkeit zu sein. Ist (5) gleich $1, 1, \ldots$, so nennen wir R einen Ring *ohne Defekt*. Insbesondere sind die Euklidischen und allgemeiner die Dedekindschen Ringe wegen (3) ohne Defekt. Wir wissen nicht, ob auch weitere Integritätsbereiche ohne Defekt existieren, und auch nicht, ob es Ringe mit dem (maximalen) Defekt $1, 2, \ldots$ gibt. (Wenn ja, so muß offenbar der Polynomring von unendlich vielen Unbestimmten über dem Ring der ganzen Zahlen so beschaffen sein.)

Der Beweis von Satz 1 wird als leichte Folgerung aus dem folgenden auch an sich interessanten determinantentheoretischen Satz gewonnen.

SATZ 2. *Bezeichne D eine Determinante $2k$-ter Ordnung ($k \geqq 1$), d den Hauptminor $k+1$-ter Ordnung, Δ die Adjungierte von d. Es gilt*

$$d\Delta = \sum_{d^*} (-1)^t \frac{1}{k\binom{k-1}{t}} d^* \Delta^*, \tag{9}$$

wobei d^* diejenigen Unterdeterminanten k-ter Ordnung der aus den ersten $k+1$ Spalten von D bestehenden Matrix M^* durchläuft, die die erste Zeile von M enthalten; t bezeichnet die Anzahl derjenigen Zeilen von d^*, die unter den letzten $k+1$ Zeilen von M^* sind, Δ^* bezeichnet die Adjungierte von d^* (in D).[2]

Um zunächst Satz 2 zu beweisen, führen wir die folgenden Bezeichnungen ein. Wir bezeichnen mit $\{x\}$ die Menge der Zahlen $1, \ldots, x$ für eine natürliche Zahl x und mit $O(X)$ die Anzahl der Elemente einer endlichen Menge X. Es sollen A, B zwei Mengen bezeichnen, für die

$$O(A) = O(B) = k; \quad 1 \in A; \quad A \subset \{2k\}; \; B \subset \{k+1\} \tag{10}$$

gilt. Bezeichne noch (A, B) das Produkt der zu den Zeilenindizes $i \in A$ und Spaltenindizes $j \in B$ gehörigen Unterdeterminante von D und ihrer Adjungierten. Dann schreibt sich die rechte Seite von (9) als

$$\sum_A \sum_B (-1)^t \frac{1}{k\binom{k-1}{t}} (A, B), \tag{11}$$

wobei

$$t = O(A \cap (\{2k\} - \{k+1\})) \tag{12}$$

ist und $\{2k\} - \{k+1\}$ die Komplementäre von $\{k+1\}$ in $\{2k\}$ bezeichnet. Wir haben zu zeigen, daß (11) der linken Seite von (9) gleich ist.

[2] Eigentlich hat die rechte Seite von (9) nur einen Sinn, wenn der Ring, aus dem die Elemente von D genommen sind, die rationalen Zahlen enthält. Aber auch für beliebige Ringe behält (9) seine Gültigkeit, wenn man rechts die Nenner wegschafft, so daß man beiderseits mit dem kleinsten gemeinsamen Vielfachen der Nenner multipliziert. Dies denken wir nötigenfalls stets ausgeführt.

Hierzu betrachten wir eine Permutation

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & 2k \\ 1' & 2' & \cdots & (2k)' \end{pmatrix} \tag{13}$$

und bezeichnen mit G_π dasjenige Entwicklungsglied von D, das zu der Permutation π der Zeilenindizes $1, \ldots, 2k$ gehört. Offenbar kommt G_π unter den Entwicklungsgliedern von (A, B) dann und nur dann vor, wenn

$$A = B' \tag{14}$$

ist, wobei B' die Menge der durch π hervorgerufenen Bilder der Elemente von B bezeichnet.

Damit (14) stattfinden kann, hierzu ist nach (10) eine Vorbedingung, daß

$$1 \in \{k+1\}' \tag{15}$$

gilt, d. h. 1 unter den Bildelementen $1', \ldots, (k+1)'$ vorkommt. Deshalb dürfen wir uns im folgenden auf die π beschränken, die der Bedingung (15) genügen.

Wir setzen für ein solches π

$$z = O(\{k+1\}' \cap (\{2k\} - \{k+1\})), \tag{16}$$

wobei also z die Anzahl derjenigen Elemente x von $\{k+1\}$ ist, für die das Bildelement x' in die Komplementäre von $\{k+1\}$ fällt. Offenbar gilt dann

$$0 \leqq z \leqq k-1. \tag{17}$$

Endlich bezeichnen wir mit j dasjenige Element von $\{k+1\}$, wofür nach (15)

$$j' = 1 \qquad (j \in \{k+1\}) \tag{18}$$

gilt.

Welche sind nun die den Bedingungen (10) und (14) genügenden Paare A, B und wie groß fällt t jedesmal aus? Vor allem sind die passenden B wegen (18) offenbar die k Mengen, die aus $\{k+1\}$ durch Streichung je eines Elementes $i \neq j$ entstehen. Dann liefert (14) auch schon das zugehörige A, denn hierfür ist (10) erfüllt. Ferner gilt nach (12) und (16) $t = z$ oder $t = z-1$, je nachdem $i' \in \{k+1\}$ oder $i' \notin \{k+1\}$. Die Anzahl dieser i ist nach (16) gleich $k-z$ bzw. z.

Folglich ist der Koeffizient von G_π in (11) gleich

$$(-1)^z \frac{k-z}{k\binom{k-1}{z}} + (-1)^{z-1} \frac{z}{k\binom{k-1}{z-1}}, \tag{19}$$

wobei im Fall $z = 0$ das zweite Glied außer acht zu lassen ist. Man sieht, daß (19) für $z = 0$ gleich 1 und für die übrigen z in (17) gleich 0 ist. Wir haben bekommen, daß (11) aus denjenigen G_π besteht, für die in (16) $z = 0$ gilt. Diese Forderung ist mit der Bedingung $\{k+1\}' = \{k+1\}$ identisch. Ein Blick auf (13) zeigt, daß diese G_π eben die Entwicklungsglieder des Produktes $d\Delta$ ausmachen, womit wir Satz 2 bewiesen haben.

Um nunmehr Satz 1 zu beweisen, berechnen wir zunächst das kleinste gemeinsame Vielfache der in (9) vorkommenden Nenner, die sich auch als

$$(k-t)\binom{k}{t} \qquad (t=0,\ldots,k-1) \tag{20}$$

schreiben lassen. Diese Zahlen sind dieselben wie

$$t\binom{k}{t}=\frac{k!}{(t-1)!\,(k-t)!} \qquad (t=1,\ldots,k).$$

In dieser Zahl ist der Exponent einer beliebigen Primzahl p gleich

$$\sum_{i=1}^{m}\left(\left[\frac{k}{p^i}\right]-\left[\frac{t-1}{p^i}\right]-\left[\frac{k-t}{p^i}\right]\right), \tag{21}$$

wobei $[z]$ für reelle z die größte ganze Zahl $\leqq z$ bezeichnet und $m(\geqq 0)$ die größte ganze Zahl mit

$$p^m \leqq k$$

ist. Hiernach ist (21) höchstens gleich m. Andererseits nimmt (21) insbesondere für $t=p^m$ diesen Höchstwert m offenbar auf. Folglich ist das kleinste gemeinsame Vielfache der Zahlen (20) das Produkt aller p^m, d. h. eben gleich der im Satz 1 mit $\bar{k}$ bezeichneten Zahl.

Multipliziert man (9) durch diese Zahl $\bar{k}$, so bekommt man, daß $\bar{k}d\Delta$ linear von den $d^*\Delta^*$ mit ganzzahligen Koeffizienten abhängt.

Hieraus gewinnt man leicht den Beweis von Satz 1. Hierzu bezeichnen wir mit d_{k-1}, d_{k+1} je eine beliebige Unterdeterminante von M von der Ordnung $k-1$ bzw. $k+1$. Dabei darf angenommen werden, daß diese Unterdeterminanten wirklich existieren, d. h. daß M mindestens $k+1$ Zeilen und Spalten enthält, weil sonst der Satz trivial ist. Aus vier Matrizen M konstruieren wir durch Nebeneinandersetzen die Matrix

$$M_1=\begin{pmatrix} M & M\\ M & M\end{pmatrix}.$$

Es ist klar, daß die von 0 verschiedenen Unterdeterminanten von M_1 vom Vorzeichen abgesehen dieselben sind, wie die von M. Diese Eigenschaft überträgt sich auch auf jede solche Matrix M_2, die aus M_1 durch Permutation der Zeilen und Spalten entsteht. Insbesondere läßt sich M_2 so wählen, daß sie eine Unterdeterminante D von der Ordnung $2k$ hat, deren Hauptminor von der Ordnung $k+1$ eben d_{k+1} und die Komplementäre dieses Minors eben d_{k-1} ist. Wendet man die vorige Bemerkung auf diese Determinante D an, so folgt, daß $\bar{k}d_{k+1}d_{k-1}$ eine bilineare Form der Unterdeterminanten k-ter Ordnung von D, d. h. auch von M ist, wobei lauter ganze Zahlen als Koeffizienten auftreten. Dies bedeutet $\bar{k}d_{k+1}\,d_{k-1}\in\mathfrak{D}_k^2$, womit Satz 1 bewiesen ist.

Das einfachste nichttriviale Beispiel für Satz 2 liefert der Fall $k=2$. Dann handelt es sich um eine Determinante $|x_{ij}|$ von 4-ter Ordnung, wofür

(9) zu folgender Identität führt:

$$\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} x_{44} = \frac{1}{2} \Bigg\{ \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} \begin{vmatrix} x_{33} & x_{34} \\ x_{43} & x_{44} \end{vmatrix} - \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix} \begin{vmatrix} x_{32} & x_{34} \\ x_{42} & x_{44} \end{vmatrix} + \begin{vmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{vmatrix} \begin{vmatrix} x_{31} & x_{34} \\ x_{41} & x_{44} \end{vmatrix} -$$
$$- \begin{vmatrix} x_{11} & x_{12} \\ x_{31} & x_{32} \end{vmatrix} \begin{vmatrix} x_{23} & x_{24} \\ x_{43} & x_{44} \end{vmatrix} + \begin{vmatrix} x_{11} & x_{13} \\ x_{31} & x_{33} \end{vmatrix} \begin{vmatrix} x_{22} & x_{24} \\ x_{42} & x_{44} \end{vmatrix} - \begin{vmatrix} x_{12} & x_{13} \\ x_{32} & x_{33} \end{vmatrix} \begin{vmatrix} x_{21} & x_{24} \\ x_{41} & x_{44} \end{vmatrix} -$$
$$- \begin{vmatrix} x_{11} & x_{12} \\ x_{41} & x_{42} \end{vmatrix} \begin{vmatrix} x_{23} & x_{24} \\ x_{33} & x_{34} \end{vmatrix} + \begin{vmatrix} x_{11} & x_{13} \\ x_{41} & x_{43} \end{vmatrix} \begin{vmatrix} x_{22} & x_{24} \\ x_{32} & x_{34} \end{vmatrix} - \begin{vmatrix} x_{12} & x_{13} \\ x_{42} & x_{43} \end{vmatrix} \begin{vmatrix} x_{21} & x_{24} \\ x_{31} & x_{34} \end{vmatrix} \Bigg\}.$$

(Man beachte, daß die sechs Elemente $x_{14}, x_{24}, x_{34}, x_{41}, x_{42}, x_{43}$ links gar nicht vorkommen, also rechts herausfallen müssen, wovon man sich leicht auch direkt überzeugt.) Wir haben nachgerechnet, daß — im Falle unbestimmter $x_{11}, \ldots, x_{44}$ — die neun Glieder der rechten Seite linear unabhängig sind, und so hat man es mit einer „unverkürzbaren Darstellung“ der linken Seite zu tun. Auch läßt sich die linke Seite nicht mit ganzen Koeffizienten linear durch Produkte von Determinanten 2-ter Ordnung darstellen. Aus diesem Beispiel folgt, daß (3) nicht immer gilt, d. h. daß es Ringe mit Defekt gibt.

Neben dem engen Zusammenhang von Satz 2 mit Satz 1 verdient Satz 2 auch ein selbständiges Interesse. Um dieses deutlich hertvortreten zu lassen, bezeichnen wir jetzt mit D die allgemeine Determinante n-ter Ordnung, worunter wir hier verstehen, daß die Elemente von D Unbestimmte über dem rationalen Zahlkörper K sind. Das Produkt $d\Delta$ einer beliebigen Unterdeterminante d mit der Adjungierten Δ nennen wir einen *Laplaceschen Bestandteil* (kurz L. B.) von D. Ist die Ordnung von d gleich m, also die von Δ gleich $n—m$, so nennen wir den Absolutwert $|n—2m|$ der Differenz beider Ordnungszahlen die *Stufe* des betrachteten L. B. Selbst $D = 1 \cdot D = D \cdot 1$ ist auch ein L. B., und zwar der einzige von der Stufe n. Die sämtlichen Stufenzahlen sind offenbar $0, 2, \ldots, n$ bzw. $1, 3, \ldots, n$, je nachdem n gerade oder ungerade ist. Da sich Satz 2 nach passender Änderung offenbar auf jede Unterdeterminante d von der Ordnung $k+1$ anwenden läßt, so folgt aus diesem Satz, daß im Fall $2|n$ jeder 2-stufige L. B. linear von den 0-stufigen L. B. über K abhängt. (Obige Definitionen und die folgenden behalten den Sinn teils auch für beliebige Determinanten.)

Nun erzeugen die sämtlichen L. B. einen endlichen Modul $\mathfrak{L}$ über K. Wegen obiger Resultate scheint uns die genauere Untersuchung dieses Moduls ein interessantes Problem zu sein. So hätte man vor allem den Rang und eine unabhängige Basis von $\mathfrak{L}$ zu bestimmen. Ferner sollte man (über Satz 2 hinaus) nach weiteren linearen Abhängigkeiten in $\mathfrak{L}$ suchen, die auch von Nutzen sein können.

Hierfür bemerken wir, daß selbst der Laplacesche Entwicklungssatz der Determinantentheorie ein Beispiel für solche Abhängigkeiten ist, da dieser Satz besagt, daß der n-stufige L. B. für jedes ganze m $\left(0 \leqq m \leqq \frac{n}{2}\right)$ linear von den $n-2m$-stufigen L. B. abhängt (mit Koeffizienten 0 und 1). Wir wissen nicht, ob jede lineare Abhängigkeit in $\mathfrak{L}$ eine „Folgerelation" aus den durch den Laplaceschen Satz angegebenen Abhängigkeiten ist.

Aus Satz 2 folgt leicht das Korollar: *Je nachdem die Ordnung n der Determinante D gerade oder ungerade ist, bilden die* 0- *bzw.* 1-*stufigen L. B. eine Basis von $\mathfrak{L}$ über K* (die aber offenbar nicht unabhängig ist). Betrachten wir nämlich einen L. B. von D, den wir als $d\Delta$ schreiben können, wobei wir annehmen dürfen, daß d ein Hauptminor und Δ die Adjungierte von d ist. Es genügt den Fall zu betrachten, daß die Ordnung von d mindestens um 2 größer als die von Δ ist. Diese Ordnungen bezeichnen wir mit $k+1$ bzw. l $(0 \leqq l < k)$. Wir ergänzen D durch Hinzufügung von $k-l-1$ letzten Zeilen und Spalten zu einer Determinante D' von der Ordnung $2k$ so, daß die neuen Elemente in der Hauptdiagonale gleich 1 und sonst gleich 0 sind. Es ist klar, daß die 0-stufigen L. B. von D' teils gleich 0, teils (bis auf das Vorzeichen) lauter L. B. von D sind von einer Stufe $\leqq k-l-1$. Wendet man also Satz 2 auf D' (statt D) an, so folgt, daß $d\Delta$ linear von den letztgenannten L. B. von D über K abhängt. Da $d\Delta$ von der Stufe $k-l+1$ ist, so haben wir als Resultat gewonnen, daß jeder s-stufige L. B. von D $(s \geqq 2)$ linear von denjenigen L. B. von D über K abhängt, die von einer Stufe $< s$ sind. Hieraus folgt die obige Behauptung durch Induktion.

Wir vermuten, daß allgemeiner bei beliebigen ganzen Zahlen r, s $(0 \leqq r < s \leqq n;\ r \equiv s \pmod 2)$ jeder s-stufige L. B. linear von den r-stufigen L. B. über K abhängt. (Die jetzt bewiesene Aussage, selbst Satz 2 und der Satz von LAPLACE bilden je ein Beispiel für solche Abhängigkeiten.) Durch die Bestätigung dieser Vermutung, genauer gesagt durch eine entsprechende Verallgemeinerung von Satz 2 würden dann als Verallgemeinerung von Satz 1 Beziehungen von der Form

$$c\,\mathfrak{D}_k\mathfrak{D}_l \subseteq \mathfrak{D}_{k+m}\mathfrak{D}_{l-m} \qquad \left(k < l,\ 0 < m \leqq \frac{l-k}{2}\right)$$

entstehen mit einer von k, l, m abhängigen natürlichen Zahl c.

(Eingegangen am 22. September 1952.)

О ДЕЛИТЕЛЯХ ОПРЕДЕЛИТЕЛЕЙ

Л. РЭДЭИ (Сегед)

(Резюме)

Под k-ом делителем определителей матрицы над коммутативным кольцом R с единцей следует понимать идеал $\mathfrak{D}_k (k \geqq 1)$, порожденный минорами порядка k. Если кольцо R евклидово, то из теоремы элементарных делителей вытекает соотношение делимости

$$\mathfrak{D}_k^2 | \mathfrak{D}_{k-1} \mathfrak{D}_{k+1} \qquad (k \geqq 1);$$

здесь $\mathfrak{D}_0 = 1$. По Штейницу подобное имеет место и в том более общем случае, когда R есть дедекиндовое кольцо, т. е. коммутативное кольцо с главными идеалами, в котором идеалы $a \neq 0$ допускают однозначное разложение в произведение простых идеалов. В общем случае имеет место соотношение

$$\mathfrak{D}_k^2 | \bar{k}\, \mathfrak{D}_{k-1} \mathfrak{D}_{k+1} \qquad (k \geqq 1),$$

где $\bar{k}$ есть наименьшее общее кратное чисел $1, \ldots, k$.

KURZER BEWEIS DER WARINGSCHEN FORMEL

Von

L. RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

Die Formel von WARING lautet so:

$$(\mathrm{A}_k)\qquad p_k = \sum_{i=1}^{k}{}^{(k)} \frac{(-1)^i k}{i}(a_1+\cdots+a_n)^i \qquad (k=1,2,\ldots),$$

wobei $a_k = (-1)^k s_k$ ist, s_k, p_k die k-te elementarsymmetrische Form bzw. die k-te Potenzsumme der Unbestimmten $x_1,\ldots,x_n$ bezeichnet, endlich bedeutet „(k)“, daß aus der Summe nur die Glieder vom Gewicht k zu behalten sind.

Die bekannten rein algebraischen Beweise sind ziemlich umständlich.[1] Wir geben hier einen sehr leichten algebraischen Beweis an, der so ermöglicht wird, daß wir gleichzeitig auch

$$(\mathrm{B}_k)\qquad p_k = -\sum_{i=0}^{k-1}{}^{(k)} \sum_{j=1}^{n} (-1)^i j a_j (a_1+\cdots+a_n)^i \qquad (k=1,2,\ldots)$$

beweisen.

Zum Beweis werden wir die Formel

$$(1)\qquad \sum_{r=1}^{n} \frac{\partial a_j}{\partial x_r} = -(n-j+1)a_{j-1} \qquad (j=1,\ldots,n;\ a_0=1)$$

benötigen. Da die linke Seite offenbar gleich ca_{j-1} ist mit einer ganzen Zahl c, so genügt es zur Bestätigung von (1) zu zeigen, daß (1) nach der Ersetzung $x_1=\cdots=x_n=1$ richtig ist. Diese Ersetzung führt (1) in die Gleichung

$$n(-1)^j\binom{n-1}{j-1} = (-1)^j(n-j+1)\binom{n}{j-1}$$

über. Diese ist richtig, folglich ist auch (1) richtig.

Als weitere Vorbereitung beweisen wir noch den folgenden

HILFSSATZ. *Gilt für ein ganzzahliges Polynom*

$$(2)\qquad \sum_{r=1}^{n} \frac{\partial f(a_1,\ldots,a_n)}{\partial x_r} = 0,$$

so ist $f(z_1,\ldots,z_n)$ *konstant.*

[1] S. z. B.: PERRON, *Algebra*, Bd. I. (Berlin, 1951), 3. Aufl. S. 151—154. Für einen sehr kurzen analytischen Beweis mit Hilfe unendlicher Reihen s. CESÀRO—KOWALEWSKI, *Elementares Lehrbuch der algebraischen Analysis und der Infinitesimalrechnung* (Leipzig, 1904), S. 398—399.

Wenn das nämlich falsch ist, so betrachten wir das in der lexikographischen Anordnung erste nichtverschwindende und nichtkonstante Glied

$$c a_1^{i_1} \dots a_n^{i_n}$$

von $f(a_1, \dots, a_n)$, wobei $c (\neq 0)$ eine ganze Zahl ist. Wir dürfen

$$i_1 = \dots = i_{j-1} = 0, \quad i_j > 0 \qquad (1 \leqq j \leqq n)$$

annehmen. Wegen (1) enthält dann die linke Seite von (2) das Glied

$$-c(n-j+1) i_j a_{j-1} a_j^{i_j - 1} a_{j+1}^{i_{j+1}} \dots a_n^{i_n} (\neq 0),$$

und zwar geht dieses allen übrigen Gliedern lexikographisch voran, kann also nicht herausfallen. Dies widerspricht der Gleichung (2), woraus der Hilfssatz folgt.

Nunmehr beweisen wir (A_k) und (B_k). Insbesondere lautet (A_1) als $p_1 = -a_1$. Da dies richtig ist, so genügt es, wenn wir die folgenden zwei Behauptungen beweisen: (bei festem k) *aus* (A_k) *folgt* (B_k), *aus* (B_k) *folgt* (A_{k+1}).

Zum Beweis setzen wir

$$S = a_1 + \dots + a_n. \tag{3}$$

Erstens nehmen wir (A_k) für ein festes k an. Wird in (A_k) x_r durch $z x_r$ ersetzt ($r = 1, \dots, n$) mit einer neuen Unbestimmten z, so gehen p_k, a_r in $z^k p_k, z^r a_r$ über. Differenziert man dann nach z und setzt $z = 1$ ein, so entsteht

$$k p_k = \sum_{i=1}^{k}{}^{(k)} (-1)^i k (a_1 + 2a_2 + \dots + n a_n) S^{i-1}.$$

Nach Dividieren durch k stimmt dies mit (B_k) überein.

Zweitens nehmen wir (B_k) für ein festes k an. Man darf in (B_k) die beiden Summen auch auf $i = k$, $j = 0$ erstrecken, weil der Summand für $i = k$ keine Gleider vom Gewicht k enthält bzw. für $j = 0$ verschwindet. Ferner gilt nach (3)

$$\sum_{i=0}^{k}{}^{(k)} \sum_{j=0}^{n} (-1)^i a_j S^i = \sum_{i=0}^{k}{}^{(k)} (-1)^i (S^i + S^{i+1}) = 0, \tag{4}$$

weil das Polynom

$$\sum_{i=0}^{k} (-1)^i (S^i + S^{i+1}) = 1 + (-1)^{k+1} S^{k+1}$$

keine Glieder vom Gewicht k enthält. Deshalb folgt aus (B_k) nach Addieren der n-fachen linken Seite von (4):

$$p_k = \sum_{i=0}^{k}{}^{(k)} \sum_{j=0}^{n} (-1)^i (n-j) a_j S^i.$$

Es genügt hier für $j = 0, \dots, n-1$ zu summieren, weil der Summand für $j=n$ verschwindet. Ersetze man noch im Summand j durch $j-1$:

$$(5)\qquad p_k = \sum_{i=0}^{k}{}^{(k)} \sum_{j=1}^{n} (-1)^i (n-j+1) a_{j-1} S^i.$$

Die innere Summe ist nach (1) und (3) gleich

$$-\sum_{j=1}^{n}\sum_{r=1}^{n}(-1)^i S^i \frac{\partial a_j}{\partial x_r} = -\sum_{r=1}^{n}(-1)^i S^i \frac{\partial S}{\partial x_r} = \sum_{r=1}^{n}(-1)^{i+1}\frac{1}{i+1}\frac{\partial S^{i+1}}{\partial x_r}.$$

Somit folgt aus (5) (bei Ersetzung von i durch $i-1$):

$$p_k = \sum_{i=1}^{k+1}{}^{(k)} \sum_{r=1}^{n} (-1)^i \frac{1}{i}\frac{\partial S^i}{\partial x_r}.$$

Nach Umordnung haben wir

$$p_k = \sum_{r=1}^{n} \frac{\partial}{\partial x_r}\left(\sum_{i=1}^{k+1}{}^{(k+1)} (-1)^i \frac{1}{i} S^i\right),$$

wobei man offenbar $(k+1)$ für (k) schreiben mußte. Da für die linke Seite

$$\sum_{r=1}^{n} \frac{\partial}{\partial x_r} \frac{p_{k+1}}{k+1}$$

eingesetzt werden darf und p_{k+1} nach dem Fundamentalsatz der symmetrischen Polynome jedenfalls ein Polynom von $a_1, \dots, a_n$ ist, so folgt aus dem Hilfssatz sofort, daß die Differenz

$$\frac{p_{k+1}}{k+1} - \sum_{i=1}^{k+1}{}^{(k+1)} (-1)^i \frac{1}{i} S^i$$

eine Konstante sein muß. Sie kann aber nur aus Gliedern vom Gewicht $k+1$ bestehen, somit muß sie verschwinden. Das bedeutet eben die Richtigkeit von (A_{k+1}). Wir haben beide Formeln (A_k), (B_k) bewiesen.

(Eingegangen am 22. September 1952.)

КРАТКОЕ ДОКАЗАТЕЛЬСТВО ФОРМУЛЫ ВАРИНГА

Л. РЭДЭИ (Сегед)

(Резюме)

Автор дает краткое, чисто алгебраическое доказательство известной формулы Варинга.

EIN BEWEISANSATZ FÜR DIE ISOPERIMETRISCHE EIGENSCHAFT DES IKOSAEDERS

Von

L. FEJES TÓTH (Veszprém)

(*Vorgelegt von* G. HAJÓS)

In einer Abhandlung vom Jahr 1841 spricht STEINER[1] eine bemerkenswerte Vermutung aus: „Dass das regelmässige Dodekaëder und Ikosaëder, jedes unter allen Körpern, die mit ihm von gleicher Gattung sind, bei gleicher Oberfläche das Maximum des Inhaltes repräsentiert, ist nicht zu bezweifeln." Unter Polyedern von gleicher Gattung werden hierbei topologisch isomorphe Polyeder verstanden.

Im Zusammenhang mit dieser Vermutung, die STEINER auch für die übrigen regulären Körper ausgesprochen bzw. bewiesen hat, schreibt STEINITZ[2] im Jahr 1927: „Für einen Beweis dieser Annahme ist auch nicht einmal ein schwacher Ansatz vorhanden, und da nach den Erfahrungen auf diesem Gebiete größte Zurückhaltung im Aussprechen von Vermutungen zu empfehlen ist, müssen wir diese Frage, insbesondere soweit sie Dodekaeder und Ikosaeder betrifft, als noch gänzlich ungeklärt bezeichnen."

Heute steht jedoch die Steinersche Vermutung nur noch für das Ikosaeder nicht fest und die in der letzteren Zeit entdeckten verschiedenartigen Extremaleigenschaften der regulären Körper[3] scheinen die Vermutung — trotz der Bemerkung von STEINITZ — auch für das Ikosaeder zu unterstützen.

Der Zweck der vorliegenden Arbeit ist sehr bescheiden: einen „schwachen Ansatz" zum Beweis der Steinerschen Vermutung für das Ikosaeder zu geben. Obwohl ein endgültiges Ergebnis in dieser Hinsicht noch nicht erzielt wurde, scheint uns die Methode, mit der wir dieses Problem anzugreifen versuchen, einer Publikation Wert zu sein.

[1] J. STEINER, Ueber Maximum und Minimum bei den Figuren in der Ebene, auf der Kugelfläche und im Raume überhaupt, *C. R. Acad. Sci. Paris*, **12** (1841), S. 479 = *Ges. Werke II*, 177—308, S. 295.

[2] E. STEINITZ, Über isoperimetrische Probleme bei konvexen Polyedern, I, *J. reine angew. Math.*, **158** (1927), S. 129—153.

[3] Eine zusammenfassende Darstellung dieses Problemenkreises befindet sich im Werk L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum* (Berlin, im Druck).

Im § 1 folgt ein kurzer historischer Überblick über die im Zusammenhang mit der Steinerschen Vermutung bisher erreichten Ergebnisse. Dann geben wir im § 2 eine Bedingung an, der ein jedes beste Polyeder mit vorgegebener Eckenzahl unterworfen ist. Unser Hauptergebnis, das im § 3 genauer formuliert und bewiesen wird, lautet etwa folgendermaßen: besitzen die regulären 5-seitigen Pyramiden eine gewisse Minimaleigenschaft, so ist die Steinersche Vermutung richtig. Dabei scheint die die Pyramide betreffende Minimalaufgabe begreiflicher zu sein, als das ursprüngliche Problem bezüglich des Ikosaeders.

§ 1. STEINER selbst konnte die obige Vermutung, abgesehen von dem schon durch LHUILIER erledigten einfachen Fall des Tetraeders, nur noch für das Oktaeder beweisen. Sein Beweis, der auf der von ihm herrührenden berühmten Symmetrisierung beruht, läßt sich nicht auf den Fall der anderen regulären Körper übertragen.

STEINITZ befaßt sich u. a. auch mit einer anderen Vermutung von STEINER, nach der unter den Polyedern vom Typus eines n-seitigen Prismas, das einer Kugel umbeschriebene reguläre Prisma das beste Polyeder ist. Diese Vermutung erweist sich für $n=3$ als richtig, wurde aber von STEINITZ für $n \geqq 8$ wiederlegt. Ob die Vermutung für $n=4$, 5, 6 und 7 richtig ist oder nicht, konnte STEINITZ nicht entscheiden. Da diese Vermutung für $n=4$ mit der zuerst erwähnten Vermutung von STEINER bezüglich des Würfels übereinstimmt, sehen wir, daß STEINITZ die Steinersche Vermutung nicht einmal für den Würfel beweisen konnte.

Im Jahr 1936 hat GOLDBERG[4] die für ein jedes konvexes n-Flach von der Oberfläche F und vom Volumen V gültige Ungleichung

$$(1,1) \qquad \frac{F^3}{V^2} \geqq 54(n-2) \operatorname{tg} \omega_n (4 \sin^2 \omega_n - 1); \qquad \omega_n = \frac{n}{n-2} \frac{\pi}{6}$$

ausgesprochen, in der Gleichheit nur für die regulären Dreikantspolyeder zutreffen sollte. Ist dies richtig, so folgt, daß das reguläre Hexaeder und Dodekaeder nicht nur unter den topologisch isomorphen Polyedern, sondern sogar unter allen konvexen Polyedern mit 6 bzw. 12 Flächen die besten sind.

GOLDBERG macht zur Unterstützung der Ungleichung (1, 1) von der Konvexität einer ziemlich verwickelten Funktion von zwei Veränderlichen Gebrauch. Da aber diese Konvexität GOLDBERG nicht nachweisen konnte, muß sein Beweis als mangelhaft angesehen werden.

In einer im Jahr 1943 abgefaßten Arbeit[5] wurde die Ungleichung (1, 1) vom Verf. durch gleiche Gedanken unterstützt neu erhalten. Dort wurde die

[4] M. GOLDBERG, The isoperimetric problem for polyhedra, *Tôhoku Math. J.*, **40** (1935), S. 226—236.

[5] L. FEJES, Über einige Extremaleigenschaften der regulären Polyeder und des gleichseitigen Dreiecksgitters, *Annali della Scoula Norm. Sup. Pisa* (2), **13** (1948), S. 51—58.

erwähnte Schwierigkeit deutlich hervorgehoben und die Konvexität der genannten Funktion durch ein Diagramm erleuchtet. Später hat Verf. zwei Beweise für (1, 1) gegeben.[6]

Vermutlich gilt für jedes konvexe Polyeder mit n Ecken die mit (1, 1) analoge Ungleichung

$$(1,2)\qquad \frac{F^3}{V^2} \geqq \frac{27\sqrt{3}}{2}(n-2)(3\,\mathrm{tg}^2\,\omega_n - 1); \qquad \omega_n = \frac{n}{n-2}\,\frac{\pi}{6},$$

wobei Gleichheit nur für die regulären Dreieckspolyeder bestehen sollte. Daraus würde folgen, daß das reguläre Oktaeder und Ikosaeder nicht nur unter den isomorphen Polyedern, sondern sogar unter allen konvexen Polyedern mit 6 bzw. 12 Ecken die besten sind.

Es sei jedoch bemerkt, daß unter den Polyedern mit 8 oder 20 Flächen nicht das Oktaeder bzw. Ikosaeder die besten Polyeder sind. Ebenso erweist sich unter den Polyedern mit 8 bzw. 20 Ecken weder der Würfel noch das Dodekaeder als extremal. Diese letzte Behauptung legt STEINITZ durch die einfache Bemerkung klar, daß das Oktaeder besser ist als der Würfel, und das Ikosaeder besser als das Dodekaeder.

Das beste konvexe Polyeder von vorgegebener Flächenzahl n ist nach einem berühmten Satz von L. LINDELÖF und H. MINKOWSKI einer Kugel umbeschrieben. Für ein solches Polyeder ist aber $V = \frac{r}{3} F$, wobei r den Kugelradius bedeutet. Folglich ist $\frac{F^3}{V^2} = \frac{27}{r^3} V = \frac{9}{r^2} F$, sodaß es sich um die Bestimmung des einer Kugel umbeschriebenen n-Flachs von kleinstmöglichem Volumen oder Oberfläche handelt. Die Beweise von (1, 1) verfolgen alle diesen Umweg. Bei Polyedern von vorgegebener Eckenzahl steht uns aber ein dem Lindelöf—Minkowskischen Satz entsprechender einfacher Satz nicht zur Verfügung. Folglich müssen bei Polyedern mit vorgegebener Eckenzahl ganz andere Methoden zur Hilfe gerufen werden als bei Polyedern mit fester Flächenzahl.

Bedeutet e, k und f die Ecken-, Kanten- und Flächenzahl eines einer Kugel umbeschriebenen Polyeders, so gilt[7]

$$(1,3)\qquad \frac{F^3}{V^2} \geqq 9k \sin\frac{\pi f}{k}\left(\mathrm{tg}^2\frac{\pi f}{2k}\,\mathrm{tg}^2\frac{\pi e}{2k} - 1\right).$$

Gleichheit gilt hierbei für und nur für die regulären Körper. Es läßt sich vermuten, daß diese Ungleichung ihre Gültigkeit für ein jedes konvexe Polyeder

[6] L. FEJES TÓTH, The isepiphan problem for n-hedra, *Amer. J Math.*, **70** (1948), S. 174—180; Extremum properties of the regular polyhedra, *Canadian J. Math.*, **2** (1950), S. 22—31.

[7] S. die in [6] angeführte zweite Arbeit.

behält. Da ein Polyeder mit n Flächen stets als ein Dreikantspolyeder, d. h. ein Polyeder mit $k=3n-6$ Kanten und $e=2n-4$ Ecken, und jedes Polyeder mit n Ecken als ein Dreieckspolyeder, d. h. als ein Polyeder mit $k=3n-6$ Kanten und $f=2n-4$ Flächen aufgefaßt werden kann, würde — bei Richtigkeit der oben ausgesprochenen Vermutung — (1, 3) die Ungleichungen (1, 1) und (1, 2) als Sonderfälle enthalten.

§ 2. Für die isoperimetrische Eigenschaft der regulären Vielecke sind verschiedene Beweise bekannt. Wir geben hier einen Beweis, der zur Übertragung auf Polyeder von vorgegebener Eckenzahl geeignet zu sein scheint.

Es läßt sich mit Hilfe des Weierstraß'schen Satzes leicht zeigen, daß es unter den konvexen n-Ecken ein nicht ausgeartetes n-Eck gibt, für das der Quotient L^2/T den kleinstmöglichen Wert erreicht, wobei L den Umfang und T den Inhalt des n-Eckes bedeutet. Wir verschieben einen Eckpunkt E des n-Eckes bei Festhaltung der übrigen Ecken mit dem infinitesimalen Vektor $d\mathbf{r}$. Bezeichnen wir die dadurch bewirkte Inhalts- und Umfangsveränderung mit dT bzw. dL, so muß im Falle eines extremalen n-Eckes

$$d\frac{L^2}{T}=2\frac{L}{T}dL-\frac{L^2}{T^2}dT=\frac{L}{T^2}(2TdL-LdT)=0$$

sein. Wir betrachten die von E nach den benachbarten Ecken A und B hinweisende Einheitsvektoren $\mathbf{e}_A$ und $\mathbf{e}_B$, sowie die auf die Seiten EA und EB senkrecht nach außen gerichteten Vektoren $\mathbf{v}_A$ und $\mathbf{v}_B$ vom Betrag $|\mathbf{v}_A|=\overline{EA}$ bzw. $|\mathbf{v}_B|=\overline{EB}$. Setzen wir $\mathbf{e}=\mathbf{e}_A+\mathbf{e}_B$ und $\mathbf{v}=\mathbf{v}_A+\mathbf{v}_B$, so gilt, wie leicht einzusehen ist, $dT=\frac{1}{2}\mathbf{v}d\mathbf{r}$ und $dL=\mathbf{e}d\mathbf{r}$. Folglich haben wir

$$\left(2T\mathbf{e}-\frac{1}{2}L\mathbf{v}\right)d\mathbf{r}=0.$$

Das kann aber für einen beliebigen Vektor $d\mathbf{r}$ nur so bestehen, wenn

$$4T\mathbf{e}=L\mathbf{v} \tag{2, 1}$$

ausfällt. Diese Bedingung muß in jedem Eckpunkt des extremalen n-Eckes erfüllt sein.

Die Vektoren $\mathbf{e}$ und $\mathbf{v}$ können nur so entgegengesetzte Richtung haben, wenn $\overline{AE}=\overline{EB}$ ist. Folglich müssen alle Seiten eines besten n-Eckes gleichlang sein. Da ferner der Quotient $\frac{|\mathbf{e}|}{|\mathbf{v}|}$ in jedem Eckpunkt denselben Wert, nämlich $\frac{L}{4T}$, annimmt, sind auch die Winkel miteinander gleich. Das extremale n-Eck wird daher durch die Bedingung (2, 1) bis auf eine Ähnlichkeit eindeutig bestimmt: es erweist sich als regulär. Damit ist der Beweis beendet.

Die Existenz eines besten konvexen Polyeders mit n Ecken läßt sich ebenfalls leicht auf den Weierstraß'schen Satz zurückführen. Es läßt sich ferner zeigen, daß es bei jedem Wert von n ein Dreieckspolyeder ist,[8] das einer unten zu besprechenden Bedingung unterworfen ist.

Verändert sich die Oberfläche F und das Volumen V eines konvexen Polyeders bei einer Verrückung eines Eckpunktes E um den infinitesimalen Vektor $d\mathbf{r}$ mit dF bzw. dV, so ist

$$d\frac{F^3}{V^2} = 3\frac{F^2}{V^2}dF - 2\frac{F^3}{V^3}dV.$$

Wir setzen voraus, daß in E nur Dreiecksflächen $f_i = EE_iE_{i+1}$ ($i = 1, \ldots, \nu$; $E_{\nu+1} = E_1$) zusammenkommen. Wir betrachten den von E aus senkrecht auf die Kante E_iE_{i+1} gerichteten Vektor $\mathbf{s}_i$ vom Betrag $|\mathbf{s}_i| = \overline{E_iE_{i+1}}$, sowie den auf die Fläche f_i senkrecht nach außen gerichteten Vektor $\mathbf{f}_i$ vom Betrag $|\mathbf{f}_i| = f_i$. Wir nennen $\mathbf{s} = \mathbf{s}_1 + \cdots + \mathbf{s}_\nu$ und $\mathbf{f} = \mathbf{f}_1 + \cdots + \mathbf{f}_\nu$ den zur Ecke E gehörigen *Seiten-* bzw. *Flächenvektor*. Es ist leicht einzusehen, daß $dF = \frac{1}{2}\mathbf{s}d\mathbf{r}$ und $dV = \frac{1}{3}\mathbf{f}d\mathbf{r}$ ist. Folglich gilt

$$d\frac{F^3}{V^2} = \frac{3F^2}{2V^2}\mathbf{s}d\mathbf{r} - \frac{2F^3}{3V^3}\mathbf{f}d\mathbf{r} = \frac{F^2}{6V^3}(9V\mathbf{s} - 4F\mathbf{f})d\mathbf{r}.$$

Im Falle eines extremalen Polyeders muß dieses Differential verschwinden. Das kann aber für eine beliebige Richtung von $d\mathbf{r}$ nur so zutreffen, wenn

$$(2, 2) \qquad 9V\mathbf{s} = 4F\mathbf{f}$$

ausfällt. Diese Bedingung muß in jeder Ecke des extremalen Polyeders erfüllt sein.

Von der Tatsache, daß das beste Polyeder nur Dreiecksflächen haben kann, machen wir im folgenden keinen Gebrauch und verzichten deshalb auf seinen Beweis. Jedenfalls zeigen die obigen Überlegungen, daß wenn es unter den Polyedern vom Typus eines Dreieckspolyeders ein bestes gibt, dieses der Bedingung (2, 2) Genüge leistet.

Die Bedingung (2, 2) entspricht der Bedingung (2, 1). Da aber die Gestalt des Polyeders durch (2, 2) im allgemeinen nicht eindeutig bestimmt wird, scheiden sich bei diesem Punkt die in der Ebene und im Raum zu verfolgenden Vorgänge.

§ 3. Wir betrachten an einer Einheitskugel vom Mittelpunkt E ein sphärisches ν-Eck $P_1 \ldots P_\nu$ vom Umfang U und nehmen an jeder Halbgerade EP_i einen beliebigen Punkt E_i an. Wir nennen das von den Dreiecken $f_i = EE_iE_{i+1}$ ($i = 1, \ldots, \nu$; $E_{\nu+1} = E_1$) bestehende Flächenstück P schlechthin eine *Pyramidenfläche*. Die Größe $K = 2\pi - U$ soll die *Krümmung* von P genannt werden. Ist das ν-Eck $P_1 \ldots P_\nu$ konvex, so sprechen wir von einer

[8] Vgl. die Arbeit [5].

konvexen Pyramidenfläche. Sind die Vielecke $E_1 \ldots E_\nu$ und $P_1 \ldots P_\nu$ regulär (und folglich homothetisch), so nennen wir die Pyramidenfläche *regulär.*

Wir erklären in derselben Weise wie in § 2 den zu P gehörigen Seiten- und Flächenvektor und stellen uns folgende

M i n i m a l a u f g a b e. Wir betrachten diejenigen konvexen Pyramidenflächen von fester Seitenzahl, Oberfläche und Krümmung, für die der Flächenvektor $\mathbf{f}$ und Seitenvektor $\mathbf{s}$ entgegengesetzte Richtungen haben: $\frac{\mathbf{s}}{|\mathbf{s}|}+\frac{\mathbf{f}}{|\mathbf{f}|}=0$. Gesucht werden unter diesen Pyramidenflächen diejenigen, für die der Quotient $\frac{|\mathbf{s}|}{|\mathbf{f}|}$ den kleinstmöglichen Wert erreicht.

Wahrscheinlich sind die extremalen Pyramidenflächen regulär. Wir wollen das an zwei Sonderfällen einigermaßen plausibel machen.

Betrachten wir zunächst das vom steifen Papier hergestellte Modell einer regulären vierseitigen Pyramidenfläche! Wir nehmen an, daß das Modell Veränderungen der Flächenwinkel zwischen 0^0 und 180^0 zuläßt. Dann ist das Modell beweglich; es besitzt einen Freiheitsgrad. Inzwischen sind alle Bedingungen unserer Extremalaufgabe erfüllt. In den beiden extremal zulässigen Lagen, wobei die Flächenwinkel 0^0 und 180^0 betragen, ist $|\mathbf{f}|$ Null, $|\mathbf{s}|$ dagegen von Null verschieden. Somit wird $\frac{|\mathbf{s}|}{|\mathbf{f}|}$ in diesen Lagen unendlich. Bringen wir unser Modell von der einen extremalen Lage durch stetige Bewegung in die andere, so nimmt der Quotient $\frac{|\mathbf{s}|}{|\mathbf{f}|}$ in der regulären Lage aus Symmetriegründen offenbar einen extremalen Wert an. Es ist zu erwarten, und es läßt sich auch leicht beweisen, daß es sich um ein absolutes Minimum handelt.

Als zweites Beispiel betrachten wir wiederum eine vierseitige Pyramidenfläche, bei der das sphäriche Viereck $P_1P_2P_3P_4$ gleichwinklig und $EE_1=EE_3$, $EE_2=EE_4$ ist. Wir halten zunächst $P_1P_2P_3P_4$ fest und ersetzen die Punkte E_i durch auf den Halbgeraden EP_i liegende Punkte E_i', sodaß

$$\overline{EE_1'}^2=\cdots=\overline{EE_4'}^2=EE_1\cdot EE_2=\cdots=EE_4\cdot EE_1$$

ausfällt. Dann bleiben die Inhalte der Dreiecksflächen und folglich auch der Flächenvektor $\mathbf{f}$ unverändert. Dagegen nimmt $|\mathbf{s}|$ — wie leicht einzusehen ist — ab. Wir können daher annehmen, daß $E_1E_2E_3E_4$ ein ebenes Rechteck ist.

Lassen wir jetzt das Viereck $P_1P_2P_3P_4$ unter der Bedingung konstanten Umfanges variieren, so wird $\frac{|\mathbf{s}|}{|\mathbf{f}|}$ in den beiden extremalen Lagen, wobei $E_1E_2E_3E_4$ in die Strecke E_1E_2 bzw. E_2E_3 ausartet, unendlich und erreicht sein Minimum, wenn $E_1E_2E_3E_4$ ein Quadrat ist.

Wir wollen nun zeigen: *Sind in der obigen Minimalaufgabe die extremalen 5-seitigen Pyramidenflächen regulär, und gibt es unter den oberflächen-*

gleichen konvexen Polyedern vom Typus des regulären Ikosaeders eines von größtmöglichem Volumen, so ist es das reguläre Ikosaeder.

Wir führen zunächst einen Begriff ein. Wir nennen eine in einem konvexen Bereich B der (x, y)-Ebene erklärte Funktion $f(x, y)$ aus naheliegenden Gründen *semikonvex,* wenn für je n Punkte $(x_1, y_1), \ldots, (x_n, y_n)$ von B

$$(3,1) \qquad \max_{1 \leqq i \leqq n} f(x_i, y_i) \geqq f\left(\frac{x_1 + \cdots + x_n}{n}, \frac{y_1 + \cdots + y_n}{n}\right)$$

ausfällt.[9]

Ist $f(x, y)$ im ganzen Bereich B positiv und gibt es eine reelle Zahl α_0, so daß für jedes $\alpha > \alpha_0$, f^α konvex ist, so ist f in B semikonvex. In diesem Fall gilt nämlich nach der Jensenschen Ungleichung

$$\left(\frac{f^\alpha(x_1, y_1) + \cdots + f^\alpha(x_n, y_n)}{n}\right)^{\frac{1}{\alpha}} \geqq f\left(\frac{x_1 + \cdots + x_n}{n}, \frac{y_1 + \cdots + y_n}{n}\right),$$

woraus sich mit $\alpha \to \infty$ eben die Ungleichung (3, 1) ergibt.

Setzen wir noch voraus, daß f stetige zweite partielle Ableitungen f_{xx}, f_{xy}, f_{yy} besitzt, so gilt dies auch für die Funktion $F(x, y) = f^\alpha(x, y)$. Bekanntlich ist F konvex, wenn in B

$$D = F_{xx} F_{yy} - F_{xy}^2 > 0$$

und außerdem $F_{xx} > 0$ oder $F_{yy} > 0$ ist. Es ist

$$\frac{D}{\alpha^2 f^{2\alpha-3}} = (\alpha - 1)(f_x^2 f_{yy} + f_y^2 f_{xx} - 2 f_x f_y f_{xy}) + f(f_{xx} f_{yy} - f_{xy}^2).$$

Wir erhalten hieraus eine hinreichende Bedingung der Semikonvexität: Es sei $f(x, y)$ eine im abgeschlossenen konvexen Bereich B erklärte positive Funktion, die in B stetige zweite partielle Ableitungen f_{xx}, f_{xy}, f_{yy} besitzt. Gilt außer der Ungleichung

$$(3,2) \qquad f_x^2 f_{yy} + f_y^2 f_{xx} - 2 f_x f_y f_{xy} > 0$$

noch $f_{xx} > 0$ oder $f_{yy} > 0$, so ist $f(x, y)$ in B semikonvex.

Mit Rücksicht auf die Stetigkeit der Ableitungen besitzt nämlich der Ausdruck an der linken Seite von (3, 2), sowie $f(f_{xx} f_{yy} - f_{xy}^2)$ je ein Minimum M bzw. m in B. Da wegen der Voraussetzung (3, 2) $M > 0$ ist, gilt $D > 0$ für $\alpha > 1 - \dfrac{m}{M}$. Ferner folgt aus $f_{xx} > 0$ wegen

$$\frac{1}{\alpha} F_{xx} = (\alpha - 1) f^{\alpha-2} f_x^2 + f^{\alpha-1} f_{xx}$$

und der vorausgesetzten Positivität von f zugleich $F_{xx} > 0$ für jedes $\alpha \geqq 1$.

[9] Es genügt das Bestehen der Ungleichung (3, 1) für je zwei Punkte vorauszusetzen. Für eine stetige Funktion bedeutet diese Ungleichung, daß kein zu der z-Achse paralleler Schnitt der Fläche $z = f(x, y)$ einen Höhepunkt im Inneren von B besitzt.

Wir wenden nun unsere Bedingung auf die Funktion

$$Q(T,K)=\frac{4\nu}{T}\left(\operatorname{cotg}\frac{2\pi-K}{2\nu}\operatorname{tg}^2\frac{\pi}{\nu}-\operatorname{tg}\frac{2\pi-K}{2\nu}\right)=\frac{\omega(K)}{T};$$

$$0<T,\quad 0<K<2\pi$$

an, die den Wert des Quotienten $\frac{\mathbf{s}^2}{\mathbf{f}^2}$ für eine reguläre ν-seitige Pyramidenfläche der Oberfläche T und Krümmung K darstellt. Mit Rücksicht auf

$$Q_T^2 Q_{KK}+Q_K^2 Q_{TT}-2Q_T Q_K Q_{TK}=Q_T^2 Q_{KK}=$$

$$=\frac{\omega^2\omega''}{T^5}=\frac{2\omega^2}{\nu T^5}\frac{\operatorname{tg}^2\frac{\pi}{\nu}-\operatorname{tg}^4\frac{2\pi-K}{2\nu}}{\operatorname{tg}^3\frac{2\pi-K}{2\nu}\cos^2\frac{2\pi-K}{2\nu}}>0$$

ist $Q(T,K)$ im Falle $\nu\geqq 4$ gewiß in jedem abgeschlossenen konvexen Teilgebiet des Halbstreifens $0<T,\ 0<K<2\pi$ semikonvex.

Wir setzen jetzt voraus, daß es unter den konvexen Polyedern vom Typus des regulären Ikosaeders ein bestes gibt. Dieses ist dann der Bedingung (2, 2) unterworfen. Ist nun unsere Vermutung bezüglich der Minimaleigenschaft der regulären 5-seitigen Pyramidenflächen richtig, so gilt für jede Ecke des besten Polyeders

$$\frac{16F^2}{81V^2}=\frac{\mathbf{s}^2}{\mathbf{f}^2}\geqq Q(T_i,K_i),\quad \nu=5,$$

wo T_i und K_i die Oberfläche bzw. Krümmung der zur Ecke vom Index i gehörigen Pyramidenfläche bedeutet. Da aber $Q(T,K)$ für

$$0<\min T_i\leqq T\leqq\max T_i,\quad 0<\min K_i\leqq K\leqq\max K_i<2\pi$$

semikonvex ist, haben wir

$$\frac{16F^2}{81V^2}\geqq Q\left(\frac{T_1+\cdots+T_{12}}{12},\frac{K_1+\cdots+K_{12}}{12}\right)=Q\left(\frac{F}{4},\frac{\pi}{3}\right)=\frac{4}{F}\,\omega\left(\frac{\pi}{3}\right),$$

d. h.

$$\frac{F^3}{V^2}\geqq 135\sqrt{3}\left(3\operatorname{tg}^2\frac{\pi}{5}-1\right).$$

Das ist eben die isoperimetrische Ungleichung für das Ikosaeder. Der Fall der Gleichheit leuchtet ein.

Zum Schluß noch einige Bemerkungen!

Es ist möglich, daß die regulären Pyramidenflächen die vermutete Minimaleigenschaft auch unter leichter zu behandelnden, allgemeineren Bedingungen behalten. Man könnte etwa versuchen, die lästige Bedingung $\frac{\mathbf{s}}{|\mathbf{s}|}+\frac{\mathbf{f}}{|\mathbf{f}|}=0$ fallen zu lassen und anstatt dessen nach dem Minimum von $\frac{\mathbf{s}^2}{\mathbf{sf}}$ zu fragen. Übrigens kann anstatt der von uns betrachteten Krümmung auch die sogenannte totale Krümmung der Pyramidenfläche, oder irgendein anderes geeignetes Funktional in Betracht gezogen werden.

Die Funktion $Q(T, K)$, die wir jetzt mit $Q(T, K, \nu)$ bezeichnen wollen, nimmt bei festen Werten von T und K mit wachsendem ν ständig ab. Dies steht im Einklang mit der Vermutung $\frac{\mathbf{s}^2}{\mathbf{f}^2} \geqq Q(T, K, \nu)$. Da mit $\nu \to \infty$ die Funktion Q einem Grenzwert zustrebt, ist zu erwarten, daß Q eine konvexe Funktion von ν ist. Das läßt sich durch Differentiation auch leicht nachweisen. Daraus folgt aber, daß $Q(T, K, \nu)$ bei festem K eine semikonvexe Funktion von T und ν ist. Ließe sich zeigen, daß $Q(T, K, \nu)$ auch als Funktion von drei Veränderlichen semikonvex ist, und wäre die Vermutung $\frac{\mathbf{s}^2}{\mathbf{f}^2} \geqq Q(T, K, \nu)$ für beliebige Werte von $\nu \geqq 3$ richtig, so könnte man mit Hilfe der obigen Überlegungen leicht die Ungleichung (1, 2) beweisen. Dieses Verfahren hätte außer der viel größeren Tragweite noch den Vorteil, daß sich auch die Existenzfrage leicht erledigen ließe.

(Eingegangen am 12. April 1952.)

МЕТОД ДОКАЗАТЕЛЬСТВА ИЗОПЕРИМЕТРИЧЕСКОГО СВОЙСТВА ИКОЗАЭДРА

Л. ФЕЕШ ТОТ (Веспрем)

(Резюме)

Известная гипотеза Й. Штейнера, соответственно которой из равноповерхностных полиэдров топологически эквивалентных данному правильному геометрическому телу само это правильное тело имеет максимальный объем, остается пока недоказанной только для случая икозаэдра. Автор доказывает, что если правильные пятисторонные пирамиды обладают некоторым экстремальным свойством, то гипотеза Штейнера верна и для икозаэдра.

ON A MONOTONIC PROPERTY OF CERTAIN STURM—LIOUVILLE FUNCTIONS

By

E. MAKAI (Budapest)

(Presented by P. TURÁN)

1. R. G. COOKE showed[1] that if the non-negative zeros of the Bessel function $J_\nu(x)$ are $x_{\nu 1}, x_{\nu 2}, \ldots, x_{\nu k}, \ldots$ in increasing order, then

$$(1) \qquad \left| \int_{x_{\nu, k-1}}^{x_{\nu k}} J_\nu(x)dx \right| > \left| \int_{x_{\nu k}}^{x_{\nu, k+1}} J_\nu(x)dx \right| \qquad (\nu > -1).$$

This inequality was proved by making use of intricate properties of the Bessel functions. Recently E. EGERVÁRY and P. TURÁN have found a simple proof of (1) in the case $\nu = 1$. Subsequently P. TURÁN raised the question whether the whole theorem could be proved in a simpler way than the original proof of COOKE. The answer turned out to be affirmative at least partially. Assertion (1) is in the case $|\nu| > \frac{1}{2}$ an almost immediate consequence of a general theorem of STURM[2] which it is advantageous to formulate in the following form with regard to a result of SZEGŐ.[3]

Let $f(x)$ and $F(x)$ be continuous functions in $a < x \leq b$ and let there $f(x) \leq F(x)$ but $f(x) \not\equiv F(x)$. Let the functions $y(x)$ and $Y(x)$ satisfy in $a < x \leq b$ the differential equations

$$(2) \qquad y'' + f(x)y = 0 \quad \textit{and} \quad Y'' + F(x)Y = 0$$

and further the following conditions:

(a) $y(x) > 0,\ Y(x) > 0$ *in* $a < x \leq b,\ Y(b) = 0$;

(b) *the limits*

$$\lim_{x\to a+0} y(x),\ \lim_{x\to a+0} Y(x),\ \lim_{x\to a+0} y'(x),\ \lim_{x\to a+0} Y'(x),\ \lim_{x\to a+0} [y'(x)Y(x) - y(x)Y'(x)]$$

exist and have non-negative values.

[1] A monotonic property of Bessel functions, *Journal of the London Mathematical Society,* **12** (1937), pp. 180—185.

[2] Ch. STURM, Sur les équations différentielles linéaires de deuxième ordre, *Journal de Mathématiques pures et appliquées,* **1** (1836), pp. 106—186, esp. p. 179.

[3] G. SZEGŐ, Inequalities for the zeros of Legendre polynomials and related functions, *Transactions of the American Mathematical Society,* **39** (1936), pp. 1—17.

Then y/Y is an increasing function of x in the interval $a < x \leqq b$. Moreover, if we assume

(c) $$\lim_{x \to a+0} \frac{y(x)}{Y(x)} \geqq 1,$$

then we have

(3) $$y(x) > Y(x) \geqq 0 \qquad (a < x \leqq b).$$

The essential idea of the proof is well known. From (2) it follows that

$$(y'Y - yY')' = [F(x) - f(x)]yY > 0 \qquad (a < x < b)$$

and, consequently, $\varphi(x) = y'Y - yY'$ is an increasing function of x. On account of assumption (b) we have

(4) $$y'Y - yY' > 0 \qquad (a < x < b),$$

consequently $(y/Y)' > 0$, that is, y/Y increases in $a < x < b$. Comparing this with assumption (c) we see that (3) is satisfied in the open interval $\langle a, b \rangle$.

For $x = b$ the formula (3) follows immediately from the inequality

$$y'(b)Y(b) - y(b)Y'(b) > 0.$$

As $Y(b) = 0$ and $Y'(b)$ cannot be positive, this inequality can hold only if $y(b) > 0$ (and $Y'(b) \neq 0$).

We shall need only the following special case of this theorem. *Let*

$$\lim_{x \to a+0} y(x) = \lim_{x \to a+0} Y(x) = 0$$

and

$$\lim_{x \to a+0} y'(x) = \lim_{x \to a+0} Y'(x) = C > 0,$$

then in the interval $a < x \leqq b$ we have $y(x) > Y(x)$, i. e. the graph of $y(x)$ lies above that of $Y(x)$.

2. From this last remark it follows by making use of an idea of STURM that *if*

(a) *in $x_0 < x < X_0$ (where X_0 is finite or infinite) $f(x)$ decreases monotonously,*

(b) *$y(x)$ is that particular solution of the differential equation*

(5) $$y'' + f(x)y = 0$$

whose three consecutive zeros are x_1, x_2 and x_3,[4] *then*

(6) $$|y(x_2 - u)| < |y(x_2 + u)| \qquad (0 < u \leqq x_2 - x_1).$$

The geometrical interpretation of (6) is the following. Let a_1 and a_2 be the arcs of the graph of $y(x)$ lying between x_1 and x_2, resp. between x_2 and x_3. Now turn the arc a_1 by 180° around its end point having the coordinates x_2 and 0, and denote its new position by a_1'. Then the arc a_1' lies in the area bounded by the x axis and the arc a_2. More precisely, if we denote by $Y(x)$ the function the graph of which is a_1', then

(6a) $$|y(x)| > |Y(x)| \qquad (x_2 < x \leqq x_2 + (x_2 - x_1)).$$

[4] We assume of course the existence of at least three zeros in (x_0, X_0).

In proving (6a) it is convenient to put $x_2 = 0$ which we can always achieve by a linear transformation. Then the function $Y(x)$ is obviously equal to $-y(-x)$. The differential equation of $Y = Y(x)$ is

$$\frac{d^2 Y}{dx^2} + f(-x) Y = 0.$$

If we take into account that $y'(x_2) = Y'(x_2)$ and $f(x)$ decreases in (x_0, X_0) or at least

$$\min_{x_1 < x < x_2} f(x) > \max_{x_2 < x < x_3} f(x)$$

— see point (d) later in this section — further if we let

$$F(x) = f(-x) > f(x) \qquad (x_2 < x \leqq x_2 + (x_2 - x_1)),$$

then we can apply the result from the end of §1 which is equivalent to (6a) and (6).

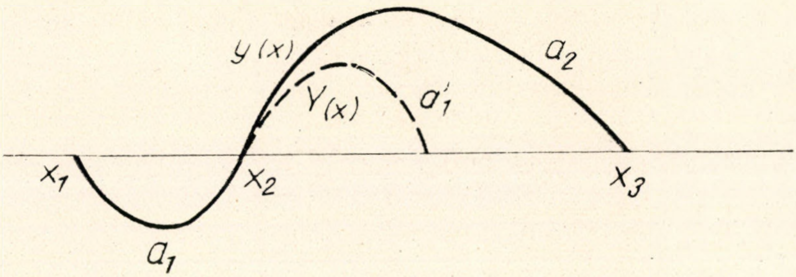


Fig. 1

The property of $y(x)$ expressed by (6) and (6a) may be completed in the following way. If there exist a smallest and a largest zero of $y(x)$ in (x_0, X_0) (say x_α resp. x_β), further x_γ is the second largest zero of $y(x)$ — the existence of x_γ being again supposed — then

$$\text{(6b)} \qquad \begin{cases} |y(x_\alpha - v)| < |y(x_\alpha + v)| & (x_\alpha - v > x_0, \ v > 0), \\ |y(x_\beta - w)| < |y(x_\beta + w)| & (x_\beta + w < X_0, \ x_\gamma < x_\beta - w < x_\beta). \end{cases}$$

The demonstration of (6b) follows exactly the same lines as that of (6).

If a function $y(x)$ satisfies requirements (6) and (6b) and the additional requirement that the zeros of $y(x)$ and $y'(x)$ alternate in (x_0, X_0), then we shall call it to be of type $T(x_0, X_0)$ regardless whether or not it satisfies a differential equation of type (5).

Now let us indicate some immediate consequences of what has been proved already.

(a) If the zeros of $y(x) \in T(x_0, X_0)$ in the interval (x_0, X_0) are $x_1, x_2, \ldots, x_n, \ldots$ in increasing order, then

$$x_2 - x_1 < x_3 - x_2 < x_4 - x_3 < \cdots$$

which is the well known theorem concerning the convexity of the zeros. (Cf. SZEGŐ, loc. cit., p. 4.)

(b) If the increasing numbers $\xi_1, \xi_2, \ldots$ are the values of x in the interval (x_0, X_0) where $|y(x)|$ has a local maximum, and the ξ's are enclosed between two consecutive zeros of $y(x)$ belonging to (x_0, X_0), then

$$|y(\xi_1)| < |y(\xi_2)| < \cdots$$

(Theorem of SONIN[5]).

(c) $$\int_{x_1}^{x_2} |y(x)|^p\, dx < \int_{x_2}^{x_3} |y(x)|^p\, dx < \cdots,$$

a special case of which is statement (a).

(d) From the way of the demonstration it is clear that the above statements hold too, if we do not assume that the function $f(x)$ decreases monotonously but only that the values of $f(x)$ are greater in $x_{i-1} < x < x_i$ than anywhere in $x_i < x < x_{i+1}$.

(e) Lastly, if $f(x)$ is monotone increasing in $x_0 < x < X_0$, the previous assertions on $y(x)$ are to be modified in an obvious way. Functions $y(x)$ satisfying the analogues of (6) and (6b) will be termed of type $T^*(x_0, X_0)$.

3. The statements of §2 can be applied immediately to the orthogonal functions of HERMITE:

$$\chi_n = \chi_n(x) = (-1)^n \cdot e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-x^2}.$$

The differential equation of these functions is

$$\frac{d^2\chi_n}{dx^2} + (2n + 1 - x^2)\chi_n = 0. \tag{7}$$

If the zeros of $\chi_n(x)$ are $x_{n1}, x_{n2}, \ldots, x_{nn}$ in decreasing order then the solutions of (7) are of type $T(x_{n,[n/2]+1}, \infty)$, and e. g.

$$\int_{x_{n1}}^{\infty} |\chi_n(x)|^p\, dx > \int_{x_{n2}}^{x_{n1}} |\chi_n(x)|^p\, dx > \cdots > \int_{x_{n,[n/2]}}^{x_{n,[n/2]-1}} |\chi_n(x)|^p\, dx > \int_{x_{n,[n/2]+1}}^{x_{n,[n/2]}} |\chi_n(x)|^p dx \tag{8}$$

$(p > 0)$. In the case of even n the last inequality holds by virtue of §2 (d).

In the case $p = 2$, the sequence (8) of inequalities was conjectured in Quantum Mechanics[6] where it has the following meaning. $\chi_n(x)$ is the n'th eigenfunction of the harmonic oscillator. If $\chi_n(x)$ is normalized then

[5] I owe I. BIHARI this remark. The present proof of SONIN's theorem requires partly less stringent restrictions on $f(x)$ than the usual proof. See G. SZEGŐ, *Orthogonal polynomials* (New York, 1939), p. 161, where it is assumed that $f(x) > 0$ if $x_0 < x < X_0$. We shall see later in §5 that by the omission of the requirement of the positivity of $f(x)$ we can state somewhat more about the extremal values of $J_\nu(x)$ than what is proved in SZEGŐ's book.

[6] Cf. G. HERZBERG, *Molekülspektren und Molekülstruktur I.* (1939), p. 58.

$\int\limits_{x_{nk}}^{x_{n,k-1}} [\chi_n(x)]^2 dx$ is the probability that the distance of the oscillating particles is between two consecutive nodes x_{nk} and $x_{n,k-1}$ of the eigenfunctions. Formula (8) shows that the farther these neighbouring nodal points are from $x=0$, the greater is the probability of finding the particle between them.

The statements of §2 hold for Hermite polynomials too defined by

$$H_n(x) = e^{\frac{x^2}{2}} \chi_n(x).$$

Here we shall only show that from the fact that $\chi_n(x)$ is of type $T(x_{n,[n/2]+1}, \infty)$ it follows that $H_n(x)$ also belongs to type $T(x_{n,[n/2]+1}, \infty)$ or, what amounts to the same, the function

$$\varphi_n(x) = e^{\frac{x^2 - x^2_{n,k}}{2}} \chi_n(x)$$

belongs to the same type.

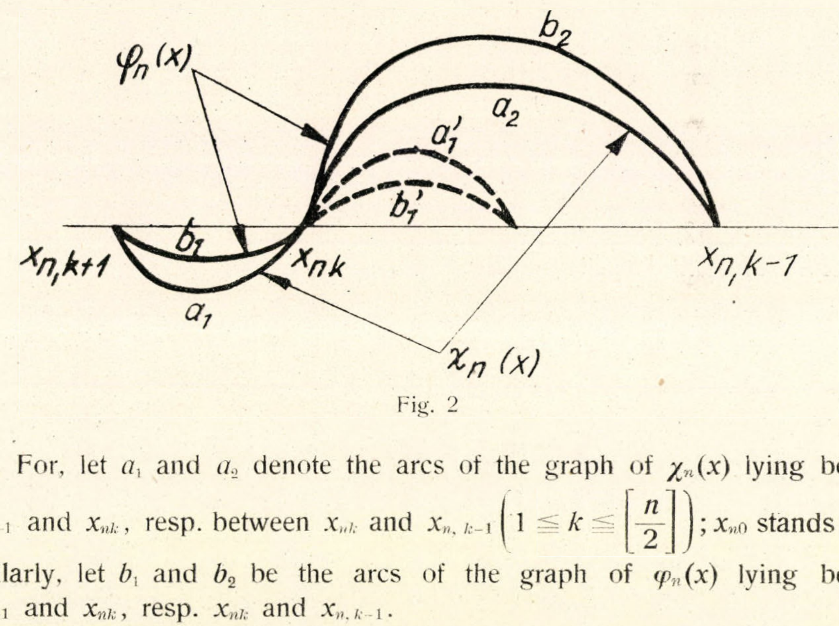


Fig. 2

For, let a_1 and a_2 denote the arcs of the graph of $\chi_n(x)$ lying between $x_{n,k+1}$ and x_{nk}, resp. between x_{nk} and $x_{n,k-1}$ $\left(1 \leq k \leq \left[\frac{n}{2}\right]\right)$; x_{n0} stands for ∞. Similarly, let b_1 and b_2 be the arcs of the graph of $\varphi_n(x)$ lying between $x_{n,k+1}$ and x_{nk}, resp. x_{nk} and $x_{n,k-1}$.

Now it is clear that

$$|\varphi_n(x)| < |\chi_n(x)| \quad \text{if} \quad x_{n,k+1} < x < x_{nk}$$

and

$$|\varphi_n(x)| > |\chi_n(x)| \quad \text{if} \quad x_{nk} < x < x_{n,k-1}.$$

Geometrically this means that the arc b_1 (resp. a_2) lies between the x axis and the arc a_1 (resp. b_2). Moreover, if we turn the arcs a_1 and b_1 around their common right end points by 180°, the new position being denoted by a_1', and b_1', then b_1' lies between a_1' and the x axis. By virtue of what has been said, a_1' lies between a_2 and the x axis and so, a fortiori, the arc b_1 lies between the arc b_2 and the x axis:

$$|\varphi_n(x_{nk}-u)| < |\varphi_n(x_{nk}+u)| \quad \text{if} \quad 0 < u \leqq x_{nk} - x_{n,k+1} \qquad \left(1 \leqq k \leqq \left[\frac{n}{2}\right]\right).$$

In other words, $\varphi_n(x)$ (and $H_n(x)$) belongs to type $T(x_{n,[n/2]+1}, \infty)$.

4. Let $x_1, x_2, \ldots, x_n$ be the zeros of the Legendre polynomial $P_n(x)$ in decreasing order and $0 < \vartheta_1 < \cdots < \vartheta_n < \pi$ the zeros of $P_n(\cos\vartheta)$. Further let $x_0 = 1, \vartheta_0 = 0$. We shall prove that

$$\int_{x_k}^{x_{k-1}} |P_n(x)|^p\,dx < \int_{x_{k+1}}^{x_k} |P_n(x)|^p\,dx \tag{9}$$

$$\left(0 \leqq p \leqq 2;\quad k = 1, 2, \ldots, \left[\frac{n}{2}\right],\quad \text{i. e.}\quad x_k > 0\right).$$

Let us consider the function $u(\vartheta) = \sqrt{\sin\vartheta}\,P_n(\cos\vartheta)$ which satisfies the differential equation

$$\frac{d^2u}{d\vartheta^2} + f(\vartheta)u = 0;\quad f(\vartheta) = \left(n+\frac{1}{2}\right)^2 + \frac{1}{\sin^2\vartheta}.$$

$f(\vartheta)$ is a decreasing function in $0 < \vartheta < \dfrac{\pi}{2}$ which, in the case of even n, is less in the interval $\langle \vartheta_{[n/2]}, \vartheta_{[n/2]+1}\rangle$ than anywhere in $\langle 0, \vartheta_{[n/2]}\rangle$. Therefore we can put by §2 (c) and (d) the symbol $<$ between the second and third member in the following sequence of inequalities:

$$\begin{gathered}\int_{\vartheta_{k-1}}^{\vartheta_k} \sin^{1-\frac{p}{2}}\vartheta\,|u(\vartheta)|^p\,d\vartheta \leqq \sin^{1-\frac{p}{2}}\vartheta_k \int_{\vartheta_{k-1}}^{\vartheta_k} |u(\vartheta)|^p\,d\vartheta < \\ < \sin^{1-\frac{p}{2}}\vartheta_k \int_{\vartheta_k}^{\vartheta_{k+1}} |u(\vartheta)|^p\,d\vartheta \leqq \int_{\vartheta_k}^{\vartheta_{k+1}} \sin^{1-\frac{p}{2}}\vartheta\,|u(\vartheta)|^p\,d\vartheta\end{gathered} \tag{9a}$$

$$\left(k = 1, 2, \ldots, \left[\frac{n}{2}\right];\quad 0 \leqq p \leqq 2\right).$$

The validity of the two symbols $\leqq$ is a trivial consequence of the mean value theorem of integrals.

On the other hand, the first and last member of the inequalities (9a) are equal to the left hand resp. right hand member of (9).

5. Now let us consider *any* solution $y = y(x)$ of BESSEL's differential equation

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right)y = 0 \qquad \left(|\nu| > \frac{1}{2}\right).$$

Let the non-negative zeros of $y(x)$ be $x_1, x_2, \ldots, x_k, \ldots$ in increasing order.

Then

$$\left|\int_{x_{k-1}}^{x_k} y(x)\,dx\right| > \left|\int_{x_k}^{x_{k+1}} y(x)\,dx\right|. \tag{10}$$

Thus, in the case $|\nu| > \frac{1}{2}$, we state somewhat more than the original assertion of COOKE. We could show by comparing $y(x)$ with

$$z(x) = \sqrt{x}\,y(x)$$

in the same way as was used in §3 that $y(x)$ is of type $T^*(0, \infty)$ (see §2 (d)) and then inequality (10) would follow.

Instead of doing so, we argue as follows. Also $z(x)$ is of the same type $T^*(0, \infty)$ as its differential equation is

$$z'' + f(x)z = 0; \quad f(x) = 1 - \frac{\nu^2 - \frac{1}{4}}{x^2},$$

hence $f'(x) > 0$ if $|\nu| > \frac{1}{2}$. Applying §2 (c) with exponent $p = 1$ to $z(x)$, we get, after dividing by $\sqrt{x_k}$ $(k > 1)$,

$$\left|\int_{x_{k-1}}^{x_k} \sqrt{\frac{x}{x_k}}\,y(x)\,dx\right| > \left|\int_{x_k}^{x_{k+1}} \sqrt{\frac{x}{x_k}}\,y(x)\,dx\right|, \tag{11}$$

the zeros of $z(x)$ being the same as those of $y(x)$. This is a sharper inequality than (10). Now in the integrand on the left of (11) we have $\sqrt{\frac{x}{x_k}} < 1$, while $\sqrt{\frac{x}{x_k}} > 1$ on the right. So the left hand (right hand) side of (11) is less (greater) than the left hand (right hand) side of (10). This completes the proof of inequality (10).

Finally if we apply §2 (b) to the special function $z = \sqrt{x}\,|J_\nu(x)|$ $\left(|\nu| > \frac{1}{2}\right)$, we see that its relative maxima form a decreasing sequence provided $x > 0$. The same is proved by SZEGÖ (loc. cit., p. 162) for the interval $\sqrt{\nu^2 - \frac{1}{4}} < x < \infty$.

(Received 16 June 1952)

ОБ ОДНОМ СВОЙСТВЕ МОНОТОННОСТИ ФУНКЦИЙ ШТУРМА—ЛИУВИЛЛЯ

Е. МАКАИ (Будапешт)

(Резюме)

В настоящей работе доказывается следующая теорема типа Штурма. Пусть $f(x)$ непрерывная функция монотонно убывающая на интервале $x_0 < x < X_0$. Пусть $y(x)$ означает одно из решений дифференциального уравнения

$$y'' + f(x)\,y = 0.$$

Предположим, что $y(x)$ имеет на сегменте $x_0 \leqq x \leqq X_0$ хотя бы три корней, три соседние из которых мы будем обозначать в возрастающем порядке через x_1, x_2 и x_3. Тогда имеет место неравенство

$$\text{(a)} \qquad |y(x_2 - u)| < |y(x_2 + u)| \qquad (0 < u \leqq x_2 - x_1).$$

С помощью этого неравенства можно доказать теорему Сонина о монотонности максимумов функции $|y(x)|$, а также неравенства (8), (9) и (10), относящиеся соответственно к Эрмитовым ортогональным функциям, к полиномам Лежандра, и к решениям дифференциального уравнения Бесселя с индексом больше $1/2$. Последние три неравенства выводятся из соотношения

$$\int_{x_1}^{x_2} |y(x)|^p\,dx < \int_{x_2}^{x_3} |y(x)|^p\,dx \qquad (p \geqq 0)$$

являющегося следствием (a).

ÜBER EINEN REIHENTHEORETISCHEN SATZ VON FEJÉR

Von

G. FREUD (Budapest)

(*Vorgelegt von* P. TURÁN)

L. FEJÉR [1], [2] hat bewiesen, daß wenn die Funktion

$$f(z)=\sum_{n=0}^{\infty} a_n z^n \tag{1}$$

für $|z|<1$ regulär, beschränkt und schlicht ist, dann konvergiert die Potenzreihe (1) an jeder Stelle $z=e^{i\varphi}$, wo der radiale Grenzwert

$$\lim_{r=1-0} f(re^{i\varphi})=f(e^{i\varphi}) \tag{2}$$

existiert, und die Summe der Reihe ist gleich $f(e^{i\varphi})$.

Der Fejérsche Beweis gründet sich auf die Behauptung, daß das konforme Bild der Kreisscheibe $|z|<1$ bei der Abbildung $w=f(z)$ einen beschränkten Flächeninhalt hat:

$$\int_0^1\int_0^{2\pi} |f'(re^{i\varphi})|^2 r\,d\varphi\,dr=\pi\sum_{n=1}^{\infty} n|a_n|^2<\infty. \tag{3}$$

Hieraus folgt der Satz aus einfachen Abschätzungen.

Wir wollen zeigen, daß die Ungleichung (3), und somit auch der Fejérsche Satz gültig bleibt, wenn $f(z)$ nicht mehr schlicht, doch eine ganze rationale Funktion der im Einheitskreise regulären, schlichten und beschränkten Funktionen $\varphi_1(z), \varphi_2(z), \ldots, \varphi_n(z)$ ist:

$$f(z)=\sum_{k_1,k_2,\ldots,k_n=0}^{N} C_{k_1,k_2,\ldots,k_n}\varphi_1^{k_1}\varphi_2^{k_2}\cdots\varphi_n^{k_n}. \tag{4}$$

Eine Funktion, die im Einheitskreise regulär und beschränkt ist, und deren Potenzreihe (1) der Forderung (3) genügt, nennen wir *eine Fejérsche Funktion*. Folglich sind auch die Funktionen, die im Einheitskreise regulär, beschränkt und schlicht sind, Fejérsche Funktionen. Aus der Forderung (3) folgt, daß die Potenzreihe (1) einer Fejérschen Funktion an jeder Stelle $z=e^{i\varphi}$, wo der Grenzwert (2) existiert, gegen $f(e^{i\varphi})$ konvergiert. Wir beweisen:

SATZ. *Die Fejérschen Funktionen bilden einen Ring (im Sinne der Algebra).*

HILFSSATZ I. *Sind $f_1(z)$ und $f_2(z)$ Fejérsche Funktionen, dann ist auch $f(z)=f_1(z)+f_2(z)$ eine Fejérsche Funktion.*

BEWEIS. Mit $f_1(z)$ und $f_2(z)$ wird auch $f_1(z)+f_2(z)$ im Einheitskreise regulär und beschränkt. Es sei

$$(5) \qquad f_1(z)=\sum_{n=0}^{\infty} a_{1n} z^n, \quad f_2(z)=\sum_{n=0}^{\infty} a_{2n} z^n,$$

dann wird

$$\sum_{n=1}^{\infty} n|a_n|^2=\sum_{n=1}^{\infty} n|a_{1n}+a_{2n}|^2 \leq 2\sum_{n=1}^{\infty} n|a_{1n}|^2+2\sum_{n=1}^{\infty} n|a_{2n}|^2<\infty,$$

da die Koeffizienten der Potenzreihen (5) der Fejérschen Funktionen $f_1(z)$ und $f_2(z)$ der Forderung (3) genügen.

HILFSSATZ II. *Sind $f_1(z)$ und $f_2(z)$ Fejérsche Funktionen, dann ist auch $f(z)=f_1(z)f_2(z)$ eine Fejérsche Funktion.*

BEWEIS. Mit $f_1(z)$ und $f_2(z)$ wird auch $f_1(z)f_2(z)$ im Einheitskreise regulär und beschränkt. K sei eine gemeinsame Schranke von $f_1(z)$ und $f_2(z)$ für $|z|<1$. Dann wird

$$\int_0^1\int_0^{2\pi} |f'(re^{i\varphi})|^2 r\,d\varphi\,dr=\int_0^1\int_0^{2\pi} |f_1'(re^{i\varphi})f_2(re^{i\varphi})+f_1(re^{i\varphi})f_2'(re^{i\varphi})|^2 r\,d\varphi\,dr<$$

$$<K\int_0^1\int_0^{2\pi} |f_1'(re^{i\varphi})|^2 r\,d\varphi\,dr+K\int_0^1\int_0^{2\pi} |f_2'(re^{i\varphi})|^2 r\,d\varphi\,dr<\infty.$$

Endlich bemerken wir, daß auch $f(z)=1$ und allgemeiner $f(z)=a_0$ trivialerweise eine Fejérsche Funktion ist.

Daher gehört auch (4) zum Ringe der Fejérschen Funktionen, womit unsere Behauptung bewiesen ist.

Aus einem einfachen Beispiel von P. TURÁN[1] folgt, daß man die Forderung, daß $f_1(z)$ und $f_2(z)$ für $|z|<1$ beschränkt seien, nicht streichen kann.

In der Tat, es sei

$$f_1(z)=f_2(z)=\sum_{n=2}^{\infty} \alpha_n z^n, \quad \alpha_n=\frac{1}{n\log^{3/4} n}.$$

Wie leicht ersichtlich, ist $\lim\limits_{x=1-0} f(x)=+\infty$, also ist $f(z)$ keine in $|z|<1$ beschränkte Funktion.

Wir setzen

$$[f(z)]^2=\sum_{n=4}^{\infty} \beta_n z^n,$$

[1] Mündliche Mitteilung.

dann wird

$$\beta_n = \sum_{\nu=2}^{n} a_\nu a_{n-\nu} > \frac{1}{n \log^{5/8} n} \int_4^n \frac{dx}{x \log^{5/8} x} =$$

$$= \frac{8}{3}\left(\frac{1}{n \log^{1/4} n} - \frac{(\log 4)^{3/8}}{n \log^{5/8} n}\right) > \frac{1}{2} \frac{1}{n \log^{1/4} n} \quad \text{für } n > 150.$$

Mithin ist $\sum n |\beta_n|^2$, als eine Majorante der divergenten Reihe $\frac{1}{4} \sum_{150}^{\infty} \frac{1}{n \log^{1/2} n}$, selbst divergent.

Also bilden die im Einheitskreise regulären, schlichten Funktionen (1), die nur der Ungleichung (3) genügen, keinen Ring.

(*Eingegangen am 16. Juni 1952.*)

Literaturverzeichnis

1. L. Fejér. Über die Konvergenz der Potenzreihe an der Konvergenzgrenze in Fällen der konformen Abbildung auf die schlichte Ebene, *Schwarz Festschrift* (Springer, 1914), S. 42—53.
2. L. Fejér. Fourierreihe und Potenzreihe, *Monatshefte für Math.*, **28** (1917), S. 64—76.

ОБ ОДНОЙ ТЕОРЕМЕ ФЭЙЕРА В ТЕОРИИ РЯДОВ

Г. ФРАЙД (Будапешт)

(Резюме)

Функцию

$$f(x) = \sum_{n=0}^{\infty} a_n z^n \tag{I}$$

называем функцией типа Фэйера, если при $|z| < 1$ она регулярна и ограниченна, а её коэффициенты удовлетворяют условию

$$\sum_{n=1}^{\infty} n\,|a_n|^2 < \infty . \tag{II}$$

Доказываем, что функции этого типа образуют кольца. Отсюда следует следующее обобщение одной теоремы Фэйера:

Пусть $\varphi_1(z), \varphi_2(z), \ldots, \varphi_n(z)$ внутри единичной окружности регулярные, однозначные и ограниченные функции и

$$f(z) = \sum_{K_1, K_2, \ldots, K_n = 0}^{N} C_{K_1 K_2 \ldots K_n} \varphi_1^{K_1} \varphi_2^{K_2} \ldots \varphi_n^{K_n} . \tag{III}$$

Тогда степенный ряд (I) функции (III) сходится для каждого $z = e^{i\varphi}$, для которого существует предел $\lim\limits_{r \to 1-0} f(re^{i\varphi}) = f(e^{i\varphi})$.

П. Туран дал пример, согласно которому функции, обладающие степенным рядом типа (II), не образуют кольца, если они неограниченны.

THE DIRECT SUM OF CYCLIC GROUPS

By

L. FUCHS (Budapest)

(*Presented by* L. RÉDEI)

§ 1. Introduction

In the theory of infinite Abelian groups[1] an important rôle is played by the groups which can be represented as the direct sum of its (finite or infinite) cyclic subgroups. In view of the basis theorem for finitely generated Abelian groups, we may regard these groups as direct generalizations of the finitely generated Abelian groups. Their importance lies in the fact that several problems in infinite group theory lead to such groups as well as in the fact that they belong to the class of those infinite groups whose structure is completely known. Therefore, it seems to be an important problem to obtain criteria which guarantee the decomposibility of an Abelian group into the direct sum of cyclic groups.

We shall find it convenient to discuss this problem in another formulation. It is evident that the stated problem is equivalent to the question of the existence of a basis of the group where by a *basis* of an Abelian group G we mean a subset $B=(a_\nu)$ of G such that G is the direct sum of the cyclic groups[2] $\{a_\nu\}$, in sign:

$$G=\sum_{a_\nu \in B}\{a_\nu\}.$$

For the sake of convenience we agree to assume that the order of each element of a basis is either infinity or a power of a natural prime (in general, we shall call such elements *primary*). Evidently, this means no proper restriction on the basis, for from each basis of G we can obtain in an obvious manner a basis all of whose elements are primary.

From the definition of direct sums it follows at once that a basis of G may be characterized as a subset B of G satisfying the following two conditions:

[1] The group operation will be written as addition. — We emphasize that we do not assume anything concerning the power of the groups considered.

[2] The symbol $\{S\}$ where S is some subset of G will denote the subgroup of G generated by S.

1. It is an *independent system* in the sense that a relation

$$m_1 a_{\nu_1} + \cdots + m_k a_{\nu_k} = 0 \qquad (a_{\nu_i} \in B)$$

(where the m_i's are rational integers) implies

$$m_1 a_{\nu_1} = \cdots = m_k a_{\nu_k} = 0.$$

This means that if[3] $O(a_{\nu_i})$ is finite, then $O(a_{\nu_i}) | m_i$ and if $O(a_{\nu_i})$ is infinite, then we have $m_i = 0$.

2. It is a *generating system*, i. e., each element a of G may be written in the form

$$a = n_1 a_{\mu_1} + \cdots + n_s a_{\mu_s} \qquad (a_{\mu_j} \in B)$$

with suitable rational integers n_j.

Clearly, 1. and 2. imply that a basis is a *maximal* independent system of G (i. e. there is no independent system in G properly containing a basis), and that a basis is a *minimal* generating system of G (i. e. no proper subset of a basis generates G). But by simple examples it may readily be shown that if a subset of G is a maximal independent or a minimal generating system of G, then it is not yet necessarily a basis of G.[4] In view of this, it seems to be natural to raise the following two dual problems.

PROBLEM 1. When is a maximal independent system of an Abelian group G a basis of G?

PROBLEM 2. When is a minimal generating system of an Abelian group G a basis of G?

The first part of the present paper is devoted to discussing both of these problems. In the literature I have failed to find a general answer to these problems. The sufficient or necessary and sufficient conditions for the decomposibility of Abelian groups into the direct sum of cyclic groups are in general inadequate for solving these problems. A partial answer to Problem 1 is contained in a recent result by A. KERTÉSZ [3] for the case of p-groups (and hence for torsion groups),[5] while a theorem due to T. SZELE [8] solves Problem 2, but his method (based on an idea of R. RADO [7]) seems to be somewhat complicated and not to be applicable to Problem 1. According to my knowledge a general method giving a complete and unified solution of the two stated problems fails to exist.

If one tries to obtain necessary and sufficient criteria for a basis of an Abelian group G in terms of the fundamental notions of infinite Abelian

[3] By $O(x)$ we denote the order of an element x of G.

[4] For example, let G be the additive group of the rational numbers with square free denominators. Then each non-zero element of G forms a maximal independent subsystem of G, and the set of the reciprocals of all primes is a minimal generating system (cf. SZELE [8]). Nevertheless G has no basis.

[5] See the remark at the end of his paper.

group theory such as order or height,[6] then one will meet unsurmountable difficulties caused by the fact that by means of these notions it is very hard to make distinction between elements of infinite order. To overcome this difficulty we shall introduce a relative concept defined for certain pairs of elements of the group. This new concept is closely connected with the common notion of order and therefore we shall call it *relative order*. The statement "a is of a greater relative order than b" is understood to have the usual meaning provided b is of finite order, and, for elements a, b of infinite order, it will be defined in case one of a and b belongs to a previously given independent subsystem of G. By making use of this new concept we shall settle Problems 1 and 2 by proving that *a maximal independent system B of G is a basis of G if and only if no element of it can be replaced by an element of G of a greater relative order without violating the independence of B* (Theorem 1), and that *a minimal generating system B of G is a basis of G if and only if no element of it can be replaced by an element of G of a smaller relative order so as to get again a generating system of G* (Theorem 2). It is to be noted that the given independent subsystem of G in the definition of relative order *is not the same* in the two cases, but this seems to be quite natural, considering that in the first case B is known to be independent, but not in the second. It is almost evident that in the first case the relative order is to be taken with respect to B, while in the second case each element of B must be considered as an independent system consisting of a single element.

In Part II we give some applications; these we shall base on the first basis criterion which is very easy to handle. It will turn out that a certain part of the theory of those infinite Abelian groups which are representable as the direct sum of cyclic groups may be derived from this common source. For example, one is able to give a new proof of KULIKOV's theorem stating that all subgroups of a group with a basis have again a basis, or to discuss the Abelian extension problem of a group with a basis by a group with a basis, and also, several known results are easy consequences of it, some of which we shall prove here, while for further applications we refer to KERTÉSZ [3]. Once in the applications, viz. in the proof of KULIKOV's theorem, we have to use the well-ordering hypothesis, but no non-trivial group-theoretic result will be made use of.

As to the notations we remark that by the letters a, b, c, x we denote group elements, by capital letters subsets of the group, while the letters such as i, k, m, n etc. are reserved for rational integers, in particular, p for a prime number. The Greek letters ν, μ may range over an arbitrary (not necessarily well-ordered) set of indices; ϱ, σ, τ are used to denote ordinal numbers.

[6] Recall that an element a of prime power order p^n is said to have the height k if k is the greatest integer for which the equation $p^k x = a$ has a solution x in G. (a is of infinite height if this equation is solvable for every natural integer k.)

PART I. THE TWO BASIS CRITERIA

§ 2. The relative order

As indicated, we introduce a new concept in terms of which we shall be able to solve our Problems 1 and 2.

What we mean by saying that an element a of an Abelian group G is of a greater relative order than an element b of G needs no explanation in case either both a and b are of finite order or a is of infinite and b of finite order. But in certain cases we can adequately define the notion "of a greater relative order" even in case both elements are of infinite order. This may be done as follows.

Assume we are given an independent system S of G.[7] We shall say that an element a of G with $O(a)=\infty$ is *of a greater relative order* than an element b of G with $O(b)=\infty$, if and only if one of them, say b, belongs to S and there exists a non-trivial dependence relation

$$ra = sb + s_1 b_1 + \cdots + s_k b_k \neq 0 \qquad (b, b_i \in S, b \neq b_i) \tag{1}$$

such that

$$|r| > |s| \neq 0. \tag{2}$$

More precisely, we must say that a is of a greater relative order than b *with respect to S*, but when there is no risk of ambiguity, we omit the reference to S. In case in (1) $|r| < |s|$ holds, a will be said to be *of a smaller relative order* than b.

If none of the elements a and b of infinite order belongs to S, or if one of them belongs to S, but there is no dependence relation of type (1) with a non-vanishing term sb, then our notion is undefined.

The requirement that one of the elements a and b whose infinite orders are to be compared in the relative sense shall belong to a given independent system, implies at once that the definition is unique. Indeed, if there is another non-trivial relation of type (1), say,

$$r'a = s'b + s_1' b_1 + \cdots + s_k' b_k \neq 0,$$

then by eliminating a from (1) and the last equation, we arrive at the relation

$$(r's - rs')b + (r's_1 - rs_1')b_1 + \cdots + (r's_k - rs_k')b_k = 0.$$

Hence we have $r's - rs' = 0$, since $b, b_1, \ldots, b_k$, as elements of S, are independent. Therefore

$$\frac{r}{s} = \frac{r'}{s'},$$

i. e. the ratio $\frac{r}{s}$ in (1) is uniquely determined by a and b. Consequently, we are led to the conclusion that the definition is actually unique.

[7] S need not be a *maximal* independent system.

Unicity implies that if we have to decide whether or not a with $O(a)=\infty$ is of a greater relative order than b with $O(b)=\infty$, it is sufficient to consider only one relation of type (1).

It is readily verified that, if S is fixed, the relative order is not necessarily transitive. For example, if in (1) we have $|s_1|>|r|>|s|\neq 0$ with $O(a)=O(b)=O(b_1)=\infty$, then b_1 is of a greater relative order than a and a is of a greater relative order than b, but the elements b_1 and b are incomparable in the sense of relative order, since they belong to S and are therefore independent with respect to S.

Finally, if we are not given an independent system S, then we agree to consider one of the elements a and b of infinite order as an independent system consisting of a single element. Then a is of a greater relative order than b if and only if there is a relation $ra=sb\neq 0$ such that $|r|>|s|$.

§ 3. The first basis criterion

Now, by making use of the notion of relative order, we are able to solve our stated Problem 1.

THEOREM 1. (THE FIRST BASIS CRITERION.)[8] *A subset B of an Abelian group G is a basis of G if and only if*

(i) *it is a maximal independent system of G,*

(ii) *no element of B can be replaced by an element of a greater relative order (with respect to B) without violating the independence.*

PROOF OF NECESSITY.[9] Assume that the group G is the direct sum of the cyclic groups $\{a_\nu\}$ where ν ranges over a certain set I of indices,

$$G=\sum_{\nu\in I}\{a_\nu\}.$$

It is evident that the set B of all a_ν's satisfies (i).

In order to show that for this set B (ii) also holds, let b be an arbitrary non-zero element of G. We clearly have a relation

$$b=m_1a_1+\cdots+m_ka_k \qquad (m_ia_i\neq 0,\ k\geqq 1) \tag{3}$$

[8] It is readily seen that for p-groups this theorem expresses the same thing as KERTÉSZ' result on principal systems [3]: *An Abelian p-group G containing no elements of infinite height is the direct sum of cyclic groups if and only if G has a principal system,* where by a principal system is meant a maximal independent subset no element of which can be replaced by an element of a greater height of G without violating independence. The corollaries of KERTÉSZ' theorem may be considered as those of Theorem 1; for them we refer to [3]. — Let us mention that this basis criterion and its consequences may be transferred word by word to Abelian groups with Euclidean operator ring. (This has also been observed by T. SZELE.)

[9] In this section the relative order is to be understood with respect to the given system B.

for some suitable notation of the a_ν's. It is obvious that no element a_ν of B other than one of $a_1, \ldots, a_k$ can be replaced by b without violating the independence of B. But an a_i of finite order with $O(a_i) < O(b)$ can not be replaced by b in the required manner, for multiplying (3) by $O(a_i)$ we get a non-trivial dependence relation containing b and some of $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_k$. Furthermore, if b is of infinite order, then not all of $a_1, \ldots, a_k$ are of finite order. But an a_j with $O(a_j) = \infty$ can not be of a smaller relative order than b, considering that (3) is a relation of type (1) such that $|1| \leqq |m_j| \neq 0$.

Proof of sufficiency. Let B be a subsystem of G satisfying (i) and (ii). We remark that there is no loss of generality in supposing the same thing as in our definition of the basis, i. e. that the order of each element of finite order of B is a power of a prime. For, from each maximal independent system we can get in an obvious manner another maximal independent system satisfying this condition, and if some element of finite order of this new system might be replaced by an element of a greater order without violating the independence, then the same would be true for the original system. (Even the converse is easily seen to hold.)

Now let us consider the cyclic groups generated by the elements a_ν of B. On account of (i), the direct sum of all these groups $\{a_\nu\}$ exists in G; we denote it by H:

$$H = \sum_{a_\nu \in B} \{a_\nu\}.$$

We are going to prove that $H = G$.

Suppose there is a b in G which does not belong to H. By (i), some multiple of b necessarily belongs to H. We may clearly assume $b \in G$ to be chosen so that $b \notin H$ but $pb \in H$ for some prime p, i. e.

$$pb = m_1 a_1 + \cdots + m_k a_k \qquad (a_i \in B,\ m_i a_i \neq 0,\ k \geqq 0). \tag{4}$$

Moreover, even the following additional requirements may be assumed to hold for b: (α) none of the coefficients m_i of the elements a_i in (4) is divisible by p; (β) among the a_i's of finite order none occurs whose order is not a power of this p; (γ) all the coefficients belonging to elements a_i of infinite order are less than p in absolute value. In fact, each term on the right hand side of (4) which can be expressed in the form px with $x \in H$ disappears by introducing $b' = b - x$ instead of b, and for this b' we have again $b' \notin H$ and $pb' \in H$ with an equation of type (4).

Now we prove that if (α)—(γ) hold for (4), then we have $k = 0$, in other words, $pb = 0$.

If there is an element a_j of infinite order in (4), then $O(b) = \infty$ and by (γ) we have $|p| > |m_j| \neq 0$. This means that b is of a greater relative order than a_j. But if we substitute b for a_j in B, the arising system is plainly again independent. This contradicts (ii) and hence there is no a_j of infinite order in (4).

In case there is some element a_i of finite order in (4), then let

$$\max_{1 \leq i \leq k} O(a_i) = O(a_1) = p^r \qquad (r \geq 1).$$

Multiplying (4) by p^{r-1} we have

$$p^r b = m_1 p^{r-1} a_1 + \cdots + m_k p^{r-1} a_k$$

which is not equal to 0 by virtue of (α). Hence, by (β), it results $O(b) = p^{r+1} > O(a_1)$. Now a_1 may be replaced by b in B without violating the independence, for in a dependence relation

$$n_1 b + n_2 a_2 + \cdots + n_k a_k = 0 \quad \text{with } n_1 \not\equiv 0 \pmod{p^{r+1}}$$

(where we may suppose $n_1 = p n_1'$ for some integer n_1') we might put in place of pb the expression (4) to obtain

$$n_1' m_1 a_1 + \cdots = 0 \qquad \text{with } n_1' m_1 \not\equiv 0 \pmod{p^r},$$

in contradiction to independence. Therefore, indeed, we have $pb = 0$.

In order to finish the proof of Theorem 1, we observe that by the maximal independence of B, b depends on B, and hence, in view of $O(b) = p$, it follows that b itself can be expressed as a linear combination of the a_ν's; in other words, $b \in H$. This contradiction establishes our first basis criterion.

§ 4. The second basis criterion

Next we turn our attention to solving Problem 2. Here we find the following dual of Theorem 1.

THEOREM 2. (THE SECOND BASIS CRITERION.) *A subsystem B of an Abelian group G is a basis of G if and only if*

(I) *it is a generating system not containing* 0,

(II) *no element of B can be replaced by an element of a smaller relative order*[10] *of G so as to get again a generating system of G.*

Before going into the proof of the theorem, we note that (II) immediately implies that B must be a *minimal* generating system of G. In fact, in the contrary case the superfluous element ($\neq 0$) of the generating system could be replaced by 0, i. e. by an element of a smaller order.

PROOF OF NECESSITY. Assume that G is the direct sum of cyclic groups:

(5) $$G = \sum_{a_\nu \in B} \{a_\nu\}$$

[10] Since in the present case B is not known to be independent, here and throughout this section the relative order is to be taken with respect to one of the elements whose infinite orders are to be compared in the relative sense.

where B is a basis of G. Each non-zero element b of G may be expressed in the form

(6) $$b = m_1 a_1 + \cdots + m_k a_k \qquad (m_i a_i \neq 0, k \geqq 1).$$

If b is of finite order, then so are the a_i's too ($i = 1, \ldots, k$). Clearly, from the elements of B only one of those $a_1, \ldots, a_k$ can be replaced by b (provided we want to have the resulting system again a generating system) whose coefficient m_i is prime to its order $O(a_i)$. But the order of such an a_i can not exceed the order of b, since by multiplying (6) by $O(b)$, all members on the right must vanish, and therefore we have $O(a_i) = O(m_i a_i) \leqq O(b)$.

If $b \in G$ is of infinite order and of a smaller relative order than a_j, then we have a relation

$$m b = m' a_j \neq 0 \qquad \text{with } |m| < |m'|.$$

This means that in (6) a_j is the only a_i of infinite order and its coefficient $|m_j| > 1$. But then a_j does not belong to the group generated by B with a_j replaced by b. This establishes the necessity of the stated condition.

Proof of sufficiency. Let $B = (a_\nu)$ be a subset of G satisfying (I) and (II). In order to show that (5) holds for G, we prove that there is no non-trivial dependence relation

(7) $$n_1 a_1 + n_2 a_2 + \cdots + n_k a_k = 0 \qquad (n_i a_i \neq 0, k \geqq 1).$$

If here there is an a_i of infinite order, say $O(a_1) = \infty$, then in B we may replace a_1 by $(1 + |n_1|) a_1$ which is of a smaller relative order than a_1 (note that $n_1 \neq 0$), and we get again a generating system of G, since a_1 obviously belongs to the group generated by the new system. Therefore, in (7) there is no a_i of infinite order; consequently, we may suppose that in (7) the orders of the a_i's are powers of one and the same prime p. Let p^t ($t \geqq 0$) be the greatest power of p which divides all the coefficients n_i, and write $n_i = p^t n_i'$. By the choice of t there is some index j ($1 \leqq j \leqq k$) with $(n_j', p) = 1$. Now the element

$$b = n_1' a_1 + \cdots + n_j' a_j + \cdots + n_k' a_k$$

satisfies $p^t b = 0$ and we have obviously $O(b) < O(a_j)$. Since n_j' is relatively prime to $O(a_j)$, we may replace a_j by b in B and we shall get again a generating system for G, because the group generated by this new system obviously contains a_j. Hence we conclude that B is an independent system of G whence the theorem follows.[11]

[11] The idea of the proof for the case of finite order is the same as I have applied to proving the basis theorem for finite Abelian groups. Cf. [2].

PART II. APPLICATIONS

§ 5. Kulikov's theorem on the subgroups of the direct sum of cyclic groups

As an application of our first basis criterion we shall prove KULIKOV's fundamental theorem [4] on the subgroups of a group with a basis:[12]

THEOREM 3. (KULIKOV.) *If H is a subgroup of a group with a basis, then H also possesses a basis, or, otherwise expressed, each subgroup of the direct sum of cyclic groups is again the direct sum of cyclic groups. (An equivalent statement is that a subdirect sum of cyclic groups has a basis.)*

Let B be a basis of G. We well-order the elements of B in some way and fix this well-ordering $\mathfrak{W}$. a_τ will denote the element of B to which the ordinal number τ is assigned.

In order to make the proof of Theorem 3 easier, we introduce the following notions.[13] Let $b \neq 0$ be an arbitrary primary element of G. We clearly have[14]

$$nb = n_1 a_{\tau_1} + n_2 a_{\tau_2} + \cdots + n_k a_{\tau_k} \neq 0 \qquad (n_i a_{\tau_i} \neq 0, k \geqq 1). \tag{8}$$

From this relation we can obtain a non-trivial dependence relation with a minimal number k of non-zero terms, by multiplying it by a suitable integer. If b is of infinite order, then in this relation only a_τ's of infinite order remain, and if b is of prime power order, then the orders of the occurring a_τ's are powers of the same prime p. Moreover, in the second case we may assume that by multiplying the relation (with the minimal k) by this prime p, all terms vanish. In both cases such a relation will be called *a minimal dependence relation belonging to b.*

If $b \neq 0$ is again a primary element of G, then we shall say that the basis element a_{τ_1} of G is *associated with b* if a_{τ_1} occurs in a minimal dependence relation (8) belonging to b and in the well-ordering $\mathfrak{W}$ we have

$$\tau_1 > \tau_2 > \cdots > \tau_k.$$

Evidently, a_{τ_1} is uniquely determined by the primary element b, and a_{τ_1} is of finite or infinite order according as b is of finite or infinite order.

[12] It seems to me that the following proof of KULIKOV's theorem is the first uniform proof. The original proof necessitated to distinguish three cases according as the group considered was torsion, mixed or torsion free group. Moreover, the proof below is a more direct one: we not only state that the subgroup H possesses a basis, but also give a method by which a basis can be constructed in case a well-ordering of a basis of G is known (for example, for countable Abelian groups).

[13] These notions depend not only on the group G, but also on the basis B of G and the well-ordering $\mathfrak{W}$ of B.

[14] We may evidently choose $n = 1$, but this is not necessary for our definition of the new notions.

Now we are going to give a method which replaces each basis element a_τ of G by an element b_τ of H or simply omits a_τ, and then we shall verify that the arising system $B'=(b_\tau)$ is a basis of the subgroup H.

For each fixed ordinal number τ, we consider the element a_τ of B. If there is no primary element b of H such that a_τ is associated with b, then we omit a_τ and take no b_τ for this τ. On the other hand, if a_τ is associated with at least one element of H, then from the set of these elements b we select one in the following manner:

1. If a_τ is of finite order, say, $O(a_\tau)=p^t$, then we take a b_τ of a possibly greatest order p^s. (Such a b_τ necessarily exists, since we clearly have $s \leqq t$.)

2. If a_τ is of infinite order, then we take a b_τ such that in (8) (with $b=b_\tau$ and $\tau_1=\tau$) the ratio

$$v=\left|\frac{n_1}{n}\right|>0$$

shall have a value as least as possible. It is readily checked that such a b_τ also exists, for by the basis property of B, v has to be a natural number, as one sees at once from (8).[15]

Now we show the system B' of these b_τ's is a basis for the subgroup H. To this end it is sufficient to verify the fulfilment of (i) and (ii) of our first basis criterion.

1°. First of all, B' is an *independent* system of H. In fact, if B' were not independent, then we should have a non-trivial dependence relation

$$q_{\sigma_1}b_{\sigma_1}+\cdots+q_{\sigma_h}b_{\sigma_h}=0 \qquad (\sigma_1>\cdots>\sigma_h) \tag{9}$$

with a minimal number h of non-zero terms. Then either all of b_σ's are of finite or all are of infinite order. In the first alternative we can suppose that a prime number p annihilates all terms. This means that the orders $O(b_\sigma)$ are powers of this fixed p and each coefficient q_σ is divisible by a power of p whose exponent is precisely one less than the exponent of p in $O(b_\sigma)$. Since the same holds for n in a minimal dependence relation (8), we may express each $q_\sigma b_\sigma$ in (9) by means of the a_τ's and then obtain a relation for the a_τ's. But by the choice of the indices we see that a_{σ_1} occurs only once and this term is not equal to zero. This contradicts the independence of the a_τ's. — A similar reasoning applies to proving that no non-trivial dependence relation may exist containing elements b_σ of infinite order. (Multiply (9) by a suitable integer and then substitute again (8).)

2°. B' is a *maximal* independent system of H. For, if not, then there would be a $b\in H$ which is independent of B'; moreover, we may clearly assume this b to be primary. Among the ordinal numbers τ of the basis

[15] Recall that (8) may be obtained from a relation of the same type with $n=1$.

elements a_τ associated with these independent b's let σ be the first one and let $b'\in H$ be an element (independent of B') whose associated basis element is a_σ. Then we have a minimal dependence relation

$$qb' = q_1 a_\sigma + q_2 a_{\sigma_2} + \cdots + q_l a_{\sigma_l} \neq 0 \qquad (\sigma > \sigma_2 > \cdots > \sigma_l).$$

Now let us consider b_σ; it necessarily exists, for H has elements with which a_σ is associated. Let

$$nb_\sigma = n_1 a_\sigma + n_2 a_{\tau_2} + \cdots + n_k a_{\tau_k} \neq 0 \qquad (\sigma > \tau_2 > \cdots > \tau_k).$$

Of course, there is no loss of generality in assuming $q_1 = n_1$. Then by subtraction we are led to the equation

$$qb' - nb_\sigma = q_2 a_{\sigma_2} + \cdots + q_l a_{\sigma_l} - n_2 a_{\tau_2} - \cdots - n_k a_{\tau_k}.$$

Hence $b'' = qb' - nb_\sigma \in H$ is either 0 or the ordinal number ϱ of the basis element a_ϱ associated with b'' is less than σ. In the first alternative the proof is finished: b' depends on b_σ; while in the second case we infer that, by the choice of σ, b'' depends on B'. If b' is of infinite order, the proof is again finished. If b' and hence b'' is of finite order, then the order of b'' is a prime p, and the dependence of b'' on B' implies a relation

$$b'' = m_1 b_{\varrho_1} + \cdots + m_r b_{\varrho_r}$$

whence

$$qb' = nb_\sigma + m_1 b_{\varrho_1} + \cdots + m_r b_{\varrho_r}$$

is a non-trivial dependence relation for b': a contradiction. The proof of (i) is thus completed.

3°. Next we turn our attention to the proof of (ii). Suppose there is an element[16] b of H which is of a greater relative order than and can be substituted, without violating independence, for some element of B'. Let

$$hb = h_1 b_{\sigma_1} + \cdots + h_s b_{\sigma_s} \neq 0 \qquad (\sigma_1 > \cdots > \sigma_s,\ s \geqq 1) \tag{10}$$

be a minimal dependence relation. It is obvious that from the elements of B' only one of $b_{\sigma_1}, \ldots, b_{\sigma_s}$ can be replaced by b so as to get again an independent system. In case b and hence all of $b_{\sigma_1}, \ldots, b_{\sigma_l}$ are of finite order, there is no loss of generality in supposing $O(b) = p^n > O(b_{\sigma_i})$ for $i = 1, \ldots, s$, since each term $h_i b_{\sigma_i}$ having the form $p^{n-1} h_i' b_{\sigma_i}$ may be brought to the left side and then we may replace b by a suitable element of H which we denote again by b. Next we substitute for each b_σ in (10) its expression (8) and then we get an expression for hb in terms of the a_τ's which shows that a_{σ_1} is associated with b. But $O(b) > O(b_{\sigma_1})$ contradicts the choice of b_{σ_1} (see 1.).

In case all terms in (10) are of infinite order, we assume b to be taken such that $0 \neq |h_i| < |h|$ $(i = 1, \ldots, s)$. By multiplying (10) by an appropriate integer q and then introducing, by means of (8), the a_τ's in place of the b_τ's,

[16] Of course, b can be assumed to be a primary element.

we obtain a relation

$$hqb = h_1 qb_{\sigma_1} + \cdots = h_1 \frac{q}{n} n_1 a_{\sigma_1} + \cdots$$

showing that a_{σ_1} is associated with b. But here the ratio of the coefficients of a_{σ_1} and b is

$$0 < \left| \frac{h_1 \frac{q}{n} n_1}{hq} \right| = \left| \frac{h_1}{h} \right| \cdot \left| \frac{n_1}{n} \right| < \left| \frac{n_1}{n} \right| = r,$$

against the choice of b_{σ_1}. This completes the proof of Theorem 3.

REMARK 1. It may be of some interest to note that the idea of the above proof of KULIKOV's theorem admits of another proof of the same theorem, a proof which does not appeal to either basis criterion. The following demonstration makes use of the same method of constructing the basis B' of the subgroup H, and therefore, it may be considered as a proof suggested directly by the first basis criterion.

Not only the basic idea, but even a large part of this second proof coincides with the proof given above. The only difference is that instead of entering into the proof of the maximality of the independence of the system B' and the fulfilment of (ii), we argue as follows.

We intend to show that B' is a generating system of H. The proof is indirect: we assume that there are elements in H which do not belong to $\{B'\}$.

If there is a $b \in H$ of finite order with this property, then we choose such a b' of a possibly least order[17] p^s with the additional requirement that in H there is no b of the same order p^s not in $\{B'\}$ whose associated basis element a_τ precedes in $\mathfrak{W}$ the basis element a_σ associated with b'. Now b' may be written in the form

$$b' = \cdots + h_1 a_\sigma + h_2 a_{\sigma_2} + \cdots + h_l a_{\sigma_l} \qquad (\sigma > \sigma_2 > \cdots > \sigma_l) \tag{11}$$

where the terms before $h_1 a_\sigma$ occur in no minimal dependence relation belonging to b'. It is evident that b_σ certainly exists and we have

$$b_\sigma = \cdots + n_1 a_\sigma + n_2 a_{\tau_2} + \cdots + n_k a_{\tau_k} \qquad (\sigma > \tau_2 > \cdots > \tau_k) \tag{12}$$

where the terms before $n_1 a_\sigma$ vanish in a minimal dependence relation belonging to b_σ. Since a_σ is associated with both of b' and b_σ, by the choice of b_σ we have $O(b') \leqq O(b_\sigma) = p^t$, and we infer that $p^r | n_1$ implies $p^r | h_1$. Consequently, we can find an integer q such that the expression of $b'' = b' - qb_\sigma$ in terms of the a_τ's, got from (11) and (12), does not contain a_σ. It is readily seen that either the basis element a_ϱ associated with b'' precedes a_σ in $\mathfrak{W}$ (i. e. $\varrho < \sigma$) or, if this is not the case, the order of b'' is less than $O(b') = p^s$; the last statement holding for in (11) and in qb_σ the terms before the term containing a_σ are annihilated by p^{s-1}. By the choice of b', in both cases we conclude

[17] It is evident that the order of such a b' must be a power of a prime.

that b'', and hence also b', belongs to $\{B'\}$. Therefore there is no b of finite order in H and not in $\{B'\}$.

However, it neither exists such a b of infinite order. For, assuming the contrary, we could choose a $b' \in H$, $\notin \{B'\}$ whose associated basis element a_σ is in the well-ordering $\mathfrak{W}$ the first one for which such a b' exists. Now taking again b_σ, the choice of b_σ implies that in (11) and (12) we have $|n_1| \leqq |h_1|$. Moreover, we can verify without difficulty the divisibility relation $n_1 | h_1$. Indeed, if we put $h_1 = qn_1 + r$ with $0 \leqq r < |n_1|$, then in $b'' = b' - qb_\sigma$ the coefficient of a_σ is equal to r, and thus, in case $r \neq 0$, b'' would have a_σ as associated basis element with a less coefficient than b_σ has, contrary to the choice of b_σ. The remaining part of the proof is the same as in the case of finite order (note that the second alternative is that b'' is of finite order). The second proof is thus completed. (If our principal aim were to give a new proof, as simple as possible, of KULIKOV's theorem, then we should prefer this second demonstration.)[18]

REMARK 2. KULIKOV's theorem suggests to ask for the existence of a basis of the factor group G/H of a group G with a basis with respect to a subgroup H of G. It is easy to illustrate by an example that in general G/H has no basis. In fact, let[19]

$$G = \mathfrak{Z}(p) + \mathfrak{Z}(p^2) + \cdots + \mathfrak{Z}(p^n) + \cdots = \sum_{n=1}^{\infty} \mathfrak{Z}(p^n),$$

and let a_n be a generator of $\mathfrak{Z}(p^n)$; then we have $O(a_n) = p^n$. We put

$$H = \{a_1 - pa_2, a_2 - pa_3, \ldots, a_n - pa_{n+1}, \ldots\}.$$

Now we show that $H \neq G$. It is easy to see that a_1 does not belong to the subgroup H; for, if we had for some integers h_i

$$a_1 = h_1(a_1 - pa_2) + \cdots + h_n(a_n - pa_{n+1}) \tag{13}$$

with

$$h_n(a_n - pa_{n+1}) \neq 0, \tag{14}$$

then by the independence of the a_i's we could conclude that $a_1 = h_1 a_1$ as well as the coefficients of $a_2, \ldots, a_{n+1}$ on the right of (13) are divisible by $p^2, \ldots, p^{n+1}$, respectively. In particular, we get $h_n \equiv 0 \pmod{p^n}$ which contradicts (14).

Therefore we see that

$$a_1 \not\equiv 0, \quad pa_2 \equiv a_1, \; \ldots, \; pa_{n+1} \equiv a_n, \; \ldots \pmod{H}$$

showing that the factor group G/H (generated by the cosets containing $a_1, a_2, \ldots$, respectively) is isomorphic to $\mathfrak{Z}(p^\infty)$. Consequently, G/H has no basis.

[18] Let us here mention that if $H = G$, our construction shows how to construct a new basis of a group G if another is given. It may be of interest to determine, if given two bases B and B' of G, under what circumstances the basis B can be well-ordered so that B' can be obtained from B by a construction described above.

[19] We denote by $\mathfrak{Z}(r)$ the cyclic group of order r and by $\mathfrak{Z}(p^\infty)$ (where p is a prime) the group of type p^∞ (PRÜFER [6]).

§ 6. Abelian extension of direct sum of cyclic groups by direct sum of cyclic groups

The first basis criterion enables us to discuss the problem of Abelian group extensions[20] of groups with a basis by groups with a basis. It is easy to show that in general we can not conclude to the existence of a basis of the extension group, as it appears from Theorem 6.

Now it is natural to raise the question as to conditions which guarantee the existence of a basis for the extension group. A complete answer to this problem is given by Theorem 6.

We shall need the following theorem which generalizes a known result of H. PRÜFER and R. BAER (see Corollary 2 in § 7).

THEOREM 4. *If for an Abelian group G there is a natural integer n such that nG is the direct sum of cyclic groups, then G itself is the direct sum of cyclic groups.*

We denote by $B=(a_\nu)$ a basis of nG and let $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r}$ be the canonical representation of n. For each ν, let $b_\nu \in G$ be an arbitrary, but fixed solution of the equation $nx=a_\nu$. The set of these b_ν's is independent, for from a non-trivial dependence relation with certain b_ν's by multiplication by an adequate divisor of n we could get a non-trivial dependence relation with certain a_ν's which is impossible.

Next we adjoin elements of G of order $p_1^{\alpha_1}$ to the set of these b_ν's so as to get an independent system which is as maximal as possible. (ZORN's lemma guarantees the existence of such a system.) Then we do the same with elements of order $p_1^{\alpha_1-1},\ldots,p_1,p_2^{\alpha_2},\ldots,p_2,\ldots,p_r^{\alpha_r},\ldots,p_r$, respectively. We state that the arising system C, consisting of the elements b_ν and the new ones (which we shall denote by c_μ), is a basis of G.

That C is an independent system needs no verification. Further, if we had an element c of G independent of C, then on account of $nc \in nG$ we could conclude $nc=0$. Now we may suppose that the order of c is a prime power, say, p_1^{α} with $\alpha \leqq \alpha_1$. This contradicts the above construction of C and therefore C is a maximal independent system in G, indeed.

No element of C can be replaced by an element c of G of a greater relative order without violating the independence (c may be assumed to be a primary element). Indeed, for the elements b_ν this is obvious from the observation that c is of a greater relative order than b_ν if and only if $nc \in nG$ is of a greater relative order than $nb_\nu=a_\nu$. On the other hand, from the replacement of a certain element c_μ by c we could infer by the construction of C that $O(c)=p_i^{\alpha}$ with $\alpha>\alpha_i$ for some i ($i=1,\ldots,r$). This is, however,

[20] We consider throughout only Abelian group extensions. — This extension problem (as well as part 1. in the proof of Theorem 6) has been suggested to the author by T. SZELE.

impossible, since then $0 \neq nc \in nG$ would depend on the a_ν's and hence the b_ν's. This establishes the statement.

From Theorems 3 and 4 we conclude:

THEOREM 5. *If for an Abelian group G there exist a natural integer n and a subgroup H such that*

$$nG \subseteq H \subseteq G$$

and H is the direct sum of cyclic groups, then also G is the direct sum of cyclic groups.

In fact, if H has a basis, then by KULIKOV's theorem nG has too, consequently, Theorem 4 implies the statement.

Now we are in a position to prove what we have indicated at the beginning of this section:

THEOREM 6. *Let the groups H and F be direct sums of cyclic groups. A necessary and sufficient condition that all Abelian group extensions G of H by F shall again be direct sums of cyclic groups is:*

in case H contains elements of infinite order: *that the orders of the elements of finite order in $G/H \simeq F$ are bounded,*

in case H is a torsion group: *that the orders of the p-primary*[21] *elements in F are bounded for each prime p which occurs as the order of some element of H.*

At first we reduce the problem to the case when F is a torsion group (the assertion shows that only the torsion subgroup of F has an effect to the existence of a basis of G). To this end we decompose G/H into two direct summands:

$$G/H = G_1/H + G_2/H$$

such that G_1/H contains all direct summands of infinite order and G_2/H does those of finite order.

We show that G_1 has a basis. Let (a_ν) be a basis of H and $(\bar{b}_\mu)$ a basis of G_1/H. Then the set B consisting of all a_ν's and of all b_μ's (where b_μ is an arbitrary element of G contained in the coset $\bar{b}_\mu$) is a basis of G_1. In fact, firstly, it is evidently a generating system of G_1, secondly, no non-trivial relation

$$n_1 a_1 + \cdots + n_k a_k + m_1 b_1 + \cdots + m_l b_l = 0 \quad (n_i a_i \neq 0, m_j b_j \neq 0) \tag{15}$$

may hold. For, this relation considered modulo the subgroup H implies $m_j \bar{b}_j = 0$ for $j = 1, \ldots, l$, and since the $\bar{b}_j$'s are all of infinite order, we hence obtain $m_j = 0$ $(j = 1, \ldots, l)$. Taking this into account we see that (15) is a dependence relation in H, whence $n_i a_i = 0$ $(i = 1, \ldots, k)$ is obtained. Thus B

[21] We say that a is a p-primary element if $O(a)$ is some power of p.

is an independent generating system of G_1, and therefore G_1 is the direct sum of cyclic groups, as we wished to prove.

Now we pass from G_1 to G where we clearly have

$$G/G_1 \simeq G_2/H,$$

consequently, to complete the proof of Theorem 6 it suffices to consider the case when F is a torsion group. Therefore, in the remaining part of the proof we may assume that F contains no elements of infinite order.

SUFFICIENCY. If the orders of the elements in F have an upper bound m, then applying Theorem 5 for G and its subgroup H with $n = m!$, we conclude that G is the direct sum of cyclic groups.

Next we show that the same conclusion holds if the orders of the p-primary elements in the factor group F are bounded for each prime p being the order of some element of the torsion group H.

We shall need the following simple

LEMMA. *Let F be a p-group and $H = H_p + H'$ a direct decomposition of the torsion group H where H_p is the p-component of H. Then each Abelian extension G of H by F has a direct decomposition*

$$G \simeq G_p + H'$$

where G_p is the p-component of G.

The statement of the Lemma follows at once from the observation that each coset of G modulo H has a representative whose order is some power of p and G_p is generated by these representatives and H_p.

Resuming the proof of the sufficiency, we decompose F into its p-components:

$$F = \sum_p{}' F_p.$$

It is quite clear that the extension G of H by F may be regarded as a group obtained by successive extensions of H by $F_2, F_3, F_5, \ldots,$ respectively. Since, by our Lemma, for each prime p, the p-component H_p of $H = \sum_p{}' H_p$ does not alter during this sequence of procedures except when we are extending just by the factor group F_p, we see that the extension G of H by F will be the direct sum of p-groups and of extensions of p-groups by p-groups (according as one of F_p and H_p equals zero or none). Our hypothesis on the orders of the elements of F implies that, for each p with $H_p \neq 0$, the elements of F_p are of bounded order; thus — by what has been proved already — it follows that each p-extension and hence G itself is the direct sum of cyclic groups.

NECESSITY. We show that if the hypothesis is not fulfilled, there is at least one extension G' of H by F which is no direct sum of cyclic groups.

1. Assume that H contains an element a of infinite order and the orders of the elements in the torsion group F have no upper bound, i. e.

$$F=\sum_{\nu\in I}\mathfrak{Z}(r_\nu)$$

where the orders r_ν of the cyclic groups $\mathfrak{Z}(r_\nu)$ are not bounded. Let G' be the Abelian group generated by H and by the elements a_ν $(\nu\in I)$ connected by the defining relations

$$r_\nu a_\nu=a \qquad \text{for each } \nu\in I.$$

Then we have $G'/H\simeq F$, and G' is no direct sum of cyclic groups, for it contains an element a of infinite order for which the equation

$$r_\nu x=a$$

is solvable for an infinity of the r_ν's.

2. Let the torsion group H contain an element a of order p^k and of height 0; further we assume that the elements of F_p are not of bounded order. The same construction of G' as in 1. shows that in the extension group G' the p-primary element a is of infinite height. Therefore G' can not be the direct sum of cyclic groups.

This completes the proof of Theorem 6.

REMARK. (*Added in the proof, 10 December 1952.*) I have observed that since the present paper was completed, a recent paper by J. DIEUDONNÉ has appeared which discusses some problems on p-groups closely related to the results of this section. For details we refer to his paper: Sur les p-groups abéliens infinis, *Portugaliae Math.*, **11** (1952), pp. 1—5.

§ 7. Some corollaries

In this section we show that several known results on the decomposibility of Abelian groups into the direct sum of cyclic groups may be obtained as corollaries of our discussions.

We begin with the simpiest one:

COROLLARY 1. (BASIS THEOREM FOR FINITELY GENERATED ABELIAN GROUPS.) *A finitely generated Abelian group G is the direct sum of a finite number of cyclic groups.*

If $g_1, g_2, \ldots, g_k$ are the generators of G and $a_1, a_2, \ldots, a_m$ $(m\leqq k)$ is a maximal independent system of G, then for each g_i there is a natural integer n_i such that $n_i g_i\in\{a_1,\ldots,a_m\}=H$. If n denotes the least common multiple of these n_i's, then we may apply Theorem 5 to the present case to obtain the statement.

COROLLARY 2. (PRÜFER—BAER.)[22] *An Abelian group G whose elements are of bounded order, is the direct sum of cyclic groups.*

[22] See PRÜFER [6] and BAER [1].

If m is an upper bound for the orders, then for $n = m!$ we clearly have $nG = 0$, and we may again make use of Theorem 5 with H the zero group.

COROLLARY 3. (PONTRJAGIN.)[23] *A countable torsion free Abelian group G is the direct sum of cyclic groups if and only if every ascending chain of subgroups of an arbitrary finite rank n contains but a finite number of different subgroups.*

Only the sufficiency of the condition needs a verification.

Let $a_1, a_2, \ldots, a_n, \ldots$ be a maximal independent system of G. For each $n = 1, 2, \ldots,$ we choose an element b_n such that in the dependence relation

$$mb_n = m_1 a_1 + \cdots + m_n a_n \qquad (m_n a_n \neq 0) \tag{16}$$

the ratio

$$v = \left| \frac{m_n}{m} \right| > 0$$

assumes its least value. Such a b_n necessarily exists, for the subgroup H_n of G consisting of all b's in G with $0 \neq qb \in \{a_1, a_2, \ldots, a_n\}$ for some integer q is of rank n, and hence, by hypothesis, finitely generated. Consequently, there is some natural integer t_n such that in (16) for each $b_n \in H$ we may choose m such that $|m| \leq |t_n|$. Following the idea of the above proof of KULIKOV's theorem it is easy to conclude that $b_1, b_2, \ldots$ form a basis for G.[24]

(Received 24 August 1952)

Bibliography

[1] R. BAER, Der Kern, eine charakteristische Untergruppe, *Compositio Math.*, **1** (1935), pp. 254—283.

[2] L. FUCHS, A simple proof of the basis theorem for finite Abelian groups, *Norske Vid. Selsk. Forh.*, **25** (1952), nr. 23.

[3] A. KERTÉSZ, On the decomposibility of abelian p-groups into the direct sum of cyclic groups, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 121—126.

[4] Л. Я. Куликов, К теории абелевых групп произвольной мощности, Матем. Сборник, **16** (58) (1945), pp. 129—162.

[5] L. PONTRJAGIN, *Topological groups* (Princeton, 1946).

[6] H. PRÜFER, Untersuchungen über die Zerlegbarkeit der abzählbaren primären Abelschen Gruppen, *Math. Zeitschrift*, **17** (1923), pp. 35—61.

[7] R. RADO, A proof of the basis theorem for finitely generated Abelian groups, *Journ. London Math. Soc.*, **26** (1951), pp. 74—75.

[8] T. SZELE, On the direct sum of cyclic groups, *Publicationes Math., Debrecen,* **2** (1951), pp. 76—78.

[9] T. SZELE, On a theorem of Pontrjagin, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 121—123.

[23] See PONTRJAGIN [5], pp. 168—169. For a simple proof see SZELE [9]. The rank of an Abelian group G is the cardinal number of a maximal independent system in G.

[24] The remaining part of the proof is simpler than the proof of KULIKOV's theorem since the group is now torsion free.

О ПРЯМОЙ СУММЕ ЦИКЛИЧЕСКИХ ГРУПП

Л. ФУКС (Будапешт)

(Резюме)

Пусть G аддитивная абелева группа произвольной мощности. Подмножество B группы G называется базисом, если G является прямой суммой циклических групп, порождённых элементами подмножества B. Очевидно, что базис B характеризуется двумя следующими свойствами:

1. Элементы его (линейно) независимы.

2. B является системой порождающей G.

Если же о некотором подмножестве S известно лишь то, что оно является максимальной независимой системой в G, или минимальной системой порождающей G, отсюда в общем случае не следует то, что S является базисом. Автор доказывает, что если среди элементов бесконечного порядка группы G некоторым образом определить понятие относительного порядка, то справедливы две следующие теоремы:

Теорема 1. Подмножество B группы G тогда и только тогда является базисом, если B является максимальной независимой системой G так, что ни один из элементов B не может быть заменён элементом группы G с большим относительным порядком — без ущерба независимости B.

Теорема 2. Подмножество B тогда и только тогда является базисом, если B является минимальной системой порождающей G и ни один элемент в B не может быть заменён элементом группы G с меньшим относительным порядком так, чтобы снова получить порождающую систему.

Специальный случай теоремы 1, относящийся к p-группам, совпадает с одним из новых результатов А. Кертеса.

Теорема 1 оказывается полезной при исследованиях разложения на прямую сумму циклических групп. Так при её помощи можно дать рассмотрения разных случаев, единое доказательство известной теоремы Куликова, согласно которой любая подгруппа прямой суммы циклических групп тоже может быть разложена на прямую сумму циклических групп. В качестве нового результата даётся полное решение следующей проблемы: найти необходимое и достаточное условие того, чтобы любое абелево расширение группы A посредством B было прямой суммой циклических групп, если только данные A и B являются прямой суммой циклических групп.

LÖSUNG EINES MARKOVSCHEN PROBLEMS BETREFFS EINER AUSDEHNUNG DES BEGRIFFES DER ELEMENTAREN FUNKTION

Von

ILONA BERECZKI (Szeged)

(*Vorgelegt von* L. KALMÁR)

1. Diese Arbeit enthält die Lösung eines Problems, welches von A. A. MARKOV im Zusammenhang mit meinem Vortrag am Ersten Ungarischen Mathematiker-Kongreß aufgeworfen wurde. Zweck jenes Vortrags[1] war die Klarstellung der Relation der Klasse der elementaren Funktionen und der Klasse der rekursiven Funktionen. Unter einer elementaren Funktion verstehe ich dabei eine arithmetische Funktion, welche aus ihren Argumenten und aus nichtnegativen ganzen Konstanten durch eine endliche Anzahl von Additionen, arithmetischen Subtraktionen, Multiplikationen, arithmetischen Divisionen, Summen- und Produktenbildungen entsteht.[2] Hier verstehe ich unter einer arithmetischen Funktion eine Funktion, deren Argumente die nichtnegativen ganzen Zahlen durchlaufen und deren Werte ebenfalls nichtnegative ganze Zahlen sind; unter arithmetischer Subtraktion verstehe ich die Bildung des absoluten Betrags der Differenz, unter arithmetischer Division die des ganzen Teils des Quotienten.[3] Unter einer rekursiven (genauer: primitiv-rekursiven) Funktion wird eine arithmetische Funktion verstanden, welche sich aus der Konstante 0 und aus der Funktion $x+1$ durch eine endliche Anzahl von Substitutionen (d. h. Bildung einer Funktion von Funktionen) und Rekursionen von der Form

$$(1)\qquad \begin{cases} \varphi(0, x, y, \ldots) = \alpha(x, y, \ldots), \\ \varphi(n+1, x, y, \ldots) = \beta(n, x, y, \ldots, \varphi(n, x, y, \ldots)) \end{cases}$$

gewinnen läßt.

[1] Vgl. meine Arbeit: Nem elemi rekurzív függvény létezése (mit einem deutschen Auszug: Existenz einer nichtelementaren rekursiven Funktion), im Erscheinen in den *Berichten des Ersten Ungarischen Mathematiker-Kongresses*.

[2] Dieser Begriff geht auf P. CSILLAG zurück. Vgl. auch L. KALMÁR, Egyszerű példa eldönthetetlen aritmetikai problémára (mit einem deutschen Auszug: Ein einfaches Beispiel für ein unentscheidbares arithmetisches Problem), *Mat. és Fiz. Lapok,* **50** (1943), S. 1—23.

[3] Die letztgenannten beiden Operationen, angewandt auf die Zahlen x und y, werden wie üblich durch $|x-y|$ und $[x/y]$ bezeichnet. Für $y=0$ verstehen wir unter $[x/y]$ etwa die Zahl 0

Die rekursiven Funktionen dienen als Hilfsmittel bei verschiedenen Untersuchungen im Gebiete der mathematischen Logik. Z. B. hat GÖDEL seinen berühmten Satz,[4] wonach es zu beliebigen, gewisse ziemlich allgemeine Bedingungen erfüllenden, Axiomensystemen arithmetische Probleme gibt, die im fraglichen Axiomensystem unlösbar sind, mittels Betrachtung gewisser rekursiven Funktionen bewiesen. Im Laufe einer Vereinfachung des Beweises des Gödelschen Satzes hat KALMÁR elementare Funktionen statt rekursiver verwendet.[5] Auch bei gewissen anderen Anwendungen ist es möglich, die rekursiven Funktionen durch elementare Funktionen zu ersetzen. So erhob sich die Frage, ob die Klasse der elementaren Funktionen nicht etwa mit der Klasse der rekursiven Funktionen identisch ist.

Es ist längst bekannt, daß jede elementare Funktion rekursiv ist. Daher läuft die obige Frage darauf hinaus, ob auch umgekehrt, jede rekursive Funktion elementar ist. Am Ersten Ungarischen Mathematiker-Kongreß habe ich gezeigt, daß dies nicht der Fall ist, da die Funktion, welche durch eine ähnliche Iteration aus der Potenzierung entsteht, wie die Potenzierung aus der Multiplikation oder die Multiplikation aus der Addition, zwar offenbar eine rekursive, jedoch keine elementare Funktion ist. MARKOV warf in diesem Zusammenhange die Frage auf, ob es richtig ist, daß auch bei einer Verallgemeinerung des Begriffes der elementaren Funktionen, wobei die Rolle der Addition, der arithmetischen Subtraktion, der Multiplikation und der arithmetischen Division die Anwendung einer endlichen Anzahl von beliebigen rekursiven Funktionen, die Rolle der Summen- und der Produktenbildung aber eine endliche Anzahl von solchen Funktionaloperationen übernimmt, die in ähnlicher Weise aus gewissen beliebigen rekursiven Funktionen entstehen, wie die Summenbildung aus der Addition oder die Produktenbildung aus der Multiplikation, die so entstehende Funktionenklasse nicht die Klasse der rekursiven Funktionen erschöpfen kann. E. EGERVÁRY warf bei gleicher Gelegenheit die Frage auf, ob es richtig ist, daß bei der Modifikation des Begriffes der elementaren Funktionen, wobei die Produktenbildung nicht verwendet wird, die Potenz nicht bereits zur modifizierten Funktionenklasse gehört.

2. In dieser Arbeit werde ich einen allgemeinen Satz („Hauptsatz") beweisen, woraus sich die Richtigkeit der Vermutungen von MARKOV und EGERVÁRY gleichzeitig ergeben. Zur Formulierung dieses Satzes benötige ich nicht nur die erwähnte Iteration der Potenzierung, sondern auch die Funktionen, die daraus durch weitere Iterationen entstehen. Nennt man die Addition die Operation erster Ordnung, die Multiplikation die Operation zweiter Ordnung, die Potenzierung die Operation dritter Ordnung, die Iteration der Poten-

[4] K. GÖDEL, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monatshefte für Math. und Phys.*, 38 (1931), S. 173—198.

[5] A. a. O.[2]

zierung die Operation vierter Ordnung usw. und bezeichnet man die Operation m-ter Ordnung, angewandt auf die Zahlen a und b, durch $a \overset{m}{\nabla} b$, so kann man diese Operationen durch eine Folge von Rekursionen wie folgt definieren:[6]

$$(2) \qquad \begin{cases} a \overset{1}{\nabla} b = a + b, \\ a \overset{2}{\nabla} 0 = 0, \\ a \overset{m+1}{\nabla} 0 = 1 \quad \text{für} \quad m = 2, 3, \ldots, \\ a \overset{m+1}{\nabla} b + 1 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} b) \quad \text{für} \quad m = 1, 2, 3, \ldots \end{cases}$$

Man sieht sofort, daß diese Gleichungen im Falle $m = 1$ in die rekursive Definition

$$\begin{cases} a \overset{2}{\nabla} 0 = 0. \\ a \overset{2}{\nabla} b + 1 = a + (a \overset{2}{\nabla} b) \end{cases}$$

des Produktes übergehen, daher ist wirklich $a \overset{2}{\nabla} b = ab$; im Falle $m = 2$ aber in die rekursive Definition

$$\begin{cases} a \overset{3}{\nabla} 0 = 1 \\ a \overset{3}{\nabla} b + 1 = a \cdot (a \overset{3}{\nabla} b) \end{cases}$$

der Potenz, daher ist wirklich $a \overset{3}{\nabla} b = a^b$. Durch eine geringe Modifikation der Definition (2) können wir erreichen, daß auch die rekursive Definition der Summe $a \overset{1}{\nabla} b = a + b$, nämlich

$$\begin{cases} a \overset{1}{\nabla} 0 = a \\ a \overset{1}{\nabla} b + 1 = (a \overset{1}{\nabla} b) + 1 \end{cases}$$

einbezogen wird. Damit die zweite dieser Gleichungen mit der letzten Gleichung von (2) für $m = 0$, d. h. mit

$$a \overset{1}{\nabla} b + 1 = a \overset{0}{\nabla} (a \overset{1}{\nabla} b)$$

übereinstimme, definieren wir die Operation 0-ter Ordnung als die Addition von 1 zur *zweiten* Zahl, auf die die Operation angewandt wird, d. h. durch $a \overset{0}{\nabla} b = b + 1$. Also lassen sich die Operationen $a \overset{m}{\nabla} b$ durch die Gleichungen

[6] Um Klammern zu ersparen, betrachte ich ∇ als ein Zeichen, welches stärker trennt, als das Pluszeichen und das Minuszeichen (und natürlich auch stärker, als das Multiplikationszeichen). Z. B. bedeutet $a \overset{m+1}{\nabla} b + 1$ dasselbe, wie $a \overset{m+1}{\nabla} (b+1)$; dagegen muß $(a \overset{m+1}{\nabla} b) + 1$ mit Klammern geschrieben werden.

$$(3)\qquad \begin{cases} a \overset{0}{\nabla} b = b+1, \\ a \overset{m+1}{\nabla} 0 = \begin{cases} a & \text{für } m=0, \\ 0 & \text{für } m=1, \\ 1 & \text{für } m \geqq 2, \end{cases} \\ a \overset{m+1}{\nabla} b+1 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} b) \end{cases}$$

definieren.

3. Aus den beiden letzten Gleichungen (3) gewinnt man der Reihe nach für $m=0, 1, 2, 3, \ldots$ je eine rekursive Definition von der Form (1) für $a \overset{1}{\nabla} b (=a+b)$, $a \overset{2}{\nabla} b (=ab)$, $a \overset{3}{\nabla} b (=a^b)$, $a \overset{4}{\nabla} b, \ldots$; also ist $a \overset{m}{\nabla} b$ für jedes (numerisch gegebene) m eine rekursive Funktion von a und b. Man kann (3) natürlich auch als eine Definition der *dreistelligen* Funktion $a \overset{m}{\nabla} b$ als Funktion von *a, b und m* auffassen. Dann ist aber (3) keine Rekursion von der Form (1), sondern eine sogenannte zweifache Rekursion (gleichzeitig nach den Argumenten m und b). Daher braucht $a \overset{m}{\nabla} b$ keine rekursive Funktion von a, m und b zu sein. In der Wirklichkeit ist sie auch keine. In der Tat sind die Gleichungen (3) durch eine geringe Modifikation aus der Definition

$$\begin{cases} \varphi(a, b, 0) = a+b \\ \varphi(a, 0, m+1) = \begin{cases} 0 & \text{für } m=0, \\ 1 & \text{für } m=1, \\ a & \text{für } m \geqq 2, \end{cases} \\ \varphi(a, b+1, m+1) = \varphi(a, \varphi(a, b, m+1), m) \end{cases}$$

der Funktion $\varphi(a, b, m)$ entstanden, die bei ACKERMANN[7] als Beispiel einer durch eine zweifache Rekursion definierten nichtrekursiven (d. h. nicht primitivrekursiven) Funktion dient; und das gleiche kann ähnlicherweise auch für die Funktion $a \overset{m}{\nabla} b$ bewiesen werden. Indessen werde ich in dieser Arbeit nicht die Eigenschaft der Funktion $a \overset{m}{\nabla} b$ benützen, daß sie keine rekursive Funktion ist; vielmehr folgt diese Eigenschaft aus dem zu beweisenden Hauptsatz. Was aber den Beweis jenes Hauptsatzes betrifft, beruht er auf gewissen Ideen, die in etwas anderer Form auch bei ACKERMANN zum Beweise der Nichtrekursivität der Funktion $\varphi(a, b, m)$ verwendet wurden.

ACKERMANNS Beweis beruht im wesentlichen darauf, daß es zu jeder rekursiven Funktion α eine natürliche Zahl m gibt, so daß α in einem gewissen Sinne von der Funktion (von a und b) $\varphi(a, b, m)$ majorisiert wird. In dieser Arbeit werden aber etwas andere Majorisierungseigenschaften als bei ACKERMANN benötigt werden und auch statt der Ackermannschen Funktion $\varphi(a, b, m)$ wird die obige Funktion $a \overset{m}{\nabla} b$ verwendet. Der Grund der Heranziehung der

[7] W. ACKERMANN, Zum Hilbertschen Aufbau der reellen Zahlen, *Math. Annalen*, 99 (1928), S. 118—133.

Funktion $a \overset{m}{\bigtriangledown} b$ statt $\varphi(a, b\, m)$ ist folgendes. Die für uns nötigen Majorisierungseigenschaften können viel einfacher aus einem analogen Satze von R. PÉTER[8] abgeleitet werden als aus den bei ACKERMANN stehenden Majorisierungsbehauptungen. PÉTER verwendet statt der Ackermannschen Funktion $\varphi(a, b, m)$ eine zweistellige Funktion[9] $\xi(m, n)$, die ebenfalls durch eine zweifache Rekursion, nämlich durch

$$(4) \quad \begin{cases} \xi(0, n) = n+1, \\ \xi(m+1, 0) = \xi(m, 1), \\ \xi(m+1, n+1) = \xi(m, \xi(m+1, n)) \end{cases}$$

definiert wird. Für diese Funktion beweist sie den folgenden Satz:

Für jede rekursive Funktion $\alpha(x, y, \ldots)$ mit einer beliebigen Anzahl von Argumenten gibt es eine nichtnegative ganze Zahl m, so daß für alle Werte der Argumente $x, y, \ldots$ die Ungleichung

$$\alpha(x, y, \ldots) < \xi(m, \max(x, y, \ldots))$$

gilt.

Nun lassen sich aus diesem Satz auf Grund eines ziemlich einfachen Zusammenhanges der Funktionen $\xi(m, n)$ und $a \overset{m}{\bigtriangledown} b$, der kein Analogon für die Funktion $\varphi(a, b, m)$ besitzt, die folgenden Sätze beweisen, die die für uns nötigen Majorisierungseigenschaften der Funktion $a \overset{m}{\bigtriangledown} b$ aussprechen:

SATZ I. *Für jede rekursive Funktion $\alpha(x, y, \ldots)$ mit einer beliebigen Anzahl von Argumenten gibt es nichtnegative ganze Zahlen p und k, für die und für alle Werte der Argumente $x, y, \ldots$ die Ungleichung*

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{p}{\bigtriangledown} k$$

besteht.

SATZ II. *Für jede zweistellige rekursive Funktion $\beta(x, y)$ gibt es eine nichtnegative ganze Zahl q, für welche und für alle Werte der Argumente x, y die Ungleichung*

$$\beta(x, y) < \max(3, x) \overset{q}{\bigtriangledown} \max(3, y)$$

besteht.

4. Auf Grund dieser Sätze werde ich dann den schon erwähnten Hauptsatz beweisen. Zur genauen Formulierung des Hauptsatzes ist es nötig, den folgenden Begriff einzuführen, welcher eine Verallgemeinerung des Begriffes der elementaren Funktionen bildet.

[8] R. PÉTER, *Rekursive Funktionen* (Budapest, 1951), insb. S. 72.

[9] Ich bezeichne diese Funktion statt der Péterschen Bezeichnung $\psi(m, n)$ mit $\xi(m, n)$, um mit ψ gewisse solche Funktionen bezeichnen zu können, die aus den mit φ bezeichneten Funktionen mittels Substitution oder Rekursion entstehen.

Es seien K und I zwei Klassen arithmetischer Funktionen; die Elemente von I seien lauter zweistellige Funktionen. Es sei $E(K, I)$ die engste Funktionenklasse (d. h. der Durchschnitt aller Funktionenklassen) mit den folgenden drei Eigenschaften.

1) $K \subseteq E$ (d. h. für $\alpha \in K$ gilt $\alpha \in E$).

2) Ist φ_0 eine r-stellige Funktion, ferner $\varphi_1, \ldots, \varphi_r$ (höchstens) s-stellige Funktionen und ist $\varphi_0 \in E, \varphi_1 \in E, \ldots, \varphi_r \in E$, so ist auch die Funktion $\psi(x_1, \ldots, x_s) = \varphi_0(\varphi_1(x_1, \ldots, x_s), \ldots, \varphi_r(x_1, \ldots, x_s))$ ein Element von E (d. h. E ist abgeschlossen in Bezug auf Substitution).

3) Ist $\varphi \in E$ und $\beta \in I$, und ist φ etwa $r+1$-stellig, so ist die durch die Rekursion

$$(5) \qquad \begin{cases} \psi(0, x_1, \ldots, x_r) = \varphi(0, x_1, \ldots, x_r) \\ \psi(n+1, x_1, \ldots, x_r) = \beta(\varphi(n+1, x_1, \ldots, x_r), \psi(n, x_1, \ldots, x_r)) \end{cases}$$

definierte Funktion ψ ein Element von E (d. h. E ist abgeschlossen in Bezug auf Rekursion der Form (5)). (Man bemerke, daß diese Rekursion eine Verallgemeinerung der Summen- und Produktenbildung bedeutet; für $\beta(x, y) = x + y$ ist ja $\psi(n, x_1, \ldots, x_r) = \sum_{i=0}^{n} \varphi(i, x_1, \ldots, x_r)$ und für $\beta(x, y) = xy$ gilt $\psi(n, x_1, \ldots, x_r) = \prod_{i=0}^{n} \varphi(i, x_1, \ldots, x_r)$.)

Man kann die Elemente der Klasse $E(K, I)$ die in Bezug auf die zu K gehörigen Ausgangsfunktionen und auf die zu I gehörigen Iterationsfunktionen elementaren Funktionen nennen. Besteht nämlich die Klasse K aus den Funktionen $x+y$, $|x-y|$, xy, $[x/y]$ und die Klasse I aus den Funktionen $x+y$ und xy, so ist $E(K, I)$, wie leicht ersichtlich, die Klasse der elementaren Funktionen.[10]

Nun lautet der Hauptsatz wie folgt.

HAUPTSATZ. *Gibt es eine nichtnegative ganze Zahl p, so daß für eine beliebige Funktion $\alpha \in K$, für eine geeignete nichtnegative ganze Zahl k und für alle Werte der Argumente $x, y, \ldots$ die Ungleichung*

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{p}{\nabla} k,$$

sowie eine nichtnegative ganze Zahl q, so daß für eine beliebige Funktion $\beta \in I$ und für alle Werte der Argumente x, y die Ungleichung

$$\beta(x, y) \leqq \max(3, x) \overset{q}{\nabla} \max(3, y)$$

[10] Um dies einzusehen, hat man zu beweisen, daß die Klasse der elementaren Funktionen in Bezug auf Substitution, ferner daß die Klasse $E(K, I)$ im Falle $K = \{x+y, |x-y|, xy, [x/y]\}$ in Bezug auf Addition, arithmetische Subtraktion, Multiplikation und arithmetische Division abgeschlossen ist. Dies beweist man auf Grund der Definition der elementaren Funktionen bzw. der Klasse $E(K, I)$ ohne Schwierigkeiten.

besteht, so gibt es eine nichtnegative ganze Zahl m, so daß für eine beliebige Funktion $\psi(x_1, \ldots, x_r) \in E(K, I)$, *für eine geeignete nichtnegative ganze Zahl k und für alle Werte der Argumente* $x_1, \ldots, x_r$ *die Ungleichung*

$$\psi(x_1, \ldots, x_r) < \max(3, x_1, \ldots, x_r) \overset{m}{\nabla} k$$

gilt; und zwar kann $m = \max(p, q+2)$ *gewählt werden.*

In **5** werde ich gewisse Hilfssätze für die Funktion $a \overset{m}{\nabla} b$ beweisen, die teilweise analog mit gewissen von ACKERMANN (a. a. O.[7]) für seine Funktion $\varphi(a, b, m)$ bewiesenen Hilfssätzen sind. In **6** folgen dann die Beweise der Sätze I und II und in **7** der Beweis des Hauptsatzes. In **8** wird aus dem Hauptsatz auf die Richtigkeit der Vermutungen von MARKOV und EGERVÁRY, ferner auf die meines a. a. O.[1] bewiesenen Satzes, endlich auf die Nichtrekursivität der Funktion $a \overset{m}{\nabla} b$ als Funktion von a, m und b gefolgert werden.

5. Zum Beweise der Sätze I und II benötige ich einige Eigenschaften der Funktion $a \overset{m}{\nabla} b$, daher beweise ich vor allem einige Hilfssätze.

HILFSSATZ 1. *Für* $m \geqq 3$ *gilt* $1 \overset{m}{\nabla} b = 1$.

BEWEIS durch vollständige Induktion nach m. Für $m = 3$ lautet unsere Behauptung $1^b = 1$, was offenbar richtig ist. Gilt die Behauptung für $m (\geqq 3)$, so gilt sie auch für $m+1$, da nach (3)

$$1 \overset{m+1}{\nabla} 0 = 1$$

und

$$1 \overset{m+1}{\nabla} b+1 = 1 \overset{m}{\nabla} (1 \overset{m+1}{\nabla} b) = 1$$

ist. Damit ist die Behauptung bewiesen.

HILFSSATZ 2. *Für* $m \geqq 2$ *gilt* $a \overset{m}{\nabla} 1 = a$.

BEWEIS durch vollständige Induktion nach m. Für $m = 2$ lautet die Behauptung $a \cdot 1 = a$. Angenommen, daß die Behauptung für $m (\geqq 2)$ richtig ist, erhält man aus (3) wegen $a \overset{m+1}{\nabla} 0 = 1$

$$a \overset{m+1}{\nabla} 1 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} 0) = a \overset{m}{\nabla} 1 = a.$$

HILFSSATZ 3. *Für* $m \geqq 1$ *gilt* $2 \overset{m}{\nabla} 2 = 4$.

BEWEIS durch vollständige Induktion nach m. Für $m = 1$ lautet die Behauptung $2 + 2 = 4$. Nehmen wir an, daß die Behauptung für $m (\geqq 1)$ schon richtig ist. Wegen $m+1 \geqq 2$ folgt aus Hilfssatz 2 $2 \overset{m+1}{\nabla} 1 = 2$, daher gewinnt man aus (3)

$$2 \overset{m+1}{\nabla} 2 = 2 \overset{m}{\nabla} (2 \overset{m+1}{\nabla} 1) = 2 \overset{m}{\nabla} 2 = 4.$$

Hilfssatz 4. *Für* $a \geqq 2$ *und* $m \geqq 3$, *oder auch für* $a \geqq 2$ *und* $b \geqq 1$, *gilt* $a \overset{m}{\nabla} b > b$.

Beweis durch vollständige Induktion nach m. Für $m = 0, 1, 2$ geht die Behauptung der Reihe nach in die Ungleichungen $b + 1 > b, a + b > b, a \cdot b > b$ über, welche für $a \geqq 2$ und $b \geqq 1$ gewiß richtig sind. Für $m = 3$ lautet die Behauptung: $a^b > b$, falls $a \geqq 2$. Man beweist dies durch vollständige Induktion nach b. In der Tat lautet für $b = 0$ die zu beweisende Ungleichung $a^0 > 0$, d. h. $1 > 0$. Angenommen, daß $a^b > b$, d. h. $a^b \geqq b + 1$, gewinnt man wegen $a > 1$

$$a^{b+1} = a^b \cdot a > a^b \geqq b + 1.$$

Nehmen wir nun an, daß die Behauptung des Hilfssatzes für $m (\geqq 3)$ richtig ist; wir beweisen dann durch vollständige Induktion nach b, daß sie auch für $m + 1$ gültig bleibt. In der Tat ist für $b = 0$ wegen (3) $a \overset{m+1}{\nabla} 0 = 1 > 0$; gilt die Behauptung, d. h. $a \overset{m+1}{\nabla} b \geqq b + 1$ für einen Wert von b, so gilt wegen (3) und der Induktionsvoraussetzung (d. h. $a \overset{m}{\nabla} c > c$ für $a \geqq 2$ und beliebiges c)

$$a \overset{m+1}{\nabla} b + 1 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} b) > a \overset{m+1}{\nabla} b \geqq b + 1.$$

Damit ist Hilfssatz 4 bewiesen.

Nun werde ich beweisen, daß die Funktion $a \overset{m}{\nabla} b$ in jeder ihrer Veränderlichen monoton wächst, falls wir für die beiden anderen Veränderlichen hinreichend große Werte einsetzen. Diese Eigenschaft der Funktion $a \overset{m}{\nabla} b$ werde ich öfters benützen, ohne mich stets besonders auf die entsprechenden Hilfssätze zu berufen.

Die Monotonität in Bezug auf b ist eine Folge von

Hilfssatz 5. *Für* $a \geqq 1$ *haben wir* $a \overset{m}{\nabla} b + 1 \geqq a \overset{m}{\nabla} b$ *und für* $a \geqq 2$ *sogar* $a \overset{m}{\nabla} b + 1 > a \overset{m}{\nabla} b$.

Beweis. Wir haben $b + 2 > b + 1$, d. h. $a \overset{0}{\nabla} b + 1 > a \overset{0}{\nabla} b$ und $a + b + 1 > > a + b$, d. h. $a \overset{1}{\nabla} b + 1 > a \overset{1}{\nabla} b$ für alle a und b. Ferner haben wir $a(b + 1) > ab$, d. h. $a \overset{2}{\nabla} b + 1 > a \overset{2}{\nabla} b$ für $a \geqq 1$. Ist $m \geqq 3$, so ist nach Hilfssatz 1

$$1 \overset{m}{\nabla} b + 1 = 1 \overset{m}{\nabla} b = 1,$$

so daß wir nur noch $a \overset{m}{\nabla} b + 1 > a \overset{m}{\nabla} b$ für $a \geqq 2$ und $m \geqq 3$ zu beweisen haben. Ist $m = 3$, so geht diese Behauptung in die Ungleichung $a^{b+1} > a^b$ für $a \geqq 2$ über, die offenbar ebenfalls richtig ist. Für $m \geqq 4$ hat man endlich nach (3) und Hilfssatz 4, wegen $a \geqq 2$ und $m - 1 \geqq 3$,

$$a \overset{m}{\nabla} b + 1 = a \overset{m-1}{\nabla} (a \overset{m}{\nabla} b) > a \overset{m}{\nabla} b,$$

wodurch die Behauptung des Hilfssatzes in sämtlichen Fällen bewiesen ist.

KOROLLAR. *Die Funktion* $a \overset{m}{\nabla} b$ *ist in Bezug auf* b *für* $a \geqq 1$ *im weiteren Sinne, für* $a \geqq 2$ *im engeren Sinne monoton wachsend; d. h. es gilt für* $a \geqq 1$ *und* $b \geqq c$ *die Ungleichung* $a \overset{m}{\nabla} b \geqq a \overset{m}{\nabla} c$, *und für* $a \geqq 2$ *und* $b > c$ *die Ungleichung* $a \overset{m}{\nabla} b > a \overset{m}{\nabla} c$.

Die Monotonität in Bezug auf a ist eine Folge von

HILFSSATZ 6. *Man hat für alle Werte von* a, b *und* m *die Ungleichung* $a+1 \overset{m}{\nabla} b \geqq a \overset{m}{\nabla} b$.

BEWEIS durch vollständige Induktion nach m. Für $m = 0, 1, 2$ lautet die zu beweisende Ungleichung der Reihe nach $b+1 \geqq b+1$, $(a+1)+b \geqq a+b$ und $(a+1)b \geqq ab$, die sämtlich offenbar richtig sind. Nehmen wir an, daß die Behauptung für $m (\geqq 2)$ schon richtig ist, und wir beweisen sie für $m+1$ durch vollständige Induktion nach b. In der Tat gilt nach (3) für $m \geqq 2$

$$a+1 \overset{m+1}{\nabla} 0 = 1 = a \overset{m+1}{\nabla} 0,$$

so daß die zu beweisende Ungleichung $a+1 \overset{m+1}{\nabla} b \geqq a \overset{m+1}{\nabla} b$ für $b = 0$ gewiß richtig ist. Ist sie für b richtig, so hat man wegen (3) und Hilfssatz 5, Korollar, ferner wegen der Induktionsvoraussetzung $a+1 \overset{m}{\nabla} c \geqq a \overset{m}{\nabla} c$,

$$a+1 \overset{m+1}{\nabla} b+1 = a+1 \overset{m}{\nabla} (a+1 \overset{m+1}{\nabla} b) \geqq a+1 \overset{m}{\nabla} (a \overset{m+1}{\nabla} b) \geqq a \overset{m}{\nabla} (a \overset{m+1}{\nabla} b) = a \overset{m+1}{\nabla} b+1.$$

KOROLLAR. *Die Funktion* $a \overset{m}{\nabla} b$ *ist in Bezug auf* a *für jedes* m *und* b *monoton wachsend im weiteren Sinne; d. h., man hat für* $c \geqq a$ *die Ungleichung* $c \overset{m}{\nabla} b \geqq a \overset{m}{\nabla} b$.

Endlich ist die Monotonität in Bezug auf m eine Folge von

HILFSSATZ 7. *Für* $a \geqq 2$ *und* $m \geqq 2$, *sowie auch für* $a \geqq 2$ *und* $b \geqq 2$, *gilt* $a \overset{m+1}{\nabla} b \geqq a \overset{m}{\nabla} b$.

BEWEIS durch vollständige Induktion nach m. Ist $m = 0$, so hat man $a+b \geqq b+1$, d. h. $a \overset{1}{\nabla} b \geqq a \overset{0}{\nabla} b$ für $a \geqq 1$. Ist $m = 1$, so hat man $ab \geqq a+b$, d. h. $a \overset{2}{\nabla} b \geqq a \overset{1}{\nabla} b$ für $a \geqq 2$ und $b \geqq 2$, da dann $ab-(a+b) = (a-1)(b-1)-1 \geqq (2-1)(2-1)-1 = 0$. Ist $m = 2$, so ist die Ungleichung $a^b \geqq ab$, d. h. $a \overset{3}{\nabla} b \geqq a \overset{2}{\nabla} b$, für $a \geqq 2$ offenbar richtig im Falle $b = 0$, da dann $a^b = 1$ und $ab = 0$; im Falle $b \geqq 1$ folgt aber aus Hilfssatz 4 wegen $a \geqq 2$

$$a^{b-1} = a \overset{3}{\nabla} b-1 > b-1,$$

d. h. $a^{b-1} \geqq b$, woraus sich durch Multiplikation mit a die zu beweisende Ungleichung $a^b \geqq ab$ ergibt. Gilt nun die Behauptung des Hilfssatzes für ein

$m \geqq 2$, so beweist man das gleiche für $m+1$, d. h. $a \overset{m+2}{\triangledown} b \geqq a \overset{m+1}{\triangledown} b$, durch vollständige Induktion nach b. In der Tat gilt für $b=0$ wegen $m \geqq 2$ und (3)

$$a \overset{m+2}{\triangledown} 0 = 1 = a \overset{m+1}{\triangledown} 0,$$

und falls die Behauptung $a \overset{m+2}{\triangledown} b \geqq a \overset{m+2}{\triangledown} b$ für einen Wert von b richtig ist, so gilt für $a \geqq 2$ wegen (3), wegen Hilfssatz 5, Korollar und wegen der Induktionsvoraussetzung $a \overset{m+1}{\triangledown} c \geqq a \overset{m}{\triangledown} c$

$$a \overset{m+2}{\triangledown} b+1 = a \overset{m+1}{\triangledown} (a \overset{m+2}{\triangledown} b) \geqq a \overset{m+1}{\triangledown} (a \overset{m+1}{\triangledown} b) \geqq a \overset{m}{\triangledown} (a \overset{m+1}{\triangledown} b) = a \overset{m+1}{\triangledown} b+1.$$

Damit ist Hilfssatz 7 bewiesen.

KOROLLAR. *Die Funktion* $a \overset{m}{\triangledown} b$ *ist in Bezug auf* m *für* $a \geqq 2$ *und* $b \geqq 2$ *monoton wachsend im weiteren Sinne; d. h., man hat für* $a \geqq 2$, $b \geqq 2$ *und beliebige* $m \geqq n$ *die Ungleichung* $a \overset{m}{\triangledown} b \geqq a \overset{n}{\triangledown} b$.

Aus den eben bewiesenen Monotonitätseigenschaften der Funktion $a \overset{m}{\triangledown} b$ folgen leicht auch die Ungleichungen, die in den folgenden drei Hilfssätzen formuliert werden.

HILFSSATZ 8. *Für* $a \geqq 3$ *und* $m \geqq 1$ *gilt* $a \overset{m}{\triangledown} a \geqq a+3$.

BEWEIS durch vollständige Induktion nach m. Für $m=1$ lautet die Behauptung $a+a \geqq a+3$, was für $a \geqq 3$ offenbar richtig ist. Nehmen wir an, die zu beweisende Ungleichung gilt für $m(\geqq 1)$ und beliebiges $a \geqq 3$; wir beweisen dann durch vollständige Induktion, daß das gleiche auch für $m+1$ besteht, d. h. $a \overset{m+1}{\triangledown} a \geqq a+3$ für $a \geqq 3$. In der Tat gilt dies für $a=3$. Es ist nämlich nach (3)

$$3 \overset{m+1}{\triangledown} 3 = 3 \overset{m}{\triangledown} (3 \overset{m+1}{\triangledown} 2).$$

Hier ist wegen (3) und Hilfssatz 2, der sich wegen $m+1 \geqq 2$ anwenden läßt, und wegen Hilfssatz 7, Korollar

$$3 \overset{m+1}{\triangledown} 2 = 3 \overset{m}{\triangledown} (3 \overset{m+1}{\triangledown} 1) = 3 \overset{m}{\triangledown} 3 \geqq 3 \overset{1}{\triangledown} 3 = 3+3.$$

Also gilt jedenfalls $3 \overset{m+1}{\triangledown} 2 \geqq 1$, so daß wegen Hilfssatz 4

$$3 \overset{m+1}{\triangledown} 3 = 3 \overset{m}{\triangledown} (3 \overset{m+1}{\triangledown} 2) \geqq 3 \overset{m+1}{\triangledown} 2 \geqq 3+3$$

gilt.

Falls nun unsere Behauptung $a \overset{m+1}{\triangledown} a \geqq a+3$ für einen Wert $a(\geqq 3)$ besteht, so gilt dasselbe auch für $a+1$, d. h. es gilt $a+1 \overset{m+1}{\triangledown} a+1 \geqq a+4$. In der Tat haben wir wegen (3)

$$a+1 \overset{m+1}{\triangledown} a+1 = a+1 \overset{m}{\triangledown} (a+1 \overset{m+1}{\triangledown} a).$$

Hier kann man Hilfssatz 4 wegen $a+1 \geqq 4$, also $a+1 \geqq 2$, und wegen $a+1 \overset{m+1}{\nabla} a \geqq a \overset{m+1}{\nabla} a \geqq a+3 \geqq 1$ (Hilfssatz 6 und Induktionsvoraussetzung) anwenden; man erhält

$$a+1 \overset{m+1}{\nabla} a+1 = a+1 \overset{m}{\nabla} (a+1 \overset{m+1}{\nabla} a) > a+1 \overset{m+1}{\nabla} a,$$

d. h.

$$a+1 \overset{m+1}{\nabla} a+1 \geqq (a+1 \overset{m+1}{\nabla} a)+1 \geqq (a+3)+1 = a+4,$$

wobei wir wiederum die Ungleichung $a+1 \overset{m+1}{\nabla} a \geqq a+3$ angewendet haben. Damit ist Hilfssatz 8 bewiesen.

Hilfssatz 9. *Für $a \geqq 3, b \geqq 2$ und $m \geqq 2$ gilt $a \overset{m}{\nabla} b \geqq b+4$.*

Beweis durch vollständige Induktion nach b. Für $b=2$ haben wir wegen (3), Hilfssatz 2 und 8, die sich wegen $m \geqq 2$ und $a \geqq 3$ anwenden lassen,

$$a \overset{m}{\nabla} 2 = a \overset{m-1}{\nabla} (a \overset{m}{\nabla} 1) = a \overset{m-1}{\nabla} a \geqq a+3 \geqq 6 = 2+4,$$

und falls die Behauptung für $b (\geqq 2)$ gültig ist, so hat man wegen (3) und Hilfssatz 5, Korollar

$$a \overset{m}{\nabla} b+1 = a \overset{m-1}{\nabla} (a \overset{m}{\nabla} b) \geqq a \overset{m-1}{\nabla} b+4 \geqq b+5,$$

da wegen Hilfssatz 4, der sich wegen $a \geqq 2$ und $b+4 \geqq 1$ anwenden läßt, $a \overset{m-1}{\nabla} b+4 > b+4$ gilt.

Hilfssatz 10. *Für $a \geqq 3$ und $m \geqq 1$ gilt $a \overset{m+1}{\nabla} 3 \geqq a \overset{m}{\nabla} a+3$.*

Beweis. Durch zweimalige Anwendung von (3) erhalten wir

$$a \overset{m+1}{\nabla} 3 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} 2) = a \overset{m}{\nabla} (a \overset{m}{\nabla} (a \overset{m+1}{\nabla} 1)).$$

Hier ist wegen $m+1 \geqq 2$ und Hilfssatz 2

$$a \overset{m+1}{\nabla} 1 = a;$$

also hat man wegen Hilfssatz 8 und Hilfssatz 5, Korollar

$$a \overset{m+1}{\nabla} 3 = a \overset{m}{\nabla} (a \overset{m}{\nabla} a) \geqq a \overset{m}{\nabla} a+3.$$

Zum Beweis des Hauptsatzes werden wir noch eine Eigenschaft der Funktion $a \overset{m}{\nabla} b$ benötigen, welche als eine Verallgemeinerung der beiden Identitäten über Potenzen

$$(a^b)^c = a^{bc}, \quad a^b a^c = a^{b+c}$$

angesehen werden können. Diese Eigenschaft wird ausgedrückt durch[11]

[11] Diese elegante Formulierung des Hilfssatzes verdanke ich L. Kalmár.

HILFSSATZ 11. *Für $a \geqq 1$ und $m \geqq 2$ haben wir*

$$(6)\qquad (a \overset{m}{\triangledown} b) \overset{m}{\triangledown} c \leqq a \overset{m}{\triangledown} bc$$

und

$$(7)\qquad (a \overset{m+1}{\triangledown} b) \overset{m}{\triangledown} (a \overset{m+1}{\triangledown} c) \leqq a \overset{m+1}{\triangledown} b+c.$$

BEWEIS. Für $m=2$ lauten die zu beweisenden Ungleichungen

$$(ab)c \leqq a(bc) \quad \text{und} \quad a^b a^c \leqq a^{b+c},$$

diese sind also richtig (in beiden gilt das Gleichheitszeichen). Nehmen wir an, daß (6) und (7) für einen Wert von $m(\geqq 2)$, für beliebige b, c und $a \geqq 1$ gültig sind. Dann beweisen wir dasselbe für $m+1$, d. h.

$$(8)\qquad (a \overset{m+1}{\triangledown} b) \overset{m+1}{\triangledown} c \leqq a \overset{m+1}{\triangledown} bc$$

und

$$(9)\qquad (a \overset{m+2}{\triangledown} b) \overset{m+1}{\triangledown} (a \overset{m+2}{\triangledown} c) \leqq a \overset{m+2}{\triangledown} b+c$$

ebenfalls für beliebige b, c und $a \geqq 1$.

Der Beweis von (8) geschieht durch vollständige Induktion in Bezug auf c. Für $c=0$ steht wegen $m \geqq 2$ und (3) auf beiden Seiten 1. Gilt (8) für einen Wert von c, so beweist man dasselbe für $c+1$ wie folgt. Nach (3) hat man

$$(10)\qquad (a \overset{m+1}{\triangledown} b) \overset{m+1}{\triangledown} c+1 = (a \overset{m+1}{\triangledown} b) \overset{m}{\triangledown} ((a \overset{m+1}{\triangledown} b) \overset{m+1}{\triangledown} c).$$

Hier ist wegen $a \geqq 1$ und Hilfssatz 6, Korollar, ferner $m+1 \geqq 3$ und Hilfssatz 1

$$a \overset{m+1}{\triangledown} b \geqq 1 \overset{m+1}{\triangledown} b = 1,$$

so daß wegen (8) und Hilfssatz 5, Korollar

$$(11)\qquad (a \overset{m+1}{\triangledown} b) \overset{m}{\triangledown} ((a \overset{m+1}{\triangledown} b) \overset{m+1}{\triangledown} c) \leqq (a \overset{m+1}{\triangledown} b) \overset{m}{\triangledown} (a \overset{m+1}{\triangledown} bc)$$

besteht. Durch Anwendung von (7) erhält man aber

$$(12)\qquad (a \overset{m+1}{\triangledown} b) \overset{m}{\triangledown} (a \overset{m+1}{\triangledown} bc) \leqq a \overset{m+1}{\triangledown} b+bc = a \overset{m+1}{\triangledown} b(c+1).$$

Aus (10), (11) und (12) gewinnt man

$$(a \overset{m+1}{\triangledown} b) \overset{m+1}{\triangledown} c+1 \leqq a \overset{m+1}{\triangledown} b(c+1),$$

d. h. (8) für $c+1$ statt c.

Den Beweis von (9) führen wir durch vollständige Induktion nach b. Für $b=0$ steht wegen (3) und Hilfssatz 1 (der sich wegen $m+1 \geqq 3$ anwenden läßt) auf der linken Seite

$$(a \overset{m+2}{\triangledown} 0) \overset{m+1}{\triangledown} (a \overset{m+2}{\triangledown} c) = 1 \overset{m+1}{\triangledown} (a \overset{m+2}{\triangledown} c) = 1,$$

auf der rechten Seite aber $a \overset{m+2}{\triangledown} c$, was wegen $a \geqq 1$ und Hilfssatz 6, Korollar, nicht kleiner als $1 \overset{m+2}{\triangledown} c$, d. h. nach Hilfssatz 1 nicht kleiner als 1 ist. Für $b=1$ steht auf der linken Seite wegen Hilfssatz 2

$$(a \overset{m+2}{\nabla} 1) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c) = a \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c),$$

was wegen (3) mit der rechten Seite $a \overset{m+2}{\nabla} 1 + c$ gleich ist. Nehmen wir nun an, daß (9) für einen Wert von $b (\geqq 1)$ gültig ist. Dann beweisen wir dasselbe für $b+1$ wie folgt. Wegen (3) und (8) (was unter Annahme der Gültigkeit von (6) und (7) bereits für alle Werte von b und c und für $a \geqq 1$ bewiesen wurde) haben wir

$$(13) \quad (a \overset{m+2}{\nabla} b+1) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c) = (a \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} b)) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c) \leqq a \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} b)(a \overset{m+2}{\nabla} c).$$

Hier ist

$$(14) \qquad (a \overset{m+2}{\nabla} b)(a \overset{m+2}{\nabla} c) \leqq (a \overset{m+2}{\nabla} b) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c).$$

In der Tat folgt diese Ungleichung für $a \geqq 2$ aus Hilfssatz 7, Korollar, da die linke Seite als $(a \overset{m+2}{\nabla} b) \overset{2}{\nabla} (a \overset{m+2}{\nabla} c)$ geschrieben werden kann und wegen Hilfssatz 6, Korollar nebst Hilfssatz 5, Korollar und Hilfssatz 2

$$a \overset{m+2}{\nabla} b \geqq 2 \overset{m+2}{\nabla} b \geqq 2 \overset{m+2}{\nabla} 1 = 2$$

gilt. Für $a = 1$ steht aber wegen Hilfssatz 1, aus dem $1 \overset{m+2}{\nabla} b = 1 \overset{m+2}{\nabla} c = 1$ und $1 \overset{m+1}{\nabla} 1 = 1$ folgt, auf beiden Seiten von (14) die Zahl 1. Aus (13), (14) und (9) folgt aber wegen Hilfssatz 5, Korollar

$$(a \overset{m+2}{\nabla} b+1) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c) \leqq a \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} b+c).$$

Auf der rechten Seite steht hier wegen (3) $a \overset{m+2}{\nabla} b+c+1$, daher gewinnen wir

$$(a \overset{m+2}{\nabla} b+1) \overset{m+1}{\nabla} (a \overset{m+2}{\nabla} c) \leqq a \overset{m+2}{\nabla} (b+1)+c,$$

d. h. (9) mit $b+1$ statt b. Damit ist Lemma 11 vollständig bewiesen.

Um die Sätze I und II aus dem in **3** angeführten Satze von R. PÉTER herleiten zu können, beweise ich zunächst einen Hilfssatz über den in **3** erwähnten Zusammenhang zwischen den Funktionen $\xi(m, n)$ und $a \overset{m}{\nabla} b$.

HILFSSATZ 12. *Für die durch die Gleichungen* (4) *definierte Funktion* $\xi(m, n)$ *gilt*

$$(15) \qquad \xi(m, n) + 3 = 2 \overset{m}{\nabla} n + 3,$$

d. h.

$$\xi(m, n) = (2 \overset{m}{\nabla} n + 3) - 3.$$

BEWEIS durch vollständige Induktion nach m. Für $m = 0$ haben wir nach (4) bzw. (3)

$$\xi(0, n) + 3 = (n+1) + 3 = n + 4$$

und

$$2 \overset{0}{\triangledown} n+3=(n+3)+1=n+4.$$

Angenommen, daß die zu beweisende Identität für m gültig ist, beweisen wir dasselbe für $m+1$, d. h.

(16) $$\xi(m+1, n)+3=2 \overset{m+1}{\triangledown} n+3,$$

durch Induktion nach n. Für $n=0$ lautet die linke Seite von (16) wegen (4)

$$\xi(m+1, 0)+3=\xi(m, 1)+3.$$

Aus der Induktionsvoraussetzung (15) folgt aber für $n=1$

$$\xi(m, 1)+3=2 \overset{m}{\triangledown} 4.$$

Andererseits steht auf der rechten Seite von (16) wegen (3) und Hilfssatz 3

$$2 \overset{m+1}{\triangledown} 3=2 \overset{m}{\triangledown}(2 \overset{m+1}{\triangledown} 2)=2 \overset{m+1}{\triangledown} 4.$$

Angenommen, daß die Identität (16) für einen Wert von n richtig ist, erhalten wir aus (4), aus der Induktionsvoraussetzung (15) (mit $\xi(m+1, n)$ statt n), ferner aus (16) und (3)

$$\xi(m+1, n+1)+3=\xi(m, \xi(m+1, n))+3=2 \overset{m}{\triangledown} \xi(m+1, n)+3=$$
$$=2 \overset{m}{\triangledown}(2 \overset{m+1}{\triangledown} n+3)=2 \overset{m+1}{\triangledown} n+4,$$

d. h. (16) mit $n+1$ statt n.

Aus Hilfssatz 12 folgt sofort die Ungleichung

(17) $$\xi(m, n)<2 \overset{m}{\triangledown} n+3.$$

Eine analoge Ungleichung für $\xi(m, n)$ wird ausgedrückt durch

HILFSSATZ 13. *Wir haben*

(18) $$\xi(m, n)<\max(3, n) \overset{m+1}{\triangledown} 3.$$

BEWEIS. Für $m=0$ steht nach (4) auf der linken Seite $\xi(0, n)=n+1$ und auf der rechten Seite $\max(3, n)+3$, was offenbar nicht kleiner als $n+3$, also größer als $n+1$ ist. Ist aber $m \geqq 1$, so folgt aus (17) und aus Hilfssatz 10

$$\xi(m, n)<2 \overset{m}{\triangledown} n+3 \leqq 2 \overset{m}{\triangledown} \max(3, n)+3 \leqq \max(3, n) \overset{m}{\triangledown} \max(3, n)+3 \leqq$$
$$\leqq \max(3, n) \overset{m+1}{\triangledown} 3.$$

6. Nun können wir die Sätze I und II beweisen.

BEWEIS DES SATZES I. Es sei $\alpha(x, y, \ldots)$ eine beliebige rekursive Funktion. Nach dem angeführten Satze von R. PÉTER gibt es eine nichtnegative ganze Zahl m, so daß

$$\alpha(x, y, \ldots)<\xi(m, \max(x, y, \ldots))$$

für alle Werte von $x, y, \ldots$; andererseits gilt nach Hilfssatz 13

$$\xi(m, \max(x, y, \ldots))<\max(3, \max(x, y, \ldots)) \overset{m+1}{\triangledown} 3=\max(3, x, y, \ldots) \overset{m+1}{\triangledown} 3.$$

Aus den beiden letzten Ungleichungen folgt die Behauptung des Satzes I mit $p = m+1$ und $k = 3$.

Beweis des Satzes II. Es sei $\beta(x, y)$ eine beliebige zweistellige rekursive Funktion. Nach dem oben angeführten Satze von R. Péter gibt es eine nichtnegative ganze Zahl m, so daß

$$\beta(x, y) < \xi(m, \max(x, y))$$

für alle Werte x und y. Für $x \geqq y$ gilt also nach (18)

$$\beta(x, y) < \xi(m, x) < \max(3, x) \overset{m+1}{\bigtriangledown} 3 \leqq \max(3, x) \overset{m+1}{\bigtriangledown} \max(3, y),$$

also gilt dann die Behauptung des Satzes II mit $q = m+1$.

Ist aber $x < y$, so hat man wegen (17)

$$\beta(x, y) < \xi(m, y) < 2 \overset{m}{\bigtriangledown} y + 3 \leqq 2 \overset{m}{\bigtriangledown} \max(3, y) + 3 \leqq \max(3, x) \overset{m}{\bigtriangledown} \max(3, y) + 3.$$

Ist hier $m \geqq 1$, so haben wir nach Hilfssatz 9

$$\max(3, y) + 3 \leqq \max(3, x) \overset{m+1}{\bigtriangledown} \max(3, y) - 1,$$

daher, nach (3),

$$\beta(x, y) < \max(3, x) \overset{m}{\bigtriangledown} (\max(3, x) \overset{m+1}{\bigtriangledown} \max(3, y) - 1) = \max(3, x) \overset{m+1}{\bigtriangledown} \max(3, y),$$

so daß die Behauptung des Satzes II wiederum mit $q = m+1$ gültig ist. Ist aber zufälligerweise $m = 0$, so hat man wegen (4)

$$\beta(x, y) < \xi(0, y) = y + 1 < 3 + y \leqq \max(3, x) + \max(3, y) =$$
$$= \max(3, x) \overset{1}{\bigtriangledown} \max(3, y),$$

so daß auch der Fall $m = 0$ keine Ausnahme bildet. Damit ist also Satz II allgemein bewiesen.

7. Nun werden wir den oben formulierten Hauptsatz beweisen. Wegen der Definition der Funktionenklasse $E(K, I)$ genügt es (mit $m = \max(p, q+2)$) folgendes zu beweisen.

1) Für jedes $\alpha \in K$ gilt für geeignetes k und für alle Werte der Argumente $x, y, \ldots$ von α die Ungleichung

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{m}{\bigtriangledown} k. \tag{19}$$

2) Gibt es zur r-stelligen arithmetischen Funktion φ_0 und zu den s-stelligen arithmetischen Funktionen $\varphi_1, \ldots, \varphi_r$ nichtnegative ganze Zahlen $k_0, k_1, \ldots, k_r$, so daß für alle Werte der Argumente

$$\varphi_0(x_1, \ldots, x_r) < \max(3, x_1, \ldots, x_r) \overset{m}{\bigtriangledown} k_0 \tag{20}$$

und für $i = 1, \ldots, r$

$$\varphi_i(x_1, \ldots, x_s) < \max(3, x_1, \ldots, x_s) \overset{m}{\bigtriangledown} k_i, \tag{21}$$

so gilt für geeignetes k und für alle Werte der Argumente die Ungleichung

$$(22)\qquad \varphi_0(\varphi_1(x_1,\ldots,x_s),\ldots,\varphi_r(x_1,\ldots,x_s)) < \max(3,x_1,\ldots,x_s)\overset{m}{\nabla} k.$$

3) Gibt es zur $r+1$-stelligen arithmetischen Funktion φ eine nichtnegative ganze Zahl l, so daß für alle Werte der Argumente

$$(23)\qquad \varphi(x_0,x_1,\ldots,x_r) < \max(3,x_0,x_1,\ldots,x_r)\overset{m}{\nabla} l,$$

so gilt für die durch die Rekursion (5) definierte Funktion ψ für geeignetes k und für alle Werte der Argumente die Ungleichung

$$(24)\qquad \psi(x_0,x_1,\ldots,x_r) < \max(3,x_0,x_1,\ldots,x_r)\overset{m}{\nabla} k.$$

Beweis von 1. Laut Voraussetzung des Hauptsatzes gilt für jedes $\alpha\in K$ für geeignetes k und für alle Werte der Argumente

$$(25)\qquad \alpha(x,y,\ldots) < \max(3,x,y,\ldots)\overset{p}{\nabla} k.$$

Hier kann wegen der Monotonität der Funktion $a\overset{m}{\nabla} b$ in Bezug auf b (für $a\geqq 1$) die Zahl k nötigenfalls vergrößert werden, so daß wir $k\geqq 2$ voraussetzen können. Dann folgt aber (19) aus (25) wegen $m\geqq p$ auf Grund der Monotonität der Funktion $a\overset{m}{\nabla} b$ in Bezug auf m (für $a\geqq 2$ und $b\geqq 2$).

Beweis von 2. Aus (20) gewinnen wir

$$(26)\qquad \begin{aligned}&\varphi_0(\varphi_1(x_1,\ldots,x_s),\ldots,\varphi_r(x_1,\ldots,x_s)) <\\ &< \max(3,\varphi_1(x_1,\ldots,x_s),\ldots,\varphi_r(x_1,\ldots,x_s))\overset{m}{\nabla} k_0.\end{aligned}$$

Wegen (21) haben wir für $i=1,\ldots,r$

$$\varphi_i(x_1,\ldots,x_s) < \max(3,x_1,\ldots,x_s)\overset{m}{\nabla} k_i \leqq \max(3,x_1,\ldots,x_s)\overset{m}{\nabla}\max(1,k_1,\ldots,k_r)$$

und wegen $m\geqq q+2\geqq 2$ und Hilfssatz 2

$$3=3\overset{m}{\nabla}1\leqq \max(3,x_1,\ldots,x_s)\overset{m}{\nabla}1\leqq\max(3,x_1,\ldots,x_s)\overset{m}{\nabla}\max(1,k_1,\ldots,k_r),$$

so daß

$$\begin{aligned}&\max(3,\varphi_1(x_1,\ldots,x_s),\ldots,\varphi_r(x_1,\ldots,x_s))\leqq\\ &\leqq \max(3,x_1,\ldots,x_s)\overset{m}{\nabla}\max(1,k_1,\ldots,k_r).\end{aligned}$$

Also folgt aus (26) und Hilfssatz 11

$$\begin{aligned}&\varphi_0(\varphi_1(x_1,\ldots,x_s),\ldots,\varphi_r(x_1,\ldots,x_s)) <\\ &< (\max(3,x_1,\ldots,x_s)\overset{m}{\nabla}\max(1,k_1,\ldots,k_r))\overset{m}{\nabla} k_0\leqq\\ &\leqq \max(3,x_1,\ldots,x_s)\overset{m}{\nabla} k_0\max(1,k_1,\ldots,k_r),\end{aligned}$$

d. h. die Ungleichung (22) mit $k=k_0\max(1,k_1,\ldots,k_r)$.

Beweis von 3. Wegen der Monotonität der Funktion $a\overset{m}{\nabla} b$ in Bezug auf b können wir voraussetzen, daß $l\geqq 1$. Wir zeigen zunächst die Gültigkeit der Ungleichung

$$(27)\qquad \psi(n, x_1, \ldots, x_r) \leqq (\max(3, n, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+1$$

durch vollständige Induktion nach n. Ist $n=0$, so folgt aus (5) und (23) wegen Hilfssatz 2 (im Falle $q \geqq 1$) bzw. $a \overset{1}{\triangledown} 1 = a+1 > a$ (im Falle $q=0$)

$$\psi(0, x_1, \ldots, x_r) = \varphi(0, x_1, \ldots, x_r) < \max(3, 0, x_1, \ldots, x_r) \overset{m}{\triangledown} l \leqq$$
$$\leqq (\max(3, 0, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} 1.$$

Gilt die Ungleichung (27) für ein n, so haben wir wegen (5) und der Voraussetzung

$$\beta(x, y) \leqq \max(3, x) \overset{q}{\triangledown} \max(3, y)$$

des Hauptsatzes

$$(28)\qquad \begin{aligned} \psi(n+1, x_1, \ldots, x_r) &= \beta(\varphi(n+1, x_1, \ldots, x_r), \psi(n, x_1, \ldots, x_r)) \leqq \\ &\leqq \max(3, \varphi(n+1, x_1, \ldots, x_r)) \overset{q}{\triangledown} \max(3, \psi(n, x_1, \ldots, x_r)). \end{aligned}$$

Hier haben wir wegen (23)

$$\varphi(n+1, x_1, \ldots, x_r) < \max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l$$

und wegen (27) und der Monotonität der Funktion $a \overset{m}{\triangledown} b$ in Bezug auf a

$$\psi(n, x_1, \ldots, x_r) \leqq (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+1,$$

ferner wegen $m \geqq 2$, Hilfssatz 2 und $l \geqq 1$

$$\begin{aligned} 3 &\leqq \max(3, n+1, x_1, \ldots, x_r) = \max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} 1 \leqq \\ &\leqq \max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l \leqq (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} 1 \leqq \\ &\leqq (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+1, \end{aligned}$$

so daß

$$\max(3, \varphi(n+1, x_1, \ldots, x_r)) \leqq \max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l$$

und

$$\max(3, \psi(n, x_1, \ldots, x_r)) \leqq (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+1.$$

Daher gewinnen wir aus (28) und (3) die Ungleichung

$$\begin{aligned} &\psi(n+1, x_1, \ldots, x_r) \leqq \\ &\leqq (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q}{\triangledown} ((\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+1) = \\ &= (\max(3, n+1, x_1, \ldots, x_r) \overset{m}{\triangledown} l) \overset{q+1}{\triangledown} n+2, \end{aligned}$$

d. h. (27) für $n+1$ statt n. Nun kann man aus (27) wegen

$$n+1 < n+2 = n \overset{1}{\triangledown} 2 \leqq \max(3, n, x_1, \ldots, x_r) \overset{1}{\triangledown} 2 \leqq \max(3, n, x_1, \ldots, x_r) \overset{m}{\triangledown} 2$$

und $q+1 \leqq m-1$ folgendermaßen weiterschließen:

$$\psi(n, x_1, \ldots, x_r) < (\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} l) \overset{q+1}{\nabla} (\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} 2) \leqq$$
$$\leqq (\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} l) \overset{m-1}{\nabla} (\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} 2),$$

da, wie gezeigt,

$$\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} l \geqq 3 \geqq 2$$

und

$$\max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} 2 \geqq n+2 \geqq 2.$$

Wegen Hilfssatz 11 (für $m \geqq 3$) bzw. $ab+ac=a(b+c)$ (für $m=2$) gewinnen wir also

$$\psi(n, x_1, \ldots, x_r) < \max(3, n, x_1, \ldots, x_r) \overset{m}{\nabla} l+2,$$

d. h. Ungleichung (24) mit $k=l+2$. Damit ist der Hauptsatz vollständig bewiesen.

8. Der Hauptsatz liefert eine Lösung des in **1** formulierten Markovschen Problems. In der Tat folgt aus dem Hauptsatz unter der Voraussetzung, daß die Funktionenklassen K und I die Bedingungen des Hauptsatzes erfüllen, daß die Funktionenklasse $E(K, I)$ nicht alle rekursiven Funktionen enthalten kann. Sie enthält nämlich nicht die Funktion $a \overset{m}{\nabla} a$ von a, wobei m die im Hauptsatze figurierende feste nichtnegative ganze Zahl (nämlich $\max(p, q+2)$) bedeutet. Sonst würde nämlich mit geeignetem k die Ungleichung

$$a \overset{m}{\nabla} a < \max(3, a) \overset{m}{\nabla} k \tag{29}$$

für alle Werte von a erfüllt sein. Hier kann man wegen der Monotonität der Funktion $a \overset{m}{\nabla} b$ in Bezug auf b (für $a \geqq 1$) die Zahl k beliebig vergrößern, so daß wir jedenfalls $k \geqq 3$ voraussetzen dürfen. Nun würde aber aus (29) für $a=k$ die Ungleichung

$$k \overset{m}{\nabla} k < \max(3, k) \overset{m}{\nabla} k$$

folgen, was wegen $\max(3, k)=k$ unmöglich ist.

Andererseits ist aber die Funktion $a \overset{m}{\nabla} a$ für jedes feste m eine rekursive Funktion von a, da sie aus der rekursiven Funktion $a \overset{m}{\nabla} b$ von a und b durch eine Substitution (von a für b) entsteht.

Im Falle, daß die Funktionenklassen K und I aus endlich vielen rekursiven Funktionen bestehen, d. h. die Rolle der Addition, der arithmetischen Subtraktion, der Multiplikation und der arithmetischen Division in der Definition der elementaren Funktionen durch eine endliche Anzahl von rekursiven Funktionen, und auch die Rolle der Summen- und Produktenbildung durch eine endliche Anzahl von Funktionenoperationen übernommen wird, die durch eine Rekursion von der Form (5) aus zweistelligen rekursiven Funktionen β definiert werden, sind die Bedingungen des Hauptsatzes erfüllt. In der Tat gibt es dann wegen Satz I und II zu jedem $\alpha \in K$ nichtnegative ganze Zahlen p und k und zu jedem $\beta \in I$ eine nichtnegative ganze Zahl q, so daß für alle Werte der Argumente die Ungleichungen

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{p}{\triangledown} k$$

$$\beta(x, y) \leqq \max(3, x) \overset{q}{\triangledown} \max(3, y)$$

bestehen. Hier kann wiederum aus Monotonitätsgründen $k \geqq 2$ vorausgesetzt werden, und die Zahlen p und q können, wegen der Monotonität von $a \overset{m}{\triangledown} b$ als Funktion von a und m für $a \geqq 2$ und $b \geqq 2$, beliebig vergrößert werden. Daher kann man für die endlich vielen $\alpha \in K$ ein gemeinsames p und für die endlich vielen $\beta \in I$ ein gemeinsames q wählen. Also kann die Funktionenklasse $E(K, I)$ für endliche Klassen rekursiver Funktionen K und I nicht die Klasse aller rekursiven Funktionen erschöpfen, in Übereinstimmung mit der Vermutung von MARKOV.

Besteht die Klasse K aus den Funktionen $x+y$, $|x-y|$, $x \cdot y$, $[x/y]$ und I aus den Funktionen $x+y$ und xy, so ist $E(K, I)$ die Klasse der elementaren Funktionen. In diesem Falle kann man $p=3$ und $q=2$ wählen. In der Tat gilt

$$x+y \leqq 2\max(x, y) \leqq 2\max(3, x, y) < 3\max(3, x, y) \leqq$$

$$\leqq \max(3, x, y)\max(3, x, y) = (\max(3, x, y))^2 = \max(3, x, y) \overset{3}{\triangledown} 2,$$

$$|x-y| \leqq \max(x, y) \leqq \max(3, x, y) < (\max(3, x, y))^2 = \max(3, x, y) \overset{3}{\triangledown} 2,$$

$$xy \leqq (\max(x, y))^2 \leqq (\max(3, x, y))^2 < (\max(3, x, y))^3 = \max(3, x, y) \overset{3}{\triangledown} 3,$$

$$[x/y] \leqq x \leqq \max(3, x, y) < (\max(3, x, y))^2 = \max(3, x, y) \overset{3}{\triangledown} 2$$

und

$$x+y \leqq \max(3, x) + \max(3, y) = \max(3, x) \overset{1}{\triangledown} \max(3, y) \leqq \max(3, x) \overset{2}{\triangledown} \max(3, y),$$

$$xy \leqq \max(3, x)\max(3, y) = \max(3, x) \overset{2}{\triangledown} \max(3, y).$$

Daher gilt in diesem Falle der Hauptsatz mit $m = \max(3, 4) = 4$, also gehört nach der obigen Bemerkung die Funktion

$$a \overset{4}{\triangledown} a = a^{\left(a^{\cdot^{\cdot^{\cdot^{\left(a^{a}\right)}}}}\right)}$$

(mit einer Anzahl a von Zahlen a auf der rechten Seite) nicht zur Klasse der elementaren Funktionen. Dies ist das Ergebnis meiner Arbeit a. a. O.[1]; dieses Ergebnis ist also ebenfalls im Hauptsatze enthalten.

Beim Egerváryschen Problem handelt es sich um die Funktionenklasse $E(K, I)$, wobei die Klasse K wiederum aus den Funktionen $x+y$, $|x-y|$, $x \cdot y$, $[x/y]$, die Klasse I aber aus der Funktion $x+y$ allein besteht; und die Vermutung von EGERVÁRY besagt, daß die Funktion a^b nicht dieser Funktionenklasse $E(K, I)$ angehört. In diesem Falle kann man nach dem obigen $p = 3$ und $q = 1$ wählen. Der Hauptsatz gilt also in diesem Falle mit $m = \max(3, 3) = 3$; nach der obigen Bemerkung gehört also die Funktion $a \overset{3}{\nabla} a = a^a$ nicht zur betrachteten Funktionenklasse $E(K, I)$; daher auch, in Übereinstimmung mit der Vermutung von EGERVÁRY, die Funktion a^b nicht, da die Funktion a^a aus derselben durch eine Substitution ensteht. Also enthält der Hauptsatz auch die Lösung des Egerváryschen Problems.

Endlich folgt aus dem Hauptsatz die Nichtrekursivität der Funktion $a \overset{m}{\nabla} b$ als Funktion von a, m und b. In der Tat, betrachten wir die Klasse $E(K, I)$, wobei K aus der einzigen Funktion $x \overset{y}{\nabla} z$ besteht, I aber die leere Menge bedeutet. Wäre nun die Funktion $x \overset{y}{\nabla} z$ rekursiv, so gäbe es wegen Satz I nichtnegative ganze Zahlen p und k, für die und für alle Werte der Argumente x, y, z die Ungleichung

$$x \overset{y}{\nabla} z < \max(3, x, y, z) \overset{p}{\nabla} k$$

besteht. Also würden die Funktionenklassen K und I die Bedingungen des Hauptsatzes (mit diesem p und mit beliebigem q, etwa mit $q = 1$) erfüllen. Nach der obigen Bemerkung gibt es also eine nichtnegative ganze Zahl m, so daß die Funktion $a \overset{m}{\nabla} a$ von a nicht zur betrachteten Funktionenklasse $E(K, I)$ gehört. Dies ist aber unmöglich, da wegen $K = E(K, I)$ die Funktion $x \overset{y}{\nabla} z$ zur Funktionenklasse $E(K, I)$ gehört, und die Funktion $a \overset{m}{\nabla} a$ daraus durch Substitution entsteht.

(Eingegangen am 1. September 1952.)

РЕШЕНИЕ ОДНОЙ ПРОБЛЕМЫ А. А. МАРКОВА, СВЯЗАННОЙ С РАСПРОСТРАНЕНИЕМ ПОНЯТИЯ ЭЛЕМЕНТАРНОЙ ФУНКЦИИ

И. БЭРЕЦКИ (Сегед)

(Резюме)

Настоящая статья дает решение проблемы ставленной А. А. Марковым на Первом Венгерском Математическом Конгрессе в связи с докладом автора. Этот доклад был посвящен выяснению вопроса о классе элементарных функций и о классе (примитивных) рекурзивных функций. Под элементарными функциями следует понимать арифметические функции, составленные из независимых переменных и из целых положительных постоянных с помощью конечного числа сложений, арифметических вычитаний, умножений, арифметических делений, и составлений сумм и продуктов. Мы назовём здесь арифметической функцией функцию, независимые переменные которой пробегают неотрицательные целые числа, и значения которой суть также неотрицательные целые числа; арифметическое вычитание есть операция $|x-y|$, а арифметическое деление операция $[x/y]$. Под рекурсивными функциями мы понимаем теоретико-числовые функции, составленные из постоянной 0 и из функции $x+1$ конечным числом подстановок (образования функции от функций) и рекурсий формы

$$\begin{cases} \varphi(0, x, y, \ldots) = \alpha(x, y, \ldots) \\ \varphi(n+1, x, y, \ldots) = \beta(n, x, y, \ldots, \varphi(n, x, y, \ldots)). \end{cases}$$

Как было уже давно узвестно, всякая элементарная функция является рекурсивной. На Первом Венгерском Математическом Конгрессе автор показала, что обратное утверждение неверно, ведь например функция получающаяся из возведения в степень такой же итерацией какой это последнее получается из умножения или умножение из сложения, есть функция очевидно рекурсивна, но не элементарна. В связи с этим Марков ставил следующий вопрос:

Пусть определение элементарной функции подвергается следующему модификацию: вместо сложения, абсолютного значения разности, умножения и целой части частного допускается применение конечного числа любых рекурсивных функций, а вместо составления сумм и продуктов применение конечного числа функциональных операций, построенных из некоторых любых рекурсивных функций так как сумма строится из сложения а произведение из умножения. Верно-ли, что класс рекурсивных функций не исчерпается элементарными функциями даже если исходить из упомянутого расширенного понимания элементарности функций?

Одновременно с Марковым Эгервари выразил мнение, что если из определения элементарных функций исключить образование произведений, то получающийся класс функций не содержит даже возведение в степень.

В настоящей работе доказывается общая теорема из которой следуют предположения Маркова и Эгервари. Определим функцию $a \overset{m}{\nabla} b$ следующей двойной рекурсией:

$$a \overset{0}{\nabla} b = b+1,$$

$$a \overset{m-1}{\nabla} 0 = \begin{cases} a & \text{при} \quad m=0, \\ 0 & \text{при} \quad m=1, \\ 1 & \text{при} \quad m \geqq 2, \end{cases}$$

$$a \overset{m+1}{\nabla} b+1 = a \overset{m}{\nabla} (a \overset{m+1}{\nabla} b).$$

Опираясь на результат Р. Петера, являющийся обобщением одной теоремы Аккермана, можно доказать следующие две теоремы.

(1) Для любой рекурсивной функции α (с произвольным числом переменных) существуют такие p и k, что

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{p}{\nabla} k$$

для всех значений $x, y, \ldots$

(2) Для всякой рекурсивной функции двух переменных β существует такой q, что для всех значений x и y

$$\beta(x, y) < \max(3, x) \overset{q}{\nabla} \max(3, y).$$

Пусть теперь K и I две классы арифметических функций, причем элементы I суть функции двух переменных. Обозначим через $E(K, I)$ самый узкий класс функций, обладающий следующими тремя свойствами:

1. $K \subseteq E$ (если $\alpha \in K$, то $\alpha \in E$).

2. Если функция ψ r переменных и функции $\varphi_1, \ldots, \varphi_r$ s переменных все принадлежат E, то и

$$\psi(\varphi_1(x_1, \ldots, x_s), \ldots, \varphi_r(x_1, \ldots, x_s)) \in E.$$

3. Если функция φ любого числа (например $r+1$) переме ных принадлежит E и β принадлежит I, то функция ψ, определенная рекурсией

$$\psi(0, x_1 \ldots, x_r) = \varphi(0, x_1, \ldots, x_r)$$

$$\psi(n+1, x_1, \ldots, x_r) = \beta(\varphi(n+1, x_1, \ldots, x_r), \psi(n, x_1, \ldots, x_r))$$

также принадлежит E. Нашу общую теорему можно теперь сформулировать так:

Если существует такой p, что для любой функции $\alpha \in K$ существует такой k, что

$$\alpha(x, y, \ldots) < \max(3, x, y, \ldots) \overset{p}{\nabla} k$$

и если есть такой q, что для любой функции $\beta \in I$

$$\beta(x, y) < \max(3, x) \overset{q}{\nabla} \max(3, y),$$

то существует такой m (а именно $m = \max(p, q+2)$) и для любого элемента ψ множества функцнй $E(K, I)$ существует такой k, что

$$\psi(x_1, \ldots, x_r) < \max(3, x_1, \ldots, x_r) \overset{m}{\nabla} k.$$

Теорема, доказана автором на Первом Венгерском Математическом Конгрессе, и решения проблем Маркова и Эгервари являются частными случаями этой более общей теоремы.

ON COMPOSED POISSON DISTRIBUTIONS, III

By

J. ACZÉL (Debrecen)

(*Presented by* A. RÉNYI)

1. The present paper is a continuation of the paper of L. JÁNOSSY, A. RÉNYI and the author [1] and of the paper of A. RÉNYI [2]. By the methods of §§ 1—2 of the first paper we prove a theorem which continues the way opened by Theorem 1 of the second paper by leaving aside the "rarity" postulate **B** needed there. Thus we get the most general form of inhomogeneous stochastic processes of random events (integer-valued Markoff-processes).

The proof is based on the solution of a system of functional equations (2) where the number of equations is infinite. There might be some interest in the quite elementary character of the proof which avoids also the use of generating functions.

2. Our proof makes use of the well-known fact that the functional equation

$$f(t_1, t_3) = f(t_1, t_2) + f(t_2, t_3)$$

(additive interval-functions) has the general solution

$$f(t_1, t_2) = L(t_2) - L(t_1).$$

As a matter of fact, putting

$$L(t) = f(t_0, t)$$

(t_0 constant),

$$f(t_1, t_2) = f(t_0, t_2) - f(t_0, t_1) = L(t_2) - L(t_1)$$

follows immediately from the functional equation and, on the other hand, the function

$$f(t_1, t_2) = L(t_2) - L(t_1)$$

with an arbitrary L evidently satisfies the equation

$$f(t_1, t_3) = f(t_1, t_2) + f(t_2, t_3).$$

This functional equation is the inhomogeneous analogue of CAUCHY's equation.

3. We prove the

THEOREM. *Let the number of random events occurring during the time-interval* (t_1, t_2) *be independent of the number of events occurring during the time-interval* (t_3, t_4) *provided that* $t_1 < t_2 \leqq t_3 < t_4$, *and let* $w_k(t_1, t_2)$ *denote the probability of exactly k events occurring in the time-interval* (t_1, t_2). *Then*

$$w_0(t_1, t_2) = e^{L(t_2)-L(t_1)},$$

(1) $$w_k(t_1, t_2) = e^{L(t_2)-L(t_1)} \sum_{\substack{r_1+2r_2+\ldots+kr_k=k \\ r_j \geqq 0}} \prod_{j=1}^{k} \frac{[C_j(t_2)-C_j(t_1)]^{r_j}}{r_j!}$$

$$[0 \leqq w_k(t_1, t_2) \leqq 1]$$

where $C_j(t)$ $(j = 1, 2, \ldots)$ *are arbitrary functions such that* $\sum_{j=1}^{\infty} [C_j(t_2) - C_j(t_1)]$ *exists and is equal to* $L(t_1) - L(t_2)$. *(General inhomogeneous composed Poisson distribution.)*

It might be remarked that the formula obtained in Theorem 1 of [2] is somewhat more special than (1).

PROOF. Evidently we have

(2) $$w_k(t_1, t_3) = \sum_{j=0}^{k} w_{k-j}(t_1, t_2)\, w_j(t_2, t_3), \qquad (k = 0, 1, 2, 3, \ldots;\ t_1 \leqq t_2 \leqq t_3)$$

because if in the time-interval (t_1, t_3) there occur exactly k events, then either k events occur in the time-interval (t_1, t_2) and no event in (t_2, t_3), or $k-1$ events in (t_1, t_2) and one in (t_2, t_3), or ..., or else no event in (t_1, t_2) and k events in (t_2, t_3). Also the formula

(3) $$\sum_{k=0}^{\infty} w_k(t_1, t_2) = 1$$

is quite evident.

We shall prove that the general solution of the system of functional equations (2) and (3) is the system of functions (1).

It would be enough to prove (1) for $k = 0$ and then to apply induction. But in order to show how the rather intricate formula (1) arises, and to facilitate the induction, we shall consider in details the cases $k = 0, 1, 2, 3$, respectively.

For $k = 0$, (2) gives

(2_0) $$w_0(t_1, t_3) = w_0(t_1, t_2)\, w_0(t_2, t_3),$$

This equation shows that if $w_0(t_1, t_2) = 0$ for some interval (t_1, t_2), then $w_0(t_1, t_2) \equiv 0$ and hence by (2) we have also $w_k(t_1, t_2) \equiv 0$ $(k = 0, 1, 2, \ldots)$, but this would be a contradiction to (3). Therefore we have $w_0(t_1, t_2) \neq 0$ for all (t_1, t_2) and hence

$$\log w_0(t_1, t_3) = \log w_0(t_1, t_2) + \log w_0(t_2, t_3).$$

Thus by **2**

$$(1_0)\qquad \log w_0(t_1, t_2) = L(t_2) - L(t_1),$$
$$w_0(t_1, t_2) = e^{L(t_2) - L(t_1)}.$$

[$L(t)$ must be a decreasing function of t, as $w_0(t_1, t_2) \leqq 1$.]

(2_0) can be solved also directly by a method similar to that of **2**.

For $k=1$ we have

$$w_1(t_1, t_3) = w_1(t_1, t_2)\, w_0(t_2, t_3) + w_0(t_1, t_2)\, w_1(t_2, t_3),$$

or dividing it by (2_0)

$$\frac{w_1(t_1, t_3)}{w_0(t_1, t_3)} = \frac{w_1(t_1, t_2)}{w_0(t_1, t_2)} + \frac{w_1(t_2, t_3)}{w_0(t_2, t_3)}.$$

Thus by **2** and (1_0) we obtain

$$\frac{w_1(t_1, t_2)}{w_0(t_1, t_2)} = C_1(t_2) - C_1(t_1),$$

$$(1_1)\qquad w_1(t_1, t_2) = e^{L(t_2) - L(t_1)}.[C_1(t_2) - C_1(t_1)].$$

[$C_1(t)$ must be an increasing function, as $w_1(t_1, t_2) \geqq 0$.]

For $k=2$, (2) gives

$$w_2(t_1, t_3) = w_2(t_1, t_2)\, w_0(t_2, t_3) + w_1(t_1, t_2)\, w_1(t_2, t_3) + w_0(t_1, t_2)\, w_2(t_2, t_3),$$

or if we divide it again by (2_0) and take (1_1), (1_0) into account, we get

$$\frac{w_2(t_1, t_3)}{w_0(t_1, t_3)} = \frac{w_2(t_1, t_2)}{w_0(t_1, t_2)} + [C_1(t_2) - C_1(t_1)]\ [C_1(t_3) - C_1(t_2)] + \frac{w_2(t_2, t_3)}{w_0(t_2, t_3)},$$

that is,

$$\frac{w_2(t_1, t_3)}{w_0(t_1, t_3)} - \frac{[C_1(t_2) - C_1(t_1)]^2}{2} - \frac{2[C_1(t_2) - C_1(t_1)]\ [C_1(t_3) - C_1(t_2)]}{2} -$$
$$- \frac{[C_1(t_3) - C_1(t_2)]^2}{2} = \frac{w_2(t_1, t_2)}{w_0(t_1, t_2)} - \frac{[C_1(t_2) - C_1(t_1)]^2}{2} + \frac{w_2(t_2, t_3)}{w_0(t_2, t_3)} -$$
$$- \frac{[C_1(t_3) - C_1(t_2)]^2}{2}.$$

Let us put

$$f(t_1, t_2) = \frac{w_2(t_1, t_2)}{w_0(t_1, t_2)} - \frac{[C_1(t_2) - C_1(t_1)]^2}{2},$$

then the last equation may be written in the form

$$f(t_1, t_3) = f(t_1, t_2) + f(t_2, t_3) \quad \text{whence} \quad f(t_1, t_2) = C_2(t_2) - C_2(t_1).$$

Hence we find

$$(1_2)\qquad w_2(t_1, t_2) = e^{L(t_2) - L(t_1)} \left(C_2(t_2) - C_2(t_1) + \frac{[C_1(t_2) - C_1(t_1)]^2}{2!} \right).$$

For $k=3$ we have

$$w_3(t_1, t_3) = w_3(t_1, t_2)\, w_0(t_2, t_3) + w_2(t_1, t_2)\, w_1(t_2, t_3) + w_1(t_1, t_2)\, w_2(t_2, t_3) + \\ + w_0(t_1, t_2)\, w_3(t_2, t_3),$$

or [using (2_0), (1_0), (1_1) and (1_2)]:

$$\frac{w_3(t_1, t_3)}{w_0(t_1, t_3)} = \frac{w_3(t_1, t_2)}{w_0(t_1, t_2)} + \left\{ C_2(t_2) - C_2(t_1) + \frac{[C_1(t_2) - C_1(t_1)]^2}{2} \right\} [C_1(t_3) - C_1(t_2)] + \\ + [C_1(t_2) - C_1(t_1)] \left\{ C_2(t_3) - C_2(t_2) + \frac{[C_1(t_3) - C_1(t_2)]^2}{2} \right\} + \frac{w_3(t_2, t_3)}{w_0(t_2, t_3)}.$$

If we put

$$f(t_1, t_2) = \frac{w_3(t_1, t_2)}{w_0(t_1, t_2)} - [C_1(t_2) - C_1(t_1)] \cdot [C_2(t_2) - C_2(t_1)] - \frac{[C_1(t_2) - C_1(t_1)]^3}{3!},$$

then we get

$$f(t_1, t_3) = f(t_1, t_2) + f(t_2, t_3) \quad \text{whence} \quad f(t_1, t_2) = C_3(t_2) - C_3(t_1),$$

and thus

$$(1_3) \qquad w_3(t_1, t_2) = e^{L(t_2) - L(t_1)} \left\{ C_3(t_2) - C_3(t_1) + [C_1(t_2) - C_1(t_1)] \cdot [C_2(t_2) - C_2(t_1)] + \\ + \frac{[C_1(t_2) - C_1(t_1)]^3}{3!} \right\}.$$

(1_0), (1_1), (1_2), (1_3) are giving the result (1) for $k=0, 1, 2, 3$.

The proof of (1) can be finished by an induction on k which is the straight continuation of our above considerations and which follows exactly the same lines as that in § 2 of [1].

Similarly, also

$$L(t_1) - L(t_2) = \sum_{j=0}^{\infty} [C_j(t_2) - C_j(t_1)]$$

can be proved from the formula

$$(3) \qquad \sum_{j=0}^{\infty} w_j(t_1, t_2) = 1$$

just as the respective fact in [1], § 2.

4. In general, most of the considerations of §§ 1, 2 in [1] remain valid if we substitute

$$w_k(t_1, t_2) \quad \text{for} \quad W_k(t_2 - t_1) = W_k(t)$$
$$C_k(t_2) - C_k(t_1) \quad \text{for} \quad c_k \cdot (t_2 - t_1) = c_k t \ \text{etc.}$$

Also in our non-homogeneous case an ordinary Poisson distribution can be derived for "rare" events. Here the rarity postulate is

$$\lim_{t_2 \to t_1} \frac{1 - v_0(t_1, t_2)}{v_1(t_1, t_2)} = 0,$$

from which

$$\lim_{t_2 \to t_1} \frac{v_k(t_1, t_2)}{v_1(t_1, t_2)} = 0 \qquad (k > 1)$$

follows by (3) and by

$$0 \leqq v_k(t_1, t_2) \leqq 1.$$

[$v_k(t_1, t_2)$ is the probability of exactly k rare events occurring in the time-interval (t_1, t_2).]

Just as in [1], § 1, we find

$$v_k(t_1, t_2) = e^{C(t_1) - C(t_2)} \frac{[C(t_2) - C(t_1)]^k}{k!} \tag{4}$$

(inhomogeneous ordinary Poisson distribution).

As A. Rényi has observed ([1], p. 87), in the case of an inhomogeneous ordinary Poisson process the inhomogeneity is not essential, as it can be eliminated by a change of the scale of time (putting $t' = C(t_2) - C(t_1)$ we have $v_k = e^{t'} \frac{t'^k}{k!}$ a homogeneous ordinary Poisson distribution), — while in the case of inhomogeneous composed Poisson processes this is in general not possible: they are "genuinely" inhomogeneous.

On the other hand, inhomogeneous composed Poisson processes can be composed of inhomogeneous ordinary Poisson processes just as in the case of homogeneous processes. In fact, comparing (1) with (4) we have

$$w_k(t_1, t_2) = \sum_{r_1 + 2r_2 + \ldots + kr_k = k} v_{r_1}^{(1)}(t_1, t_2) \cdot v_{r_2}^{(2)}(t_1, t_2) \ldots v_{r_k}^{(k)}(t_1, t_2) \prod_{j=k+1}^{\infty} v_0^{(j)}(t_1, t_2),$$

where

$$v_n^{(j)}(t_1, t_2) = e^{C_j(t_1) - C_j(t_2)} \cdot \frac{[C_j(t_2) - C_j(t_1)]^n}{n!};$$

that is: also the most general inhomogeneous Markoff process of random events is the sum of an infinity of independent ordinary Poisson processes, the k-th process consisting of the random occurrence of k-tuples of events; — in accordance with the remark of A. N. Kolmogoroff ([1], p. 211).

(Received 8 September 1952)

Bibliography

[1] L. Jánossy, A. Rényi and J. Aczél, On composed Poisson distributions, I, *Acta Math. Acad. Sci. Hung.*, **1** (1950), pp. 209—224.

[2] A. Rényi, On composed Poisson distributions, II, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 83—98.

ОБОБЩЕННЫЕ РАСПРЕДЕЛЕНИЯ ТИПА ПУАССОНА, III

Я. АЦЕЛ (Дебрецен)

(Резюме)

Если $w_k(t_1, t_2)$ вероятность того, что некоторое событие произойдёт в промежуток времени (t_1, t_2) k раз, и предположим, что числа событий, происходящих в неперекрывающиеся промежутки времени, независимы, то

$$w_k(t_1, t_2) = e^{L(t_2)-L(t_1)} \sum_{\substack{r_1+2r_2+\cdots+kr_k=k \\ r_j \geqq 0}} \prod_{j=1}^{k} \frac{[C_j(t_2)-C_j(t_1)]^{r_j}}{r_j!}$$

$$L(t_2)-L(t_1) = -\sum_{k=1}^{\infty} [C_k(t_2)-C_k(t_1)] \qquad (0 \leqq w_k(t_1, t_2) \leqq 1).$$

Результат получается решением системы функциональных уравнений:

$$w_k(t_1, t_3) = \sum_{j=0}^{k} w_j(t_1, t_2)\, w_{k-j}(t_2, t_3)$$

$$\sum_{k=0}^{\infty} w_k(t_1, t_2) = 1.$$

ON FULLY DECOMPOSABLE ABELIAN TORSION GROUPS

By

A. KERTÉSZ (Debrecen)

(*Presented by* L. RÉDEI)

§ 1. Introduction

A group is said to be fully decomposable, if it can be represented as the direct sum of directly indecomposable groups, the latter being groups which do not decompose into the direct sum of two of its proper subgroups. As we know that among the abelian torsion groups only the groups $C(p^m)$ $(0 \leqq m \leqq \infty)$ are directly indecomposable ([3], [6][1]), the fully decomposable abelian torsion groups are those which are direct sums of cyclic and quasi-cyclic groups.[2] A criterion for abelian torsion groups to be fully decomposable was kown so far only for the enumerable case, in the form of a theorem of PRÜFER [5] (see Corollary 6 in § 4 of the present paper). In what follows I shall give a criterion for abelian torsion groups of arbitrary power to be fully decomposable (§ 2). This criterion is the generalization of a former result of mine, stating when an abelian p-group of arbitrary power is the direct sum of cyclic groups [2]. Generalizing this result from a different point of view, L. FUCHS obtained recently valuable new results on groups decomposable into the direct sum of cyclic groups [1]. L. FUCHS generalized the criterion of [2], by giving a condition based, instead of the height of group elements, on their "relative order", the latter being closely related to the order of group elements. On the other hand, both the criterion of [2] and that of the present paper are based on the concept of the height of group elements. The criterion in [2] reads as follows: An arbitrary abelian p-group is the direct sum of cyclic groups if and only if the group contains no element of infinite height and there exists a principal system of elements in the group. In the present paper I show that if the concept of elements of infinite height is split in a suitable manner into that of elements of *outer* resp. *inner* infinite height, then — with unchanged definition of principal system (see § 2) — the same criterion holds for an arbitrary abelian p-group to be the direct

[1] The numbers in brackets refer to the Bibliography given at the end of the paper.

[2] For notation and terminology see the next two paragraphs.

sum of cyclic and quasicyclic groups, provided that the term "element of infinite height" is replaced by "element of outer infinite height". (See Theorem in § 2.) Since an abelian torsion group is the direct sum of uniquely determined p-groups, our theorem evidently solves the problem in question for arbitrary torsion groups too. From this criterion we can easily deduce several corollaries (§§ 3, 4) which are partly new and partly well-known theorems. In the proof of Corollary 1, I am indebted for an important idea to L. FUCHS (see [4]). It may be mentioned that the present paper can be read without knowledge of [2]; the proofs are very simple, making use only of the best known fundamental concepts of group theory.

The notations and terminology used are the following. By capital letters we denote groups or some systems of group elements, by the letters $x, a, b, \ldots, g$ group elements, while the other small Latin letters are reserved for rational integers, in particular p for a prime number. The Greek letter ν may range over an arbitrary (not necessarily ordered) set of indices. In what follows, by a group we shall mean always an additively written abelian group with more than one element. A subgroup generated by the elements $a, b, \ldots$ of a group is denoted by $\{a, b, \ldots\}$. A group every element of which is of finite order, is called a *torsion group*. It is well known that a torsion group may be represented as the direct sum of its uniquely determined *primary components*, each of which is a *p-group*, i. e. a group containing only elements of p-power order. We denote the order of a group element a by $O(a)$. The *height* in G of an element a of the p-group G is defined as follows. An element $a \neq 0$ of the p-group G is said to have the height $h = H(a)$ if the equation $p^n x = a$ is solvable in G for $n \leqq h$, but not for $n > h$. We define $H(a) = \infty$ if $p^n x = a$ has a solution $x \in G$ for each natural number n. We emphasize that $H(a)$ is defined only for $a \neq 0$.[3] An important rôle is played in our investigations by the concepts of elements of inner resp. outer infinite height. We say that an element a with $H(a) = \infty$ of the p-group G is an *element of inner infinite height* if $p^t x = a$ has a solution $x \in G$ of infinite height for each natural number t. In the contrary case, when there exists a t for which $p^t x = a$ admits only solutions $x \in G$ of finite height, we call a an *element of outer infinite height*. We remark that if G is the direct sum of its subgroups $B_1, B_2, \ldots$ and $g = b_1 + b_2 + \ldots$ ($b_\nu \in B_\nu$), then evidently $H(g) \leqq H(b_\nu)$ ($\nu = 1, 2, \ldots$) holds.

The elements $a_1, \ldots, a_n$ of the group G are called *independent* if $r_1 a_1 + \ldots + r_n a_n = 0$ implies $r_1 a_1 = \ldots = r_n a_n = 0$. An arbitrary set S of elements of G is independent, if every finite subset of S is independent. The independence so defined is therefore a property of finite character and, con-

[3] By the *height* of an element g of G we always mean the number $H(g)$ defined above, i. e. the height refers always to the whole group G, even when an element is, for the moment, considered as an element of some subgroup of G.

sequently, by virtue of ZORN's lemma (or the equivalent lemma of TUKEY), every subset R of G contains a *maximal independent system* S. If $R = G$, we call S a maximal independent system of G. We denote by $C(p^m)$ in the case $m < \infty$ the cyclic group of order p^m, and in the case $m = \infty$ the group of type (p^∞) or *quasicyclic group*, i. e. the group $\{a_1, a_2, \ldots\}$ defined by the relations

(1) $$a_1 \neq 0, \quad pa_1 = 0, \quad pa_2 = a_1, \ \ldots .$$

§ 2. The main criterion

We recall the definition of principal system in [2], on which our following theorem is based.

DEFINITION. A maximal independent system P of the abelian p-group G will be called a *principal system* of G, if no element of P can be exchanged for an element of a greater height of G without violating the independence of the system.[3] (Of course, elements of infinite height are considered to be of a greater height than any element of finite height.)

REMARK. Each element of finite height of a principal system P is of order p, for an element $a \in P$ of order p^k $(k > 1)$ would be exchangeable for the element $p^{k-1}a$ of greater height. On the other hand, an element on infinite height can obviously be exchanged for an element of order p. Consequently, in what follows we may always assume that *a principal system contains only elements of order p.*

Now we state the following

THEOREM. *An abelian p-group G is fully decomposable — i. e. it decomposes into the direct sum of cyclic and quasicyclic groups — if and only if*

1) *G contains no element of outer infinite height and*

2) *there exists a principal system of elements in G.*

PROOF. The necessity of conditions 1) and 2) can easily be verified. In fact, let us suppose that G is the direct sum of cyclic and quasicyclic groups C_ν. Then 1) obviously holds. In order to prove the validity of 2), we choose an element a_ν of order p from each direct summand C_ν of G. We show that the set of all a_ν's is a principal system P of G. Indeed, P is a maximal independent system in G; moreover, for an arbitrary element $b \in G$ of order p^k a relation of the form

$$p^{k-1}b = r_1 a_1 + \ldots + r_n a_n$$

holds (with a suitable notation for the elements of P), showing that $a_1, \ldots, a_n$ are the only elements of P one of which can be replaced by b without violating the independence of the system P. Since $H(b) \leq H(p^{k-1}b) \leq H(a_i)$ $(i = 1, \ldots, n)$, no element of P can be exchanged for an element of a greater height of G.

In order to prove the sufficiency of conditions 1) and 2), let us consider an arbitrary abelian p-group without elements of outer infinite height and containing a principal system P. If a_ν is an element of infinite height of P, then from 1) we infer the existence of an infinite system of elements $a_\nu^{(1)}, a_\nu^{(2)}, \dots$ in G, such that

$$pa_\nu^{(1)} = a_\nu, \quad pa_\nu^{(2)} = a_\nu^{(1)}, \quad \dots$$

hold. Then (see (1))

$$\{a_\nu^{(1)}, a_\nu^{(2)}, \dots\} = C_\nu = C_\nu(p^{m_\nu}) \qquad (m_\nu = \infty)$$

is a quasicyclic subgroup of G containing a_ν. On the other hand, if for an element $a_\nu \in P$

$$H(a_\nu) = h_\nu < \infty$$

holds, then we determine an element $c_\nu \in G$ such that

$$p^{h_\nu} c_\nu = a_\nu. \tag{2}$$

In this case let

$$\{c_\nu\} = C_\nu = C_\nu(p^{m_\nu}) \qquad (m_\nu = h_\nu + 1).$$

Now we state that G is the direct sum of the groups C_ν. As a matter of fact, the independence of the system P guarantees that the subgroup D of G generated by the C_ν's is the *direct sum* of these groups:

$$D = \sum_\nu C_\nu.$$

Therefore only the fact $D = G$ remains to be proved.

Before proving this, we make the following remark: *A relation*

$$pg = d_1 + \dots + d_m \qquad (g \in G,\ d_i \in C_i) \tag{3}$$

implies $d_i = pd_i'$ $(d_i' \in C_i,\ i = 1, \dots, m)$. For let us suppose the contrary. Then there exists a relation

$$pg^* = r_1 c_1 + \dots + r_k c_k \qquad (g^* \in G) \tag{4}$$

for which $p \nmid r_j$ $(j = 1, \dots, k)$. In fact, each d_i in the right member of (3) which can be expressed in the form $d_i = pd_i'$ $(d_i' \in C_i)$ — and for a quasicyclic C_i this is always the case — disappears by introducing $g' = g - d_i'$. Now, among the elements $a_1, \dots, a_k$ corresponding to the elements $c_1, \dots, c_k$ in (4) let a_1 be one of maximal height:

$$h_1 = H(a_1) \geqq H(a_j) \qquad (j = 1, \dots, k).$$

Thus, by virtue of (2), (4) implies

$$a' = p^{h_1+1} g^* = r_1 a_1 + \cdots.$$

This means, however, that the element a_1 of P can be exchanged for an element a' with $H(a') > H(a_1)$, without violating the independence of the system. This contradicts the principal property of the system P.

Now suppose that D is a proper subgroup of G. Then there exists an element $g \in G$ such that

(5) $$g \notin D, \quad pg \in D,$$

i. e. a relation (3) holds. Thus $d_i = pd'_i$ $(d'_i \in C_i, i = 1, \dots, m)$ and so we have

$$p(g - d'_1 - \dots - d'_m) = 0.$$

Therefore

(6) $$g' = g - d'_1 - \dots - d'_m$$

— as an element $\neq 0$, see (5) — is an element of order p, and consequently, by the maximal property of the system P, g' is a linear combination of some a_ν's, i. e. $g' \in D$. Hence (6) implies $g \in D$ which contradicts (5), completing the proof of the Theorem.

§ 3. Abelian torsion groups in which the heights of the elements of finite height are bounded

As the first application of the Theorem in § 2 we obtain the following

COROLLARY 1. *If in an abelian p-group G the heights of the elements of finite height are bounded, then G is fully decomposable, i. e. G is the direct sum of cyclic and quasicyclic groups.*

PROOF. At first we show that a p-group G in which the height of any element of finite height is $\leqq h$ $(< \infty)$, cannot contain elements of outer infinite height. In fact, if a is an arbitrary element of infinite height of G, then there exists an $y \in G$ for which

$$p^{n+h+1} y = p^n (p^{h+1} y) = a$$

holds. Therefore the equation $p^n x = a$ has a solution $x = p^{h+1} y$ of infinite height in G (for each natural number n), $p^{h+1} y$ being obviously an element of infinite height.[4]

Now, we construct a principal system in G. Consider a maximal independent system of elements of infinite height in G, extend this system to an independent system by adjoining a maximal system of elements of height h, then by adjoining a maximal system of elements of height $h-1$, and so on. The set of elements thus obtained is obviously a principal system in G.

The following generalization of the concept of height enables us to extend Corollary 1 immediately to arbitrary torsion groups. Let a be an arbitrary element of the torsion group G, and $O(a) = p_1^{s_1} \dots p_r^{s_r}$ $(p_1, \dots, p_r$ are distinc primes and $s_1, \dots, s_r > 0)$. If there exists a maximal number n_0 among the integers of the form $n = p_1^{t_1} \dots p_r^{t_r}$ $(t_i \geqq 0)$ for which the equation $nx = a$

[4] I am indebted for this idea to L. FUCHS.

has a solution $x \in G$, then we say that the element a is of height n_0. In the contrary case we call a an element of infinite height.[5]

By making use of this definition, from Corollary 1 we are led at once to

COROLLARY 2. *If the heights of the elements of finite height of an abelian torsion group G are bounded, then G is fully decomposable, i.e. G is the direct sum of cyclic and quasicyclic groups.*

An immediate consequence of Corollary 2 is the following well known theorem:

COROLLARY 3. *If every element $(\neq 0)$ of an abelian torsion group G is of infinite height, then G is the direct sum of quasicyclic groups.*

§ 4. Further applications

A theorem of KULIKOV [4], [2] is generalized by the following

COROLLARY 4. *An abelian p-group G is the direct sum of cyclic and quasicyclic groups if and only if*

1. *G contains no element of outer infinite height, and*
2. *G is the union of a countable ascending chain*

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots \tag{7}$$

of its subgroups A_n such that, in each A_n, among the elements of finite height there exists an element of maximal height h_n.

PROOF. The necessity of conditions 1. and 2. is evident. In order to prove their sufficiency, we construct in G a principal system as follows. Choose a maximal independent system of elements $\in A_1$ of order p and of infinite height ("∞-layer"), adjoin a maximal system of elements $\in A_1$ of height h_1 ("h_1-layer"), then one of height $h_1 - 1$ ("$h_1 - 1$-layer"), and so on, such that the resulting system P_1 shall be a maximal independent system of A_1.[3] In an analoguous way we extend P_1 to a maximal independent system P_2 of A_2 by adjoining a maximal "∞-layer", then a maximal "h_2-layer" and so on. We proceed likewise in constructing successively the systems $P_3, P_4, \ldots$, each P_n being a *maximal independent system* of A_n. Now we assert that the union P of all systems $P_1, P_2, \ldots$ is a principal system in G. Clearly, P is a maximal independent system of G, and *the elements of P are all of order p.* Hence for an arbitrary element $g \in G$ of order p^k a representation

$$p^{k-1} g = r_1 a_1 + \ldots + r_n a_n \tag{8}$$

[5] We should like to call the reader's attention to the fact that in the case of p-groups this definition of height does not accord with that given in § 1, but for elements of infinite heights the two definitions coincide.

holds, $a_1, \ldots, a_n$ being suitable elements of P, ordered so that

(9) $$a_i \in A_m \text{ implies } a_{i-s} \in A_m \qquad (s = 1, \ldots, i-1).$$

Equation (8) means that $a_1, \ldots, a_n$ are the only elements of P one of which may be replaced by g without violating the independence of the system P. Therefore it is sufficient to show that

(10) $$h = H(p^{k-1}g) \leqq H(a_i) \qquad (i = 1, \ldots, n),$$

for, by $H(g) \leqq H(p^{k-1}g)$, (10) implies $H(g) \leqq H(a_i)$ $(i = 1, \ldots, n)$ which proves that P is in fact a principal system of G.

Suppose (10) is not true and let

(11) $$H(a_i) < h \leqq H(a_{i+t}) \qquad (t = 1, \ldots, n-i).$$

Then, by (8) and (9), the element

$$g' = p^{k-1}g - (r_{i+1}a_{i+1} + \ldots + r_n a_n) = r_1 a_1 + \ldots + r_i a_i \in A_m$$

with $H(g') \geqq h > H(a_i)$ (see (11)) may replace the element $a_i \in P_m$ without violating the independence of P_m, which contradicts the maximality of the $H(g')$-layer of P_m, qu. e. d.

An immediate consequence of Corollary 4 is

COROLLARY 5. *A countable abelian p-group is the direct sum of cyclic and quasicyclic groups if and only if it contains no element of outer infinite height.*

For a countable abelian p-group is the union of a countable ascending chain of finite groups.

Corollary 5 is only another formulation of the following theorem of PRÜFER [5]:

COROLLARY 6. *A countable abelian p-group G is the direct sum of cyclic and quasicyclic groups if and only if each element of infinite height of G is contained in a subgroup $C(p^\infty)$ of G.*

(Received 26 September 1952)

Bibliography

[1] L. FUCHS, The direct sum of cyclic groups, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 177—195.

[2] A. KERTÉSZ, On the decomposibility of abelian p-groups into the direct sum of cyclic groups, *Acta Math. Acad. Sci. Hung.*, **3** (1952), pp. 121—126.

[3] Л. Я. Куликов, К теории абелевых групп произвольной мощности, Матем. Сборник, **9** (51) (1941), pp. 165—182.

[4] Л. Я. Куликов, К теории абелевых групп произвольной мощности, Матем. Сборник, **16** (58) (1945), pp. 129—162.

[5] H. PRÜFER, Untersuchungen über die Zerlegbarkeit der abzählbaren primären Abelschen Gruppen, *Math. Zeitschrift*, **17** (1923), pp. 35—61.

[6] T. SZELE, Sur la décomposition des groupes abéliens, *C. R. Acad. Sci. Paris*, **229** (1949), pp. 1052—1053.

О ПОЛНО РАЗЛОЖИМЫХ ПЕРИОДИЧЕСКИХ АБЕЛЕВЫХ ГРУППАХ

А. КЕРТЕС (Дебрецен)

(Резюме)

Автор, обобщая один из своих прежних результатов, доказывает следующую теорему:

Абелева p-группа произвольной мощности является прямой суммой прямо-неразложимых абелевых групп тогда и только тогда, если она не содержит элементов внешно-бесконечной высоты, и обладает максимальной независимой системой элементов, ни один из которых не может быть заменен элементом большей высоты без нарушения независимости. (Под элементом внешно-бесконечной высоты следует при этом понимать элемент бесконечной высоты, для которого уравнение $p^t x = a$, где $t > 0$ есть подходящее целое число, допускает только решения x конечной высоты)

Из этой теоремы легко выводится ряд теоретико-групповых теорем.

CONTRIBUTION TO THE THEORY OF SEMISIMPLE RINGS

By

L. FUCHS (Budapest) and T. SZELE (Debrecen)

(*Presented by* L. RÉDEI)

§ 1. Introduction

In the theory of semisimple rings (i. e. rings with minimum condition on left ideals and without nonzero nilpotent left ideals) the most important theorem, due to WEDDERBURN—ARTIN, states that a ring R is semisimple if and only if it may be represented as the direct sum of a finite number of minimal (and non-nilpotent) left ideals.[1] In the proof of the necessity of this condition an important step is the demonstration that[2] every left ideal L of R contains an idempotent generator e, that is, $L = Re$ holds with $e^2 = e$. This is equivalent to the fact that each *left ideal* L of R has a *right unity*, i. e. an element e such that $xe = x$ for every x in L. (See Lemma 1 below.) Now the question arises as to whether there exist other rings too with this property. In this note we intend to show that the answer to this question is negative: the rings every left ideal of which contains a right unity coincide with the semisimple rings (Theorem 1). Hence it will result that the following four statements concerning a ring R are equivalent:

(i) R is semisimple,

(ii) every left ideal L of R contains an idempotent generator,

(iii) every left ideal L of R possesses a right unity,

(iv) R is the direct sum of a finite number of minimal (and non-nilpotent) left ideals.

This result will enable us to get a deeper insight into the structure of semisimple rings; indeed, we shall see that, in the derivation of (iv) from (i), the hypotheses (the minimum condition on left ideals and the absence of

[1] See e. g. [1], [2], [3] or [7]. (The numbers in brackets refer to the Bibliography given at the end of this note.)

[2] See e. g. [2, p. 28]. In certain discussions this is stated only for *minimal* left ideals L of the ring. (Minimal means that minimal in the set of all nonzero left ideals of the ring.) — We note that for the sake of uniformity the left ideal (ring) consisting of a single element 0 will also be regarded as a left ideal (ring) with unity.

nonzero nilpotent left ideals) are necessary only in the proof of (ii), while there is no further need of them in deriving (iv) from (ii) or (iii).

Next we shall consider rings every *left ideal* of which is assumed to have a *left unity* rather than right unity. It is interesting that in such rings a left unity of a left ideal L is at the same time a *twosided* unity. Hence these rings are again semisimple; moreover, it will turn out that they may be characterized by the property of being the direct sum of a finite number of *skewfields* which are twosided ideals of the whole ring (Theorem 2).

Our final problem to be discussed here consists in determining all rings every *subring* of which possesses a *left* (or right or twosided) *unity*. We shall show that a ring R enjoys this property if and only if it is the direct sum of a finite number of rings (twosided ideals of R) each of which is isomorphic o a subring of the algebraic closure $\overline{P}_{p_i}$ of the prime field P_{p_i} of prime characteristic p_i. Such a ring is obviously a field,[3] namely, an absolutely algebraic field[4] of characteristic p_i. Therefore, a ring with the mentioned property is necessarily commutative, and is finite or enumerable. Conversely, for each cardinal number $c \leqq \aleph_0$ there exists at least one ring of power c every subring of which contains a unity.

Finally, let us mention that the results of this note may be enumerated also as follows:

(1) every left ideal in a ring R has a right unity if and only if R is the direct sum of a finite number of *complete matrix rings* (over some skewfield),

(2) every left ideal in R has a left unity if and only if R is the direct sum of a finite number of *skewfields,*

(3) every subring in R has a left unity if and only if R is the direct sum of a finite number of *special fields;* these fields are absolutely algebraic modular fields.[4]

§ 2. A characterization of semisimple rings

First we prove the equivalence of conditions (ii) and (iii) stated in § 1.

LEMMA 1. *A left ideal L of an arbitrary ring R is generated by an idempotent element e, $L = Re$, if and only if L has a right unity.*

The necessity of the condition in this lemma is evident in view of the fact that the idempotent generator e itself is a right unity for L. In order to establish the sufficiency, let e' denote a right unity of the left ideal L. Then $Re' \subseteq L$. On the other hand, $L \subseteq R$ implies $L = Le' \subseteq Re'$ whence $L = Re'$ follows, indeed.[5]

[3] We refer to Corollary in § 4.

[4] By an absolutely algebraic field we mean an algebraic extension of a prime field. We shall call a field modular if its characteristic is a prime number.

[5] We remark that we have in addition proved that $L = L'e'$ for each left ideal L' of R containing L.

In the further course of this section we shall suppose that *the ring R is a non-trivial ring in which every left ideal possesses a right unity.* Then, in particular, R itself has a right unity, further, it is evident that R can contain no nilpotent left ideals other than 0 (i. e. its radical is the zero ideal).

In such rings R we have the following simple but fundamental lemma.

LEMMA 2. *If L and L_1 are left ideals of R such that $L_1 \subseteq L$, then there exists a direct decomposition*

$$L = L_1 + L_2$$

for some suitable left ideal L_2 of R.

In fact, denoting by e_1 a right unity of L_1, we have the Peirce decomposition

$$x = xe_1 + (x - xe_1) \qquad (x \in L)$$

where $xe_1 \in L_1$ and the elements $x - xe_1$ form a left ideal L_2 of R, contained in L. Since $L_2 e_1 = 0$, the decomposition of L into L_1 and L_2 is direct. This establishes Lemma 2.

Now let M_1 be a maximal left ideal of R, different from R. By ZORN's lemma, such an M_1 necessarily exists, for the proper left ideals of R may be characterized by the property of containing no right unity of R. Then, by Lemma 2, we have

$$R = L_1 + M_1$$

for some nonzero left ideal L_1. If M_1 is different from the zero ideal, then let M_2 be a maximal left ideal of R properly contained in M_1. Apply again Lemma 2 to get $M_1 = L_2 + M_2$ where $L_2 \neq 0$ is a left ideal of R. Thus proceeding, we conclude

$$R = L_1 + L_2 + M_2 = L_1 + L_2 + L_3 + M_3 = \dots$$

where each L_i $(i = 1, 2, 3, \dots)$ is a left ideal of R. This process must terminate after a finite number of steps, for in the contrary case the ring R would contain the infinite (discrete!) direct sum

$$R' = L_1 + L_2 + L_3 + \dots$$

which is a left ideal in R, but can contain, of course, no right unity. Therefore we have

$$R = L_1 + L_2 + \dots + L_n \qquad (L_i \neq 0) \tag{1}$$

for some natural integer n.

The left ideals L_i in (1) must be minimal in R, for if not, then there would exist a left ideal L' with[6] $0 \subset L' \subset L_i$ for some i, whence $M_i \subset M_i + L' \subset \subset M_i + L_i = M_{i-1}$, contrary to the choice of M_i in M_{i-1} (we have put $M_0 = R$). Consequently, R is the direct sum of a finite number of minimal left ideals whence (iv) is proved. This completes the proof of

[6] The sign $\subset$ is used to denote proper inclusion.

THEOREM 1. *A ring is semisimple if and only if every left ideal of it has a right unity.*[7]

§ 3. Rings in which every left ideal has a left unity

Now we are going to examine how our result in § 2 modifies if we assume the presence of a *left* unity, instead of a right unity, for the left ideals of the ring. First of all we show:

LEMMA 3. *If in a ring R every left ideal has a left unity, then this left unity is a twosided unity.*

Let L be an arbitrary left ideal in R and e' a left unity of L. Then we have the Peirce decomposition

$$x = xe' + (x - xe') \qquad (x \in L).$$

The set of all $x - xe'$ with $x \in L$ is a left ideal L' in R, hence it contains a left unity $y - ye'$. Thus we have

$$x - xe' = (y - ye')(x - xe') = y(x - xe') - ye'(x - xe') = \\ = y(x - xe') - y(x - xe') = 0$$

for each $x - xe' \in L'$, e' being a left unity of L. Therefore we conclude that $x = xe'$ for each $x \in L$, i. e. e' is a twosided unity in L, q. e. d.

Now let R be *a ring in which every left ideal has a left unity*, or, what is by Lemma 3 an equivalent hypothesis, *every left ideal contains a (twosided) unity*. Then, by Theorem 1, it follows that R is the direct sum of a finite number of minimal left ideals S_i:

$$R = S_1 + S_2 + \cdots + S_n \qquad (S_i \neq 0). \tag{2}$$

It is easy to see that the left ideals S_i are twosided ideals in R. Indeed, if $e = e_1 + e_2 + \cdots + e_n$ $(e_i \in S_i)$ is the unity of R, then for an arbitrary $a_i \in S_i$ we have

$$a_i = a_i e = a_i e_1 + \cdots + a_i e_i + \cdots + a_i e_n$$

whence (2) implies

$$a_i e_k = \begin{cases} a_i & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

Therefore e_i is a right and hence, by hypothesis, a twosided unity in S_i, so that we obtain

$$a_i a_k = a_i(e_k a_k) = (a_i e_k) a_k = 0 \qquad \text{for } i \neq k,$$

i. e.

$$S_i S_k = 0 \qquad \text{for } i \neq k.$$

[7] It may be noted that a similar problem on rings with unity has been solved by J. VON NEUMANN [5]. He has proved that every *principal* left (or right) ideal in a ring with unity has a right (left) unity if and only if the ring is regular in the sense that for each a there is an x such that $axa = a$.

Consequently, (2) is a decomposition of R into the direct sum of the twosided ideals S_i which are minimal as left ideals of R. Since a left ideal of S_i is at the same time a left ideal of the whole ring R, we infer that S_i contains no non-trivial left ideals, i. e., S_i is a skewfield.[8]

On the other hand, it is obvious that the direct sum $R = S_1 + S_2 + \cdots + S_n$ of n arbitrary skewfields contains 2^n left ideals each of which is the direct sum of some of the skewfields S_i. Hence any left ideal in this ring R has a (left) unity and we are led to the following result.

THEOREM 2. *If every left ideal of a ring R has a left unity, then R is the direct sum of a finite number of skewfields, and conversely: if R is the direct sum of a finite number of skewfields, then every left ideal in R has a unity.*

§ 4. The presence of a unity in every subring

This section is devoted to proving the following

THEOREM 3. *If every subring of a ring R has a left (or right) unity, then R is the direct sum of a finite number of fields F_i each of which is an absolutely algebraic modular field. Conversely, if R is the direct sum of a finite number of such fields F_i, then every subring of R has a (twosided) unity.*

Simultaneously we shall obtain the following corollary which also seems to be of some interest.

COROLLARY. *Every subring $\neq 0$ of a skewfield S is again a skewfield if and only if S is an absolutely algebraic field of prime characteristic.*

For the proof, let us assume that R is a ring every subring of which has a left unity. Then, by Theorem 2, we have a direct decomposition

$$R = F_1 + F_2 + \cdots + F_n \tag{3}$$

where each F_i $(i = 1, 2, \ldots, n)$ is a skewfield. F_i must have a prime characteristic p_i, for a skewfield of zero characteristic contains a subring isomorphic to the ring of even integers and this ring has no left unity. Moreover, each element x of the skewfield F_i is algebraic over the prime field P_{p_i} of F_i, for if $x \in F_i$ were transcendental over P_{p_i} then the ring of all polynomials of x with coefficients in P_{p_i} and without constant term would be a subring of F_i without left unity. Hence, by an important theorem of N. JACOBSON,[9] the skewfield F_i is commutative. Consequently, F_i is an absolutely algebraic field

[8] Cf. e. g. SZELE [6]. — That S_i are skewfields follows also from the known result that if R has zero radical then $L = Re$ is a minimal left ideal in R if and only if eRe is a skewfield. (See, for example, [2, p. 36].) For completeness we prefer to give here an independent proof.

[9] See Theorem 8 on p. 701 in [4].

of characteristic p_i. Thus the necessity of the condition in the above theorem and in the corollary is proved.

In order to show the sufficiency, let T denote some subring of the algebraic closure $\overline{P_p}$ of the prime field P_p. Each element $t \in T$ generates a finite ring $P_p[t]$ over P_p. Since $P_p[t]$ contains no divisors of zero, it is a field. Consequently, $P_p[t]$ and so a fortiori T contains t^{-1} which proves that T is a field. This establishes the corollary.

Finally, let R be a ring of type (3) where every component F_i is a subring (i. e. a subfield) of a field $\overline{P_{p_i}}$. Let U be some subring of R and u an element of U with a maximal number k of nonzero components a_i $(a_i \in F_i)$. As the numeration of the direct summands F_i is irrelevant, we can put

$$u = a_1 + a_2 + \cdots + a_k \qquad (a_1, a_2, \ldots, a_k \neq 0;\ a_i \in F_i).$$

Now take into account that, for a suitable natural integer m_i, we have $a_i^{m_i} = e_i$ $(i = 1, 2, \ldots, k)$, and this implies the existence of a natural integer exponent m such that

$$u^m = e_1 + e_2 + \cdots + e_k$$

where e_i denotes the unity of the field F_i $(i = 1, 2, \ldots, k)$. Considering that $u^m \in U$ is the unity of the direct sum $F_1 + F_2 + \cdots + F_k$, u^m will certainly be the unity of U if we shall have established the inclusion relation

$$U \subseteq F_1 + F_2 + \cdots + F_k.$$

For, let us suppose the contrary. Then, with a suitable notation, U contains an element

$$v = b_1 + b_2 + \cdots + b_{k+1} + \cdots \qquad (b_i \in F_i)$$

with $b_{k+1} \neq 0$. But in this case the element

$$u + (v - u^m v)$$

belongs to U and has more than k nonzero components. This contradiction completes the proof.

(Received 13 November 1952)

Bibliography

1. A. A. ALBERT, *Structure of algebras,* Amer. Math. Soc. Coll. Publ., XXIV (New York, 1939).
2. E. ARTIN, C. J. NESBITT and R M. THRALL, *Rings with minimum condition,* Univ. of Michigan Publ. in Math., 1 (Ann Arbor, 1946).
3. M. DEURING, *Algebren,* Ergebnisse d. Math. u. Grenzgeb. IV 1 (Berlin, 1935).
4. N. JACOBSON, Structure theory for algebraic algebras of bounded degree, *Annals of Math. (Princeton),* (2), **46** (1945), pp. 695—707.
5. J. VON NEUMANN, *Continuous Geometry* (Princeton, 1937).
6. T. SZELE, Die Ringe ohne Linksideale, *Buletin Stiintific,* **1** (1949), pp. 788—789.
7. B. L. VAN DER WAERDEN, *Moderne Algebra,* II, Grundlehren d. Math. Wiss. XXXIV, 2nd ed., (Berlin, 1940).

К ТЕОРИЮ ПОЛУПРОСТЫХ КОЛЕЦ

Л. ФУКС (Будапешт) и Т. СЕЛЕ (Дебрецен)

(Резюме)

Как известно, левые идеалы полупростого кольца могут быть порождены идемпотентным элементом, или, что то же самое, имеют правый единичный элемент. Можно спросить, существуют-ли также другие кольца с этим свойством. Авторы доказывают, что таких других колец нет: только для полупростых колец верно то, что все левые идеалы обладают правым единичным элементом.

В дальнейшем исследуются кольца, в которых все левые идеалы обладают левой единицей. Оказывается, что тогда левая единица является необходимо и правой. Кольца с этим свойством совпадают с кольцами представимыми в виде прямой суммы конечного числа тел.

Наконец, авторы доказывают: Для того, чтобы всякое подкольцо коль а R имело единичный элемент, необходимо и достаточно, чтобы R являлось прямой суммой конечного числа полей, каждое которых изоморфно подполю алгебраического замыкания конечного простого поля.

A CRITERION FOR DISTRIBUTIVE LATTICES

By
KIYOSHI ISÉKI (Kobé, Japan)
(*Presented by* L. RÉDEI)

The object of this short note is to give a simple proof of the following

THEOREM. *A lattice is distributive if, and only if, every meet-irreducible ideal is prime.*

A proof of this theorem may be found in my articles [3], [4].[1] The idea of the proof here given is due to L. FUCHS [2]. Before proceeding to the proof, we shall give some definitions which are needed below.[2]

A set I of elements of a lattice L is called an *ideal* of L if

1. if $x, y \in I$, then $x \cup y \in I$,

2. if $x \in I$ and $a \in L$, then $a \cap x \in L$.

An ideal I is said to be *meet-irreducible* if it is not equal to the set-intersection of two ideals each distinct from I. A proper ideal I is said to be *prime*, if $x \cap y \in I$ implies $x \in I$ or $y \in I$.

We shall first prove the following

LEMMA. Any ideal I of a lattice L is the set-intersection of all meet-irreducible ideals which contain I.

PROOF. Since L is meet-irreducible, there is at least one meet-irreducible ideal containing I. Thus the ideal I is contained in the set-intersection of all meet-irreducible ideals containing I. Let a be an element of $L—I$. By ZORN's lemma, I can be extended to an ideal M which does not contain a and is maximal with respect to the property of not containing a. The ideal M is meet-irreducible. Suppose that M is the intersection of two ideals A, B each distinct from M. By the maximal property of M, both ideals A and B must contain the element a. Hence $A \cap B$ contains a, in contradiction to $A \cap B = M$. Therefore there is a meet-irreducible ideal M containing I and not containing a given element of $L—I$. Thus the ideal I is the set-intersection of all meet-irreducible ideals containing I. The Lemma is proved.

[1] Numbers in brackets refer to the bibliography at the end of this Note.

[2] The undefined terminologies and notations are those of G. BIRKHOFF, *Lattice theory*, Amer. Math. Soc. Colloquium Publications (New York, 1948).

PROOF OF THE THEOREM. In order to prove the theorem, let us suppose that every meet-irreducible ideal is prime. We shall consider the principal ideal I generated by $a\cup(b\cap c)$ for given three elements a, b and c. Let M be a meet-irreducible ideal which contains I. Clearly $a\in M$ and $b\cap c\in M$. Since M is also prime, we have $b\in M$ or $c\in M$. Therefore $a\cup b\in M$ or $a\cup c\in M$. Hence $(a\cup b)\cap(a\cup c)\in M$. By the Lemma, I is the set-intersection of all meet-irreducible ideals M which contain the ideal I, consequently, we have $(a\cup b)\cap(a\cup c)\in I$. Hence $a\cup(b\cap c)\geqq(a\cup b)\cap(a\cup c)$. Since $a\cup(b\cap c)\leqq\leqq(a\cup b)\cap(a\cup c)$ holds in every lattice, L is distributive.

On the other hand, it was shown by G. BIRKHOFF and O. FRINK [1] that in distributive lattices, prime ideals and meet-irreducible ideals are identical.

Therefore the Theorem is proved.

(*Received 15 November 1952*)

References

[1] G. BIRKHOFF and O. FRINK, Representations of lattices by sets, *Trans. Amer. Math. Soc.*, **64** (1948), pp. 299—316.

[2] L. FUCHS, Über die Ideale arithmetischer Ringe, *Comm. Math. Helv.*, **23** (1949), pp. 334—341.

[3] K. ISÉKI, Une condition pour qu'un lattice soit distributif, *C. R. Acad. Sci. Paris*, **230** (1950), pp. 1726—1727.

[4] K. ISÉKI, Contribution to lattice theory, to appear in the *Publicationes Math., Debrecen*.

КРИТЕРИЙ ОТНОСИТЕЛЬНО ДИСТРИБУТИВНЫХ СТРУКТУР

К. ИСЕКИ (Кобе, Япония)

(Резюме)

Автор дает новое доказательство установленного им раньше факта, что структура является дистрибутивной тогда и только тогда, если все ее пересеченно-неприводимые идеалы просты.

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 du Commerce Extérieur des Livres et Journaux „Kultúra" (Budapest, VI., Sztálin-út 21. Compte-courant No. 45-790-057-50-032) ou à l'étranger ch.z 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, französicher, 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.

28 forint

INDEX

Rényi, A., On projections of probability distributions.
А. Реньи, О проекциях распределений вероятностей 131
Rédei, L., Über die Determinantenteiler.
Л. Рэдэи, О делителях определителей 143
Rédei, L., Kurzer Beweis der Waringschen Formel.
Л. Рэдэи, Краткое доказательство формулы Варинга 151
Fejes Tóth, L., Ein Beweisansatz für die isoperimetrische Eigenschaft des Ikosaeders.
Л. Фееш Тот, Метод доказательства изопериметрического свойства икозаэдра 155
Makai, E., On a monotonic property of certain Sturm—Liouville functions.
Е. Макаи, Об одном свойстве монотонности функций Штурма—Лиувилля . 165
Freud, G., Über einen reihentheoretischen Satz von Fejér.
Г. Фройд, Об одной теореме Фэйера в теории рядов 173
Fuchs, L., The direct sum of cyclic groups.
Л. Фукс, О прямой сумме циклических групп 177
Bereczki, I., Lösung eines Markovschen Problems betreffs einer Ausdehnung des Begriffes der elementaren Funktion.
И. Бэрецки, Решение одной проблемы Маркова, связанной с распространением понятия элементарной функции 197
Aczél, J., On composed Poisson distributions, III.
Я. Ацел, Обобщенные распределения типа Пуассона, III 219
Kertész, A., On fully decomposable abelian torsion groups.
А. Кертес, О полно разложимых периодических абелевых группах 225
Fuchs, L., and *Szele, T.*, Contribution to the theory of semisimple rings.
Л. Фукс и Т. Селе, К теорию полупростых колец 233
Iséki, K., A criterion for distributive lattices.
К. Исеки, Критерий относительно дистрибутивных структур 241

Megjelent 750 példányban

Technikai szerkesztő: Fuchs László
Akadémiai Kiadó (Budapest, V., Alkotmány-u. 21.) Felelős: Mestyán János

Csongrádmegyei Nyomdaipari Vállalat, Szeged 524520 Felelős vezető: Vincze György

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS III. FASCICULUS 4

MAGYAR TUDOMÁNYOS AKADÉMIA
BUDAPEST, 1952

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 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) külföld számára pedig a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalatnál (Budapest, VI., Sztálin-út 21. Bankszámla 45-790-057-50-032 sz.) vagy külföldi képviseleteinél 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.

VOLLIDEALRINGE IM WEITEREN SINN, I

Von

L. RÉDEI (Szeged), korrespondierendem Mitglied der Akademie

§ 1. Einleitung

Unter einem *Vollidealring i. w. S.* (= im weiteren Sinn) verstehe ich einen Ring, dessen alle Unterringe (zweiseitige) Ideale sind.[1] In der vorliegenden Mitteilung I dieser Arbeit bestimme ich die durch ein Element erzeugten Vollidealringe i. w. S., in einer späteren Mitteilung II werde ich dann das Problem allgemein erledigen.

Stets bezeichne I den Ring der ganzen Zahlen, $I[x]$ den Polynomring des Unbestimmten x über I. Dann ist das Produkt $xI[x]$ der Unterring von $I[x]$ bestehend aus den durch x teilbaren Polynomen. Abgesehen von x bezeichnen kleine lateinische Buchstaben (auch mit Indizes versehen) Elemente von I, insbesondere bezeichnet p eine Primzahl.

SATZ. *Die sämtlichen durch ein Element erzeugten Vollidealringe i. w. S. sind (bis auf Isomorphie) die Restklassenringe*

$$\text{(I)} \qquad xI[x]/x(x-a)\prod_p \mathfrak{P},$$

$$\text{(II)} \qquad xI[x]/x\prod_p \mathfrak{P},$$

wobei man die Produkte $\prod_p \mathfrak{P}$ auf endlich viele verschiedene p zu erstrecken hat, die in (I) *der Bedingung $p|a$ unterworfen sind, und $\mathfrak{P}$ in* (I) *eins der Ideale*

$$\text{(1)} \qquad (x(x-b),p) \qquad (p\nmid b),$$

$$\text{(2)} \qquad (x-b,p),$$

in (II) *eins der Ideale*

$$\text{(3)} \qquad (x(x-a)(x-b),p(x-a),p^n) \qquad (p|a;\ \ p\nmid b;\ \ n\geqq 2),$$

$$\text{(4)} \qquad ((x-a)(x-b)+p^{n-1}bc,p(x-a),p^n) \qquad (p|a;\ \ n\geqq 2),$$

$$\text{(5)} \qquad (x-a,p^n) \qquad (n\geqq 1),$$

$$\text{(6)} \qquad (x^2(x-a),p) \qquad (p\nmid a),$$

$$\text{(7)} \qquad (x(x-a),p)$$

von $I[x]$ bezeichnet. (In (3) bis (7) ändern sich a, b, c, n mit p.)

[1] Unter einem Vollidealring schlechthin verstehe ich einen Ring, in dem sogar alle Untermoduln Ideale sind. Diese Ringe habe ich neulich bestimmt. Vgl. L. RÉDEI, Die Vollidealringe, *Monatshefte f. Math.*, 56 (1952), S. 89—95. Sowohl die Vollidealringe als auch die Vollidealringe i. w. S. bilden je ein ringtheoretisches Analogon der sogenannten Hamiltonschen Gruppen (s. den Anfang der zitierten Arbeit).

BEMERKUNG. Zur richtigen Deutung dieses Satzes muß bemerkt werden, daß das x-fache eines Ideals von $I[x]$ stets ein Ideal von $xI[x]$ ist (s. Hilfssatz 1). Die Produkte $\prod_p \mathfrak{P}$ in (I), (II) dürfen auch leer (d. h. gleich 1) sein. Man sieht, daß die Ringe (I), (II) unendlich bzw. endlich sind. Es ist leicht die im Satz angegebenen Ringe abstrakt zu definieren. Z. B. wenn in (II) nur ein $\mathfrak{P}$ vorkommt und dies gleich (3) ist, so handelt es sich um den durch die Gleichungen

$$\varrho^4-(a+b)\varrho^3+ab\varrho^2=0,\ p(\varrho^2-a\varrho)=0,\ p^n\varrho=0 \qquad (p|a;\ p\nmid b;\ n\geqq 2)$$

definierten Ring mit dem Erzeugenden ϱ.

Wir machen schon hier die Bemerkung, von der wir öfters Gebrauch machen werden, daß die Eigenschaft von Ringen, Vollidealring i. w. S. zu sein, gegenüber homomorphe Abbildungen und Bildung von Unterringen invariant ist. D. h. wenn R ein Vollidealring i. w. S. ist und S, T ein homomorphes Bild bzw. einen Unterring von R bedeutet, so sind auch S, T Vollidealringe i. w. S. Die zweite Behauptung ist klar. Zum Beweis der ersten Behauptung betrachten wir einen Unterring S' von S. In der gegebenen Homomorphie $R\sim S$ entspricht diesem Unterring S' von S ein Unterring R' von R. Da nach der Annahme R' ein Ideal in R ist, so folgt, daß auch S' ein Ideal in S ist. Dies bedeutet die Richtigkeit der Behauptung.

Obwohl wir uns, wie gesagt, mit den Vollidealringen i. w. S. in voller Allgemeinheit erst in Mitteilung II beschäftigen werden, trotzdem wollen wir auf eine bemerkenswerte Eigenschaft dieser Ringe schon jetzt hinweisen. Wir sehen, daß die Ideale in (1), (2) und in (3) bis (7) je ein Hauptpolynom[2] vom Grade $\leqq 2$ bzw. $\leqq 3$ enthalten. Ähnliches gilt dann offenbar auch über die Produkte $\prod_p \mathfrak{P}$ in (I) und (II). Folglich enthalten die Ideale $x(x-a)\prod_p \mathfrak{P}$, $x\prod_p \mathfrak{P}$ in (I) und (II) je ein Hauptpolynom in $xI[x]$ vom Grade $\leqq 4$. Das hat zur Folge, daß sich die im Satz angegebenen Ringe durch ein algebraisches Element vom Grade $\leqq 3$ erzeugen lassen.[3] Da nach obigem jedes Element eines Vollidealringes i. w. S. einen Vollidealring i. w. S. erzeugt, so führt unser Satz zum folgenden:

KOROLLAR. *Alle Elemente eines Vollidealringes i. w. S. sind algebraisch vom Grade* $\leqq 3$.

[2] Ein Polynom nennen wir ein *Hauptpolynom*, wenn sein Anfangskoeffizient 1 ist.

[3] Wir sagen, daß ein Ringelement ϱ algebraisch vom Grade $n(\geqq 1)$ ist, wenn es ein Hauptpolynom

$$f(x)=x^{n+1}+a_1x^n+\ldots+a_nx$$

mit rationalen Koeffizienten $a_1,\ldots,a_n$ gibt, wofür $f(\varrho)=0$ gilt und dabei n minimal ist. (Natürlich müssen die a_i für die meisten Ringe ganz sein, damit $f(\varrho)$ einen Sinn hat.) Insbesondere für Körper stimmt diese Definition des Grades mit der klassischen Definition überein.

§ 2. Erste Reduktion des Beweises

Es ist klar, daß die durch ein Element erzeugten Ringe mit den aus dem Ring $xI[x]$ gebildeten Restklassenringen isomorph sind. Folglich kommt es bei jedem Problem der durch ein Element erzeugten Ringe, insbesondere also auch bei dem Beweis unseres Satzes, nur auf die Ideale des Ringes $xI[x]$ an. Über diese beweisen wir den folgenden:

HILFSSATZ 1. *Die Ideale des Ringes $xI[x]$ sind die x-fachen der Ideale des Ringes $I[x]$.*

Jedes Ideal von $xI[x]$ ist nämlich von der Form $x\mathfrak{a}$, wobei $\mathfrak{a}$ eine Teilmenge von $I[x]$ ist. Mit $x\mathfrak{a}$ zusammen muß auch $\mathfrak{a}$ einen Modul bilden. Betrachten wir ein beliebiges $f(x)\,(\in I[x])$. Es läßt sich

$$f(x)=g(x)+a \qquad (g(x)\in x\,I[x];\ a\in I)$$

setzen. Wegen der Annahme gilt $g(x)x\mathfrak{a}\subseteq x\mathfrak{a}$, also $g(x)\mathfrak{a}\subseteq\mathfrak{a}$. Ferner gilt nach obigem $a\mathfrak{a}\subseteq\mathfrak{a}$. Aus beiden folgt $f(x)\mathfrak{a}\subseteq\mathfrak{a}$. Hiernach ist $\mathfrak{a}$ ein Ideal von $I[x]$.

Umgekehrt, wenn letzteres gilt, so ist $x\mathfrak{a}$ ein Untermodul von $xI[x]$. Ist $f(x)\in xI[x]$, so gilt noch mehr $f(x)\in I[x]$, woraus wegen der Annahme $f(x)\mathfrak{a}\subseteq\mathfrak{a}$ folgt. Dies ergibt $f(x)x\mathfrak{a}\subseteq x\mathfrak{a}$, somit ist $x\mathfrak{a}$ ein Ideal von $xI[x]$. Hilfssatz 1 haben wir bewiesen.

Nachher sei $\mathfrak{a}$ stets ein Ideal von $I[x]$. Nach Hilfssatz 1 und der vorangeschickten Bemerkung sind die Restklassenringe

$$R=xI[x]/x\mathfrak{a} \tag{8}$$

(bis auf Isomorphie) die sämtlichen Ringe, die sich durch ein Element erzeugen lassen. Wir nennen das Ideal $\mathfrak{a}$ *zulässig*, wenn (8) ein Vollidealring i. w. S. ist. Hiernach haben wir nur die zulässigen Ideale zu bestimmen, und zwar ist unser Satz offenbar mit der folgenden Behauptung äquivalent: *Die sämtlichen zulässigen Ideale sind die in* (I) *und* (II) *vorkommenden Produkte* $(x-a)\prod\limits_{p}\mathfrak{P}$ *bzw.* $\prod\limits_{p}\mathfrak{P}$.

Die Grundlage zur Bestimmung der zulässigen Ideale bildet der folgende:

HILFSSATZ 2. *Ein Ideal $\mathfrak{a}$ von $I[x]$ ist dann und nur dann zulässig, wenn für jedes Polynom $\varphi(x)(\in xI[x])$ die Kongruenz*

$$\zeta(\varphi(x))\equiv x\varphi(x) \pmod{x\mathfrak{a}} \tag{9}$$

mit einem Polynom $\zeta(x)(\in xI[x])$ lösbar ist.

BEMERKUNG. Da die linke Seite von (9) durch x teilbar ist, so dürfen wir (9) auch in der Form

$$x^{-1}\zeta(\varphi(x))\equiv\varphi(x) \pmod{\mathfrak{a}} \tag{9'}$$

schreiben. Indem wir das „unbekannte" Polynom $\zeta(x)$ in der Form $\sum\limits_{i=1}^{m} z_i x^i$

annehmen ($z_i \in I$), können wir (9') weiter durch

$$(9'') \qquad \sum_{i=1}^{m} x^{-1}\varphi(x)^i z_i \equiv \varphi(x) \pmod{\mathfrak{a}}$$

ersetzen. Wenn wir bei einem $\varphi(x)(\in xI[x])$ kurz über die Lösbarkeit von (9) (oder (9')) bzw. (9'') sprechen, so soll das stets bedeuten, daß es ein Polynom $\zeta(x)(\in xI[x])$ bzw. (eine ganze Zahl $m \geqq 1$ und) ganze Zahlen $z_1, \ldots, z_m$ gibt, wofür (9) (d. h. (9')) bzw. (9'') gilt. Ist Lösbarkeit für alle $\varphi(x)$ vorhanden, so sagen wir kurz, daß (9) (d. h. (9')) bzw. (9'') (für das betrachtete $\mathfrak{a}$) *lösbar* ist. (Nebenbei bemerken wir, daß im Falle der Lösbarkeit sich stets eine Lösung $\zeta(x)$ vom Grade $m \leqq 3$ finden wird; das ist ja wegen des Satzes eine Notwendigkeit.)

Zum Beweis von Hilfssatz 2 bezeichne α die Restklasse x (mod $x\mathfrak{a}$) im Ring $xI[x]$. Dann ist α ein erzeugendes Element des Ringes R in (8), folglich sind die $\varphi(\alpha)$ die sämtlichen Elemente von R.

Ist nun $\mathfrak{a}$ zulässig, d. h. R ein Vollidealring i. w. S., so muß der durch ein $\varphi(\alpha)$ erzeugte Unterring ein Ideal sein. Dieses enthält dann das Produkt $\alpha\varphi(\alpha)$, folglich gilt

$$(10) \qquad \zeta(\varphi(\alpha)) = \alpha\varphi(\alpha)$$

mit einem passenden $\zeta(x) \in xI[x]$. Nach der Bedeutung von α besagt (10) dasselbe wie (9), womit wir bewiesen haben, daß (9) für jedes zulässige $\mathfrak{a}$ lösbar ist.

Umgekehrt nehmen wir für ein $\mathfrak{a}$ die Lösbarkeit von (9) an. Bezeichne R' einen Unterring von R. Ein beliebiges Element von R' läßt sich in der Form $\varphi(\alpha)$ schreiben. Aus der Lösbarkeit von (9) folgt das Bestehen einer Gleichung (10) mit einem passenden $\zeta(x) \in xI[x]$. Da die linke Seite von (10) in R' liegt, so liegt auch $\alpha\varphi(\alpha)$ in ihm. Hiernach gilt $\alpha R' \subseteq R'$. Da α ein erzeugendes Element von R ist, so folgt hieraus $RR' \subseteq R'$. Dies bedeutet, daß R' ein Ideal von R ist. Wir haben bekommen, daß R ein Vollidealring i. w. S. ist, was der Zulässigkeit von $\mathfrak{a}$ gleichkommt. Hilfssatz 2 haben wir bewiesen.

Unser übriges Verfahren wird darin bestehen, daß wir auf Grund von Hilfssatz 2 unter allen Idealen von $I[x]$ die zulässigen aussuchen. Dabei werden wir die von KRONECKER und HENSEL[4] ausgearbeitete Idealtheorie von $I[x]$ zu Hilfe nehmen. Zur Bequemlichkeit des Lesers stellen wir alle Tatsachen dieser Idealtheorie zusammen, die wir zu unseren Ausführungen gebrauchen, jede solche Tatsache nennen wir aber erst an der Stelle, wo sie zur Anwendung kommt.

Da $I[x]$ offenbar kein Vollidealring i. w. S. ist, so ist $\mathfrak{a} = 0$ kein zulässiges Ideal. Deshalb soll im folgenden für die Ideale $\mathfrak{a}$ von $I[x]$ stillschwei-

[4] KRONECKER—HENSEL, *Vorlesungen über Zahlentheorie* (Leipzig, 1901), S. 1—509. (Vgl. die Seiten 194—240 dieses Werkes.)

gend nur der Fall $\mathfrak{a} \neq 0$ berücksichtigt werden. Den größten gemeinsamen Teiler der Elemente von $\mathfrak{a}$ nennen wir den *Hauptidealfaktor* von $\mathfrak{a}$ und bezeichnen ihn stets mit $F(x)$; dabei dürfen wir annehmen, daß der Anfangskoeffizient von $F(x)$ positiv ist. Ist $F(x)=1$, so nennen wir $\mathfrak{a}$ *primitiv*. Stets ist $F(x)^{-1}\mathfrak{a}$ ein primitives Ideal.

HILFSSATZ 3. *Ist $\mathfrak{a}$ ein zulässiges Ideal von $I[x]$, so ist $F(x)=x-a$ oder $F(x)=1$.*

Da nämlich $\mathfrak{a} \subseteqq (F(x))$, also $x\mathfrak{a} \subseteqq (xF(x))$ ist, so folgt aus Hilfssatz 2 noch mehr die Lösbarkeit von

$$\zeta(\varphi(x)) \equiv x\varphi(x) \pmod{xF(x)}. \tag{11}$$

Insbesondere für $\varphi(x)=x^2$ lautet (11) als

$$\zeta(x^2) \equiv x^3 \pmod{xF(x)}.$$

Hiernach ist $F(x)$ gewiß durch keine ganze Zahl >1 teilbar. Deshalb gilt

$$F(x)=a_0x^k+\cdots+a_k \qquad ((a_0,\ldots,a_k)=1\,;\quad a_0>0).$$

Setzen wir in (11) $\varphi(x)=cx$ $(c \geqq 1)$ ein:

$$\zeta(cx) \equiv cx^2 \pmod{xF(x)}. \tag{12}$$

Da $\zeta(x)$ das konstante Glied 0 hat, so folgt hieraus nach Dividieren durch x und Weglassen von durch c^2 teilbaren Gliedern

$$bc \equiv cx \pmod{F(x),\, c^2},$$

wobei b den Koeffizienten des Gliedes ersten Grades von $\zeta(x)$ bezeichnet. Nach Dividieren durch c entsteht

$$x-b \equiv 0 \pmod{F(x),\, c}.$$

Wendet man dies mit einer Primzahl c an, die kein Teiler von a_0 ist, so sieht man, daß $F(x)$ vom Grade $\leqq 1$ sein muß.

Ist $F(x)$ konstant, so muß nach dem schon bewiesenen $F(x)=1$ sein. Im anderen Falle gilt

$$F(x)=a_0x+a_1 \qquad ((a_0, a_1)=1\,;\quad a_0>0).$$

Wenden wir (12) mit $c=a_0$ an und setzen für x die Zahl $-\dfrac{a_1}{a_0}$ ein, so entsteht wegen $F\left(-\dfrac{a_1}{a_0}\right)=0$ die Gleichung

$$\zeta(-a_1)=\frac{a_1^2}{a_0}.$$

Da die linke Seite ganz ist, so folgt $a_0=1$, d. h. $F(x)=x+a_1$. Hilfssatz 3 haben wir bewiesen.

Nach KRONECKER—HENSEL[4] lassen sich die primitiven Ideale von $I[x]$ so charakterisieren, daß sie eine positive ganze Zahl und ein Hauptpolynom vom Grade $\geqq 1$ enthalten. Enthält ein solches Ideal eine Potenz (>1) einer Primzahl p, so nennen wir es kurz ein *p-Ideal*. Nach KRONECKER—HENSEL[4] zerfällt jedes Ideal $\mathfrak{a}$ von $I[x]$ in das Produkt des Hauptidealfaktors $F(x)$ und von p_i-Idealen $(i=1,\ldots,k)$, die zu verschiedenen Primzahlen $p_1,\ldots,p_k$ gehören und bis auf die Reihenfolge eindeutig bestimmt sind. Die letzteren nennen wir die *p-Komponenten* von $\mathfrak{a}$ $(p=p_1,\ldots,p_k)$.

HILFSSATZ 4. *Ein Ideal $\mathfrak{a}$ von $I[x]$, das kein Hauptideal ist, ist dann und nur dann zulässig, wenn alle $F(x)\mathfrak{P}$ zulässig sind, wobei $F(x)$ (wie auch schon in Hilfssatz 3) den Hauptidealfaktor von $\mathfrak{a}$ bezeichnet und $\mathfrak{P}$ jede p-Komponente von $\mathfrak{a}$ durchläuft.*

„Nur dann" folgt wegen $\mathfrak{a}\subseteq F(x)\mathfrak{P}$ sofort aus Hilfssatz 2.

Zum Beweis von „dann" setzen wir

$$\mathfrak{a}=F(x)\mathfrak{D}, \tag{13}$$

wobei $\mathfrak{D}(\neq 1)$ ein primitives Ideal ist. Bezeichne d (>1) die kleinste in $\mathfrak{D}$ enthaltene positive ganze Zahl. Man setze

$$d=P_1\ldots P_k \qquad (k\geqq 1), \tag{14}$$

wobei $P_1,\ldots,P_k$ Potenzen von verschiedenen Primzahlen sind. Für die Ideale

$$\mathfrak{P}_i=(\mathfrak{D},P_i) \qquad (i=1,\ldots,k) \tag{15}$$

gilt offenbar

$$\mathfrak{D}=(\mathfrak{D},d)=\bigcap_{i=1}^{k}(\mathfrak{D},P_i)=\bigcap_{i=1}^{k}\mathfrak{P}_i.$$

Wegen $(\mathfrak{P}_i,\mathfrak{P}_j)=1$ $(i\neq j)$ folgt hieraus nach bekanntem Satz[5]

$$\mathfrak{D}=\prod_{i=1}^{k}\mathfrak{P}_i.$$

Da nach (14), (15) P_i die kleinste positive ganze Zahl in $\mathfrak{P}_i$ ist, so sind $\mathfrak{P}_1,\ldots,\mathfrak{P}_k$ nach (13) eben die sämtlichen p-Komponenten von $\mathfrak{a}$.

Angenommen nunmehr, daß alle $F(x)\mathfrak{P}_i$ $(i=1,\ldots,k)$ zulässig sind, betrachte man ein beliebiges $\varphi(x)\,(\in xI[x])$ und bestimme nach Hilfssatz 2 je eine Lösung $\zeta_1(x),\ldots,\zeta_k(x)$ von

$$\zeta_i(\varphi(x))\equiv x\varphi(x) \pmod{xF(x)\mathfrak{P}_i} \qquad (\zeta_i(x)\in xI[x];\ i=1,\ldots,k). \tag{16}$$

Wir setzen

$$d_i=P_i^{-1}d \qquad (i=1,\ldots,k). \tag{17}$$

Dann sind $d_1,\ldots,d_k$ nach (14) relativ prime ganze Zahlen. Wir nehmen eine Lösung von

$$d_1q_1+\cdots+d_kq_k=1 \tag{18}$$

und zeigen, daß

$$\zeta(x)=d_1q_1\zeta_1(x)+\cdots+d_kq_k\zeta_k(x) \tag{19}$$

eine Lösung von (9) ist, woraus nach Hilfssatz 2 die Richtigkeit von Hilfssatz 4 folgen wird.

[5] B. L. VAN DER WAERDEN, *Moderne Algebra*, Bd. II, 2. Aufl. (Berlin, 1940), S. 41.

Zum Beweis der Behauptung definieren wir die Polynome $\varkappa_1(x),\dots,\varkappa_k(x),\varkappa(x)$ durch

(20) $$x\varkappa_i(x)=\zeta_i(\varphi(x)) \qquad (i=1,\dots,k),$$

(21) $$\varkappa(x)=d_1q_1\varkappa_1(x)+\cdots+d_kq_k\varkappa_k(x).$$

Aus (19) bis (21) folgt

(22) $$\zeta(\varphi(x))=x\varkappa(x).$$

Ferner gilt nach (16), (20)

(23) $$\varkappa_i(x)\equiv\varphi(x) \pmod{F(x)\mathfrak{P}_i} \qquad (i=1,\dots,k).$$

Da nach der Annahme insbesondere $F(x)\mathfrak{P}_i$ zulässig ist, so folgt aus Hilfssatz 3, daß nur die Fälle $F(x)=x-a$, $F(x)=1$ möglich sind. Weiter betrachten wir diese zwei Fälle getrennt.

Fall $F(x)=x-a$. Aus (23) folgt

(24) $$\varkappa_i(a)=\varphi(a) \qquad (i=1,\dots,k).$$

Wird dies aus (23) subtrahiert, so entsteht nach Dividieren durch $F(x)\,(=x-a)$

$$\frac{\varkappa_i(x)-\varkappa_i(a)}{x-a}\equiv\frac{\varphi(x)-\varphi(a)}{x-a} \pmod{\mathfrak{P}_i} \qquad (i=1,\dots,k),$$

wobei beide Seiten Polynome in $I(x)$ sind. Multipliziert man beiderseits mit d_i, so entsteht eine richtige Kongruenz mod $d_i\mathfrak{P}_i$. Da ferner wegen $d\in\mathfrak{D}$ und (15), (17) $d_i\mathfrak{P}_i\subseteq\mathfrak{D}$ ist, so gilt noch mehr

$$d_i\frac{\varkappa_i(x)-\varkappa_i(a)}{x-a}\equiv d_i\frac{\varphi(x)-\varphi(a)}{x-a} \pmod{\mathfrak{D}} \qquad (i=1,\dots,k).$$

Wird auch noch mit q_i multipliziert, so entsteht nach Addieren wegen (18), (21)

$$\frac{\varkappa(x)-\varkappa(a)}{x-a}\equiv\frac{\varphi(x)-\varphi(a)}{x-a} \pmod{\mathfrak{D}}.$$

Andererseits folgt aus (18), (21), (24)

$$\varkappa(a)=\varphi(a).$$

Somit gilt wegen $x-a=F(x)$ und (13)

$$\varkappa(x)\equiv\varphi(x) \pmod{\mathfrak{a}},$$

d. h.

(25) $$x\varkappa(x)\equiv x\varphi(x) \pmod{x\mathfrak{a}}.$$

Dies bedeutet nach (22) eben das Bestehen von (9).

Fall $F(x)=1$. Jetzt schließen wir kürzer. Aus (23) folgt nach Multiplikation durch d_i ähnlich wie vorher

$$d_i\varkappa_i(x)\equiv d_i\varphi(x) \pmod{\mathfrak{D}} \qquad (i=1,\dots,k).$$

Nach (13) ist jetzt $\mathfrak{D}=\mathfrak{a}$. Wenn also mit q_ix multipliziert und dann addiert wird, so entsteht wieder (25). Dies bedeutet nach (22) auch jetzt das Bestehen von (9), womit wir Hilfssatz 4 bewiesen haben.

HILFSSATZ 5. *Ein Hauptideal* $\mathfrak{a}$ *von* $I(x)$ *ist dann und nur dann zulässig, wenn* $\mathfrak{a}=x-a$ *oder* $\mathfrak{a}=1$ *ist.*

Hiervon ist „nur dann“ wegen Hilfssatz 3 richtig. Die Richtigkeit von „dann“ folgt aus Hilfssatz 2, da (9) in den Fällen $\mathfrak{a}=x-a$, $\mathfrak{a}=1$ für jedes $\varphi(x)\,(\in xI[x])$ durch $\zeta(x)=ax$ bzw. durch $\zeta(x)=0$ befriedigt ist.

Aus den Hilfssätzen 3, 4, 5 sehen wir, daß unser Satz auf folgende Behauptung reduziert ist: *Bezeichne* $F(x)$ *das Polynom* $x-a$ *oder* 1, *ferner bezeichne* $\mathfrak{P}$ *ein p-Ideal. Dann und nur dann ist* $F(x)\mathfrak{P}$ *zulässig, wenn* $\mathfrak{P}$ *den Fällen* $F(x)=x-a$, $F(x)=1$ *entsprechend eins der in* (1), (2) *bzw.* (3) *bis* (7) *angegebenen Ideale ist.*

§ 3. Zweite Reduktion des Beweises

Wegen des eben gesagten legen wir uns ein Ideal

$$\mathfrak{a}=F(x)\mathfrak{P} \qquad (F(x)=x-a \text{ oder } F(x)=1) \tag{26}$$

vor, wobei $\mathfrak{P}$ ein p-Ideal in $I[x]$ ist. Wir haben zu untersuchen, welche dieser $\mathfrak{a}$ zulässig sind.

Um das spätere nicht unterbrechen zu müssen, teilen wir hier mit, wie die sämtlichen p-Ideale $\mathfrak{P}$ nach KRONECKER—HENSEL[4] bestimmt sind. (Das diesbezügliche macht den Hauptinhalt der Kronecker—Henselschen Theorie der Ideale von $I[x]$ aus.)

Jedes p-Ideal läßt sich in der Form

$$\mathfrak{P}=(F_0(x),\ldots,F_t(x)) \qquad (t\geqq 1) \tag{27}$$

annehmen, so daß

$$F_i(x)=p^{d_i}(x^{n_i}+\ldots) \tag{28}$$

das p^{d_i}-fache eines Hauptpolynoms n_i-ten Grades in $I[x]$ ist und

$$n_0>\ldots>n_t=0, \qquad 0=d_0<\ldots<d_t, \tag{29}$$

ferner

$$\left\{\begin{aligned} p^{d_1}F_0(x) &= f_{11}(x)F_1(x)+\ldots+f_{1t}(x)F_t(x),\\ p^{d_2-d_1}F_1(x) &= f_{22}(x)F_2(x)+\ldots+f_{2t}(x)F_t(x),\\ &\;\;\vdots\\ p^{d_t-d_{t-1}}F_{t-1}(x) &= f_{tt}(x)F_t(x),\\ F_t(x) &= p^{d_t} \end{aligned}\right. \tag{30}$$

gelten, wobei $f_{jj}(x)\,(\in I[x])$ ein Hauptpolynom vom Grade $n_{j-1}-n_j$ $(j=1,\ldots,t)$ und $f_{ij}(x)\,(\in I[x])$ ein Polynom vom Grade $<n_{j-1}-n_j$ $(1\leqq i<j\leqq t)$ ist. Wir nennen die Polynome $F_0(x),\ldots,F_t(x)$ mit den gesagten Eigenschaften eine *Kronecker—Henselsche Basis* von $\mathfrak{P}$.[6]

[6] Vollständigkeitshalber bemerken wir noch, daß die Zahlen (29) durch $\mathfrak{P}$ eindeutig bestimmt sind. Ferner läßt sich die Kronecker—Henselsche Basis von $\mathfrak{P}$ auf genau eine Weise so wählen, daß die Koeffizienten von $f_{ij}(x)$ aus den Zahlen $0,1,\ldots,p^{d_j-d_{j-1}}-1$ sind $(1\leqq i<j\leqq t)$. Um Mißverständnisse zu vermeiden, bemerken wir, daß (27) im allgemeinen keine „kürzeste“ Basisdarstellung ist.

Außerdem gilt, daß sich die Elemente von $\mathfrak{P}$ eindeutig in der Form

(31) $$f_0(x)F_0(x)+\ldots+f_t(x)F_t(x) \qquad (f_i(x)\in I[x];\, i=0,\ldots,t)$$

angeben lassen, so daß $f_i(x)F_i(x)$ von kleinerem Grade ist als $F_{i-1}(x)(i=1,\ldots,t)$. Der späteren Anwendung halber bemerken wir, daß die Bedingungen (30) insbesondere für den Fall $t=2$ so lauten:

(32) $$\begin{aligned} p^{d_1}F_0(x) &= f_{11}(x)F_1(x)+f_{12}(x)F_2(x),\\ p^{d_2-d_1}F_1(x) &= \qquad\qquad\quad f_{22}(x)F_2(x),\\ F_2(x) &= p^{d_2}. \end{aligned}$$

Aus diesen Gleichungen folgt

(33) $$F_0(x)=f_{11}(x).\,p^{-d_1}F_1(x)+p^{d_2-d_1}f_{12}(x).$$

Im folgenden halten wir für das Ideal $\mathfrak{P}$ in (26) alle Bezeichnungen (27) bis (31) fest.

Hilfssatz 6. *Ist das Ideal* $\mathfrak{a}$ *in* (26) *zulässig, so gilt*[7]

(34) $$F(x)F_0(x)\,|\,x^2(x-b) \pmod p \qquad (p \nmid b)$$

mit einem passenden b.

Da nämlich nach (26) bis (28) $x\mathfrak{a}\subset(xF(x)F_0(x),p)$ gilt, so muß nach Hilfssatz 2

(35) $$\zeta(\varphi(x))\equiv x\varphi(x) \pmod{xF(x)F_0(x),p}$$

für jedes $\varphi(x)\,(\in xI[x])$ mit einem $\zeta(x)\,(\in xI[x])$ lösbar sein.

Hieraus folgern wir zunächst, daß mit passenden $k_0,\ldots,k_{p-1}(\geqq 0)$

(36) $$F(x)F_0(x)\equiv\prod_{i=0}^{p-1}(x-i)^{k_i} \pmod p$$

gilt.

Die Faktoren links sind Hauptpolynome, ferner ist nach (26) stets $F(x)\,|\,x-a$. Wäre also (36) falsch, so gilt

(37) $$f(x)\,|\,F_0(x) \pmod p$$

mit einem mod p irreduziblen Polynom $f(x)\,(\in I[x])$ vom Grade $r\,(\geqq 2)$. Wir setzen in (35)

$$\varphi(x)=x^{p^r-1}$$

ein. Da hierfür bekanntlich $\varphi(x)\equiv 1 \pmod{f(x),p}$ gilt, so folgt aus (35), (37) offenbar

$$\zeta(1)\equiv x \pmod{f(x),p}.$$

Hiernach ist $f(x)\,|\,x-\zeta(1)$ (mod p), somit ist $f(x)$ linear. Dieser Widerspruch beweist (36).

[7] Mit $f(x)\,|\,g(x) \pmod p$ bezeichnen wir, daß $f(x)$ ein Teiler von $g(x)$ mod p ist.

Wir brauchen zur Richtigkeit von Hilfssatz 6 nur noch zu zeigen, daß in (36)

$$k_1+\ldots+k_{p-1}\leqq 1,\quad k_0\leqq 2 \tag{38}$$

gilt.

Wäre (38_1) falsch, so gilt

$$(x-c)(x-d)\,|\,F(x)F_0(x)\ (\mathrm{mod}\ p) \qquad (p\nmid c,d) \tag{39}$$

mit geeigneten c, d. Je nachdem $c\equiv d\ (\mathrm{mod}\ p)$ oder $c\not\equiv d\ (\mathrm{mod}\ p)$ gilt, setze man

$$\varphi(x)=x^{2p-2}-2x^{p-1}=(x^{p-1}-1)^2-1 \text{ bzw. } \varphi(x)=x^{p-1}.$$

Da entsprechend

$$\varphi(x)\equiv-1\ (\mathrm{mod}\,(x-c)(x-d),p) \text{ bzw. } \varphi(x)\equiv 1\ (\mathrm{mod}\,(x-c)(x-d),p)$$

gilt, so folgt aus (35), (39) offenbar

$$\zeta(-1)\equiv -x\ (\mathrm{mod}\,(x-c)(x-d),p) \text{ bzw. } \zeta(1)\equiv x\ (\mathrm{mod}\,(x-c)(x-d),p).$$

Beide sind unmöglich, womit wir (38_1) bewiesen haben.

Wäre (38_2) falsch, also $k_0\geqq 3$, so folgt aus (35), (36) offenbar

$$\zeta(\varphi(x))\equiv x\varphi(x)\ (\mathrm{mod}\ x^4,p).$$

Insbesondere für $\varphi(x)=x^2$ ergibt sich

$$\zeta(x^2)\equiv x^3\ (\mathrm{mod}\ x^4,p).$$

Dies ist unmöglich, womit wir (38_2) und auch Hilfssatz 6 bewiesen haben.

HILFSSATZ 7. *Ist das Ideal* $\mathfrak{a}$ *in* (26) *zulässig, so gilt im Fall* $F(x)=x-a$

$$\mathfrak{P}=(F_0(x),p), \tag{40}$$

im Fall $F(x)=1$ *gilt entweder* (40) *oder mit einem passenden* a

$$\mathfrak{P}=(x-a,p^n) \qquad (n\geqq 2) \tag{41}$$

oder

$$\mathfrak{P}=(F_0(x),p(x-a),p^n) \qquad (p^{n-1}\,|\,F_0(a);\ n\geqq 2), \tag{42}$$

wobei $F_0(x)$ *ein Hauptpolynom ist, dessen Grad in* (40) *und* (42) $\geqq 1$ *bzw.* $\geqq 2$ *ist.*[8]

Aus Hilfssatz 2 folgt nämlich für $\varphi(x)=px$, daß

$$\zeta(px)\equiv px^2\ (\mathrm{mod}\ xF(x)\mathfrak{P})$$

mit einem $\zeta(x)\,(\in xI[x])$ lösbar ist. Da mit passendem c

$$\zeta(px)\equiv cpx\ (\mathrm{mod}\ p^2)$$

gilt, so folgt

$$cpx\equiv px^2\ (\mathrm{mod}\ xF(x)\mathfrak{P},p^2),$$

also

$$p(x-c)\equiv 0\ (\mathrm{mod}\ F(x)\mathfrak{P},p^2). \tag{43}$$

[8] Selbstverständlich hat a in (41), (42) nichts mit dem a in (26) zu tun.

Nach (28), (29_2) gilt

$$p^2 | F_i(x) \qquad (2 \leqq i \leqq t).$$

Hiernach und nach (31) folgt aus (43) eine Kongruenz

$$p(x-c) \equiv F(x)(f_0(x)F_0(x)+f_1(x)F_1(x)) \pmod{p^2} \tag{44}$$

mit geeigneten Polynomen $f_0(x), f_1(x)$, so daß der Grad von $f_1(x)F_1(x)$ kleiner ist als der von $F_0(x)$. Da ferner $F(x), F_0(x)$ Hauptpolynome sind und nach (28), (29_2)

$$p | F_1(x)$$

gilt, so folgt aus (44) ($p | f_0(x)$ und)

$$x-c \equiv F(x)\left(g(x)F_0(x)+f_1(x)\frac{F_1(x)}{p}\right) \pmod{p} \tag{45}$$

mit einem $g(x)$ ($\in I[x]$).

Im Fall $F(x)=x-a$ folgt aus (45) ($c \equiv a \pmod{p}$ und)

$$1 \equiv g(x)F_0(x)+f_1(x)\frac{F_1(x)}{p} \pmod{p}.$$

Da das zweite Produkt rechts von kleinerem Grade als $F_0(x)$ ist, so ergibt sich ($p | g(x)$ und)

$$1 \equiv f_1(x)\frac{F_1(x)}{p} \pmod{p}.$$

Hiernach gilt wegen (28) offenbar $F_1(x)=p$, $n_1=0$. Aus (29_1) folgt auch $t=1$. Man sieht also nach (27), daß jetzt (40) gilt.

Im Fall $F(x)=1$ hat man nach (45)

$$x-c \equiv g(x)F_0(x)+f_1(x)\frac{F_1(x)}{p} \pmod{p}, \tag{46}$$

wobei — wir wiederholen — das zweite Produkt rechts von kleinerem Grade ist als $F_0(x)$.

Erstens wenn $F_0(x)$ vom Grade 1 ist, so muß $F_1(x)$ konstant sein. Dies bedeutet nach (28), daß $F_1(x)=p^{d_1}$ ($d_1 \geqq 1$) ist, ferner folgt aus (29) $t=1$. Folglich ist nach (27) $\mathfrak{P}=(F_0(x), p^{d_1})$. Da $F_0(x)$ ein Hauptpolynom vom Grade 1, also von der Form $x-a$ ist, so kommt jetzt $\mathfrak{P}$ unter den Idealen (40), (41) vor.

Zweitens wenn $F_0(x)$ vom Grade $\geqq 2$ ist, so folgt aus (46) notwendig $p | g(x)$,

$$x-c \equiv f_1(x)\frac{F_1(x)}{p} \pmod{p}.$$

Hiernach ist entweder $F_1(x)=p$ oder $F_1(x)=p(x-a)$. Wegen (27) bis (29) muß entsprechend ($t=1$ und)

$$\mathfrak{P}=(F_0(x), p)$$

bzw. ($t=2$ und)

$$\mathfrak{P}=(F_0(x), p(x-a), p^{d_2})$$

sein. Im letzeren Fall gilt auch $d_1=1, F_1(a)=0$, woraus nach (33)

$$F_0(a)\equiv 0 \pmod{p^{d_2-1}}$$

folgt. Wir haben gewonnen, daß jetzt (40) oder (42) gilt. Hilfssatz 7 haben wir bewiesen.

Wir zeigen nunmehr den folgenden:

HILFSSATZ 8. *Das Ideal* $\mathfrak{a}=F(x)\mathfrak{P}$ *von* $I[x]$ *mit dem Hauptidealfaktor* $F(x)$ *und der (einzigen) p-Komponente* $\mathfrak{P}$ *kann nur dann zulässig sein, wenn* e r s t e n s $F(x)=x-a$ ($p \nmid a$) *gilt und* $\mathfrak{P}$ *eins der folgenden Ideale ist:*

Fall 1. $\mathfrak{P}=(x^2, p)$,

Fall 2. $\mathfrak{P}=(x, p)$,

z w e i t e n s $F(x)=x-a$ ($p|a$) *gilt und* $\mathfrak{P}$ *eins der folgenden Ideale ist:*

Fall 3. $\mathfrak{P}=(x(x-b), p) \qquad (p \nmid b)$,

Fall 4. $\mathfrak{P}=(x-b, p)$,

d r i t t e n s $F(x)=1$ *gilt und* $\mathfrak{P}$ *eins der folgenden Ideale ist:*

Fall 5. $\mathfrak{P}=(x^2(x-a), p) \qquad (p \nmid a)$,

Fall 6. $\mathfrak{P}=(x(x-a), p)$,

Fall 7. $\mathfrak{P}=(x-a, p^n)$,

Fall 8. $\mathfrak{P}=(x^2(x-b)-a^2(a-b)-p^{n-1}c, p\; x-a), p^n)$
$(p \nmid b;\; p|a(a-b);\; n \geqq 2)$,

Fall 9. $\mathfrak{P}=(x(x-b)-a(a-b)-p^{n-1}c, p(x-a), p^n)$
$(p|a(a-b),\; n \geqq 2)$.

Nach Hilfssatz 3 kann nämlich nur $F(x)=x-a$ oder $F(x)=1$ sein. Wenn $F(x)=x-a$ und $p \nmid a$ gilt, so ist nach Hilfssatz 7 nur

$$\mathfrak{P}=(F_0(x), p)$$

möglich, ferner folgt aus Hilfssatz 6, daß

$$F_0(x)\,|\,x^2 \pmod{p}$$

gelten muß. Da $F_0(x)$ jetzt nur mod p in Betracht kommt, so darf $F_0(x)|x^2$ also $F_0(x)=x^2$ oder $F_0(x)=x$ gesetzt werden. Das sind eben die Fälle 1, 2.

Wenn $F(x)=x-a$ und $p|a$ gilt, so ist der einzige Unterschied, daß jetzt

$$F_0(x)\,|\,x(x-b) \pmod{p} \qquad (p \nmid b)$$

gelten muß, weshalb man jetzt $F_0(x)|x(x-b)$ setzen darf (mit einem anderen b). Alle Möglichkeiten sind jetzt durch $F_0(x)=x(x-b)\,(p \nmid b)$ und $F_0(x)=x-b$ erschöpft. Somit sind jetzt eben die Fälle 3, 4 enstanden.

Wenn $F(x)=1$ gilt, so sind nach Hilfssatz 7 nur

$$\mathfrak{P}=(F_0(x), p),$$
$$\mathfrak{P}=(x-a, p^n) \qquad (n \geqq 2),$$
$$\mathfrak{P}=(F_0(x), p(x-a), p^n) \qquad (p^{n-1} | F_0(a);\ n \geqq 2)$$

möglich.

Im ersten Fall kommt $F_0(x)$ nur mod p in Betracht, weshalb man nach Hilfssatz 6

$$F_0(x) | x^2(x-b) \qquad (p \nmid b)$$

setzen darf, wodurch offenbar die Fälle 5, 6, 7 entstehen (insbesondere Fall 7 nur mit $n=1$).

Im zweiten Fall haben wir es mit Fall 7 (für $n \geqq 2$) zu tun.

Im dritten Fall gilt nach Hilfssatz 6

$$F_0(x) | x^2(x-b) \pmod{p} \qquad (p \nmid b).$$

Da jetzt der Grad von $F_0(x)$ nur 3 oder 2 sein kann, so gilt entsprechend

$$F_0(x)=x^2(x-b)+pf(x) \quad (p \nmid b) \text{ bzw. } F_0(x)=x(x-b)+pg(x)$$

mit je einem Polynom $f(x), g(x)\, (\in I[x])$ vom Grade $\leqq 2$ bzw. $\leqq 1$. Ferner gilt jetzt $p(x-a) \in \mathfrak{P}$, somit kommen $f(x), g(x)$ nur mod $(x-a)$ in Betracht. Deshalb darf

$$F_0(x)=x^2(x-b)+pc_1 \quad (p \nmid b) \text{ bzw. } F_0(x)=x(x-b)+pc_2$$

gesetzt werden. Wegen $p^{n-1} | F_0(a)$ gilt dabei

$$pc_1 \equiv -a^2(a-b) \pmod{p^{n-1}} \text{ bzw. } pc_2 \equiv -a(a-b) \pmod{p^{n-1}},$$

ferner gilt in beiden dieser Fälle notwendig auch $p | a(a-b)$. Wir sehen somit, daß jetzt die Fälle 8, 9 entstehen. Hilfssatz 8 haben wir bewiesen.

§ 4. Schluß des Beweises

Wir haben nur noch zu untersuchen, welche Ideale $\mathfrak{a}$ im Hilfssatz 8 wirklich zulässig sind. Hierzu brauchen wir noch den folgenden:

HILFSSATZ 9. *Es seien gegeben ein Ideal* $\mathfrak{a}$ *von* $I[x]$ *mit*

(47) $$\mathfrak{a} \supseteqq (x^k - bx^{k-1} - c, p(x-a)) \qquad (p | c;\ k \geqq 2)$$

und ein Polynom in $xI[x]$

(48) $$\varphi(x)=s_k x^k + \ldots + s_1 x = x\psi(x),$$

wobei also

(49) $$\psi(x)=s_k x^{k-1} + \ldots + s_1$$

ist. Man setze

(50) $$A=\psi(a), \quad B=\psi(b).$$

Dann gilt[9]

$$(51)\qquad \varphi(x)^j \equiv b^{j-k} B^j x^k + c\,\frac{a^{j-k}A^j - b^{j-k}B^j}{a-b}\,x \pmod{x\mathfrak{a}} \qquad (j \geqq k),$$

$$(52)\quad \varphi(x)^{k-1} \equiv \frac{B^{k-1}-s_1^{k-1}}{b}\,x^k + s_1^{k-1}x^{k-1} + c\,\frac{b(A^{k-1}-s_1^{k-1})-a(B^{k-1}-s_1^{k-1})}{ab(a-b)}\,x \pmod{x\mathfrak{a}},$$

$$(53)\qquad x\varphi(x) \equiv (s_k b + s_{k-1})x^k + s_{k-2}x^{k-1} + \ldots + s_1 x^2 + c s_k x \pmod{x\mathfrak{a}}.$$

Zunächst zeigen wir hierzu

$$(54)\qquad x^i \equiv b^{i-k}x^k + c\,\frac{a^{i-k}-b^{i-k}}{a-b}\,x \pmod{x\mathfrak{a}} \qquad (i \geqq k).$$

Aus (47) folgt nämlich

$$(55)\qquad x^{k+1} \equiv bx^k + cx \pmod{x\mathfrak{a}}.$$

Ferner folgt aus (47) $px \equiv pa \pmod{\mathfrak{a}}$, also $px^2 \equiv pax \pmod{x\mathfrak{a}}$. Wegen $p|c$ folgt hieraus

$$(56)\qquad cx^2 \equiv cax \pmod{x\mathfrak{a}}.$$

Nun sehen wir, daß (54) für $i = k$ trivial ist. Multiplizieren wir (54) mit x und setzen dann rechts (55), (56) ein, so entsteht

$$x^{i+1} \equiv b^{i-k}(bx^k + cx) + c\,\frac{a^{i-k}-b^{i-k}}{a-b}\,ax \pmod{x\mathfrak{a}}.$$

Nach leichter Umformung sehen wir, daß dies dasselbe ist, wie (54) für $i+1$ statt i. Wir haben somit (54) mit Induktion bewiesen.

Nunmehr beweisen wir (51). Aus (48) folgt nach dem Polynomialsatz

$$(57)\qquad \varphi(x)^j = \sum (i_1 \ldots i_k) s_1^{i_1} \ldots s_k^{i_k}\, x^i,$$

wobei man über die $i_1, \ldots, i_k (\geqq 0)$ mit $i_1 + \ldots + i_k = j$ zu summieren hat und

$$(i_1 \ldots i_k) = \frac{j!}{i_1! \ldots i_k!}, \qquad i = i_1 + 2i_2 + \ldots + ki_k$$

bezeichnet. Offenbar ist $i \geqq j$. Wenn also $j \geqq k$ ist, so gilt noch mehr $i \geqq k$. Hiernach folgt aus (57), (54)

$$(58)\quad \varphi(x)^j \equiv \sum (i_1 \ldots i_k) s_1^{i_1} \ldots s_k^{i_k} \left(b^{i-k}x^k + c\,\frac{a^{i-k}-b^{i-k}}{a-b}\,x \right) \pmod{x\mathfrak{a}} \quad (j \geqq k).$$

Andererseits setzen wir in (57) $x = a$ ein. Wegen (48), (50_1) entsteht

$$a^j A^j = \sum (i_1 \ldots i_k) s_1^{i_1} \ldots s_k^{i_k}\, a^i.$$

Ähnliches gilt auch mit b, B statt a, A. Nach Einsetzung in (58) bekommen wir eben (51).

[9] Jeder Nenner in den Formeln (51), (52) ist ein Teiler des Zählers, wenn man in diesen a und b als Unbestimmte auffaßt. Wir denken in solchen Fällen die Kürzung mit dem Nenner stets ausgeführt, so daß ein Polynom (in a und b) entsteht. Nach dieser Bemerkung sind (51), (52) auch für $a = b$ sinnvoll.

Der Beweis von (52) entsteht durch eine leichte Modifikation des vorigen. Betrachten wir nämlich den Fall $j=k-1$ von (57). Das zu

(59) $$i_1=k-1,\quad i_2=\ldots=i_k=0$$

gehörende Glied der rechten Seite ist

(60) $$s_1^{k-1}x^{k-1},$$

und die übrigen Glieder sind vom Grade $\geqq k$. Wird auf die letzteren (54) angewendet, so bekommt man, daß $\varphi(x)^{k-1}-s_1^{k-1}x^{k-1}$ mod $x\mathfrak{a}$ kongruent derjenigen Summe ist, die aus (58) durch Weglassen des zu (59) gehörenden Summanden

(61) $$s_1^{k-1}\left(b^{-1}x^k+c\frac{a^{-1}-b^{-1}}{a-b}x\right)$$

entsteht. Folglich entsteht aus (51) eine richtige Kongruenz, wenn man $j=k-1$ einsetzt, aber links und rechts (60) bzw. (61) subtrahiert. So ergibt sich eben (52).

Endlich folgt (53) sofort aus (48) und dem Fall $i=k+1$ von (54). Hilfssatz 9 haben wir bewiesen.

Wir haben nur noch zu überprüfen, welche der in den Fällen 1 bis 9 des Hilfssatzes 8 angegebenen Ideale $\mathfrak{a}=F(x)\mathfrak{P}$ zulässig sind. Es wird sich herausstellen, daß ein solches $\mathfrak{a}=F(x)\mathfrak{P}$ dann und nur dann zulässig ist, wenn $F(x)=x-a$ und $\mathfrak{P}$ eins der Ideale (1), (2) bzw. $F(x)=1$ und $\mathfrak{P}$ eins der Ideale (3) bis (7) ist. Hierdurch wird wegen der Hilfssätze 3, 4, 5 der Beweis des Satzes erbracht.

Wir schicken noch als unmittelbare Folgerung von Hilfssatz 2 voran, daß jedes Ideal von $I[x]$, das ein zulässiges Ideal enthält, selbst zulässig ist. Umgekehrt, ein nicht zulässiges Ideal von $I[x]$ enthält also nur nicht zulässige Ideale.

Fälle 1, 2

Wir zeigen, daß $\mathfrak{a}$ in den Fällen 1, 2 von Hilfssatz 8 nicht zulässig ist. Da jedes Ideal $\mathfrak{a}$ des Falles 1 in einem Ideal $\mathfrak{a}$ des Falles 2 enthalten ist, so genügt es wegen der eben gemachten Bemerkung nur die auf Fall 2 bezügliche Behauptung zu beweisen.

Wir haben also zu beweisen, daß das Ideal

(62) $$\mathfrak{a}=(x-a)(x,p)\qquad (p\nmid a)$$

nicht zulässig ist. Zu diesem Zweck setzen wir

(63) $$\varphi(x)=x^2-ax+px.$$

Auf Grund von Hilfssatz 2 genügt es, wenn wir aus der Annahme

(64) $$\xi(\varphi(x))\equiv x\varphi(x)\ (\mathrm{mod}\ x\mathfrak{a})\qquad (\xi(x)\in xI[x])$$

zu einem Widerspruch kommen.

Wegen

$$x\mathfrak{a}=(x^3-ax^2,p(x^2-ax))$$

folgt aus (63)

$$x\varphi(x) \equiv px^2 \equiv pax \pmod{x\mathfrak{a}}. \tag{65}$$

$$\varphi(x)^2 \equiv (x-a+p)x\varphi(x) \equiv (x-a+p)pax \equiv p^2ax \pmod{x\mathfrak{a}}.$$

Aus diesen beiden folgt leicht durch Induktion, daß allgemein

$$\varphi(x)^k \equiv c_k p^2 ax \pmod{x\mathfrak{a}} \qquad (k \geqq 2)$$

gilt, wobei die c_k ganze Zahlen sind. Hiernach folgt aus (64) eine Kongruenz

$$\varphi(x)z_1 + p^2axz_2 \equiv x\varphi(x) \pmod{x\mathfrak{a}}$$

mit irgendwelchen ganzen Zahlen z_1, z_2. Wegen (62), (63), (65) ergibt sich hieraus nach Dividieren durch x

$$(x-a+p)z_1 + p^2az_2 \equiv pa \pmod{(x-a)(x,p)}. \tag{66}$$

Nach Einsetzung von a statt x entsteht die Gleichung

$$pz_1 + p^2az_2 = pa. \tag{67}$$

Wird dies aus (66) subtrahiert, so ergibt sich nach Dividieren durch $x-a$

$$z_1 \equiv 0 \pmod{x,p}.$$

Dies bedeutet $p|z_1$, somit folgt aus (67) der Widerspruch $p|a$. Die Behauptung haben wir bewiesen.

Fälle 3, 4

Wir zeigen, daß $\mathfrak{a}$ in den Fällen 3, 4 von Hilfssatz 8 zulässig ist. Da jedes Ideal $\mathfrak{a}$ des Falles 4 ein Ideal $\mathfrak{a}$ des Falles 3 enthält, so genügt es nur die auf Fall 3 bezügliche Behauptung zu beweisen.

Wir haben also zu beweisen, daß das Ideal

$$\mathfrak{a} = (x-a)(x(x-b),p) \qquad (p|a;\ p \nmid b) \tag{68}$$

zulässig ist.

Es gilt

$$\mathfrak{a} = (x^2(x-b) - ax(x-b),\, p(x-a)).$$

Ferner gilt $px \equiv pa \pmod{\mathfrak{a}}$, also wegen $p|a$ auch $ax \equiv a^2 \pmod{\mathfrak{a}}$ und $ax^2 \equiv a^3 \pmod{\mathfrak{a}}$. Somit haben wir

$$\mathfrak{a} = (x^2(x-b) - a^2(a-b),\, p(x-a)) \qquad (p|a;\ p \nmid b). \tag{69}$$

Hiernach erfüllt $\mathfrak{a}$ die Bedingung (47) (mit $k=3$, $c=a^2(a-b)$), weshalb sich aus Hilfssatz 9 folgendes ergibt. Wird

$$\varphi(x) = qx^3 + rx^2 + sx, \tag{70}$$

$$A = qa^2 + ra + s, \quad B = qb^2 + rb + s \tag{71}$$

gesetzt, so gilt:

$$\varphi(x)^j \equiv b^{j-3}B^jx^3 + (a^{j-1}A^j - a^2b^{j-3}B^j)x \pmod{x\mathfrak{a}} \qquad (j \geqq 3), \tag{72}$$

$$\varphi(x)^2 \equiv \frac{B^2-s^2}{b}x^3 + s^2x^2 + \frac{a}{b}[b(A^2-s^2) - a(B^2-s^2)]x \pmod{x\mathfrak{a}}, \tag{73}$$

$$x\varphi(x) \equiv (qb+r)x^3 + sx^2 + qa^2(a-b)x \pmod{x\mathfrak{a}}. \tag{74}$$

Da $x\mathfrak{a}$ nach (68) ein Hauptpolynom 4-ten Grades enthält, so enthält jede Restklasse in $xI[x]/x\mathfrak{a}$ ein Polynom von der Form (70). Um also die Zulässigkeit von $\mathfrak{a}$ zu beweisen, genügt es nach Hilfssatz 2 (vgl. (9″)) die Lösbarkeit von

$$\sum_{j=1}^{m} \varphi(x)^j z_j \equiv x\varphi(x) \pmod{x\mathfrak{a}}$$

auszuweisen. Nach Einsetzung von (70), (72), (73), (74) und Dividieren durch x hat man

$$(75)\quad (qx^2+rx+s)z_1+\left\{\frac{B^2-s^2}{b}x^2+s^2x+\frac{a}{b}[b(A^2-s^2)-a(B^2-s^2)]\right\}z_2+$$

$$+\sum_{j=3}^{m}\{b^{j-3}B^jx^2+(a^{j-1}A^j-a^2b^{j-3}B^j\}z_j\equiv(qb+r)x^2+sx+qa^2(a-b) \pmod{\mathfrak{a}}.$$

Eine Kongruenz

$$f(x)\equiv g(x) \pmod{\mathfrak{a}} \qquad (f(x), g(x)\in I[x])$$

bedeutet wegen (68) offenbar dasselbe wie

$$f(a)=g(a)$$

und

$$\frac{f(x)-f(a)}{x-a}\equiv\frac{g(x)-g(a)}{x-a} \pmod{x^2-bx, p}.$$

Die Anwendung dieses einfachen Prinzips verwandelt die Kongruenz (75) (unter Berücksichtigung von (71_1)) in die Gleichung und Kongruenz:

$$(76)\qquad Az_1+aA^2z_2+\sum_{j=3}^{m}a^{j-1}A^jz_j=aA,$$

$$(77)\quad \{q(x+a)+r\}z_1+\left\{\frac{B^2-s^2}{b}(x+a)+s^2\right\}z_2+\sum_{j=3}^{m}b^{j-3}B^j(x+a)z_j\equiv$$

$$\equiv(qb+r)(x+a)+s \pmod{x^2-bx, p}.$$

Wir haben somit zu zeigen, daß dieses System (76), (77) lösbar ist.

Die Gleichung (76) schreibt sich einfacher so:

$$(78)\qquad A\left(-a+\sum_{j=1}^{m}a^{j-1}A^{j-1}z_j\right)=0.$$

In der Kongruenz (77) sind beide Seiten linear in x, weshalb „(mod x^2-bx, p)“ sich durch „(mod p)“ ersetzen läßt. Nach Koeffizientenvergleichung zerfällt also (77) wegen $p|a$ in die zwei Kongruenzen

$$qz_1+\frac{B^2-s^2}{b}z_2+\sum_{j=3}^{m}b^{j-3}B^jz_j\equiv qb+r \pmod{p},$$

$$(79)\qquad rz_1+s^2z_2\equiv s \pmod{p}.$$

Die erste dieser Kongruenzen läßt sich wegen $p\nmid b$ durch b multiplizieren.

Wird dann noch (79) addiert, so entsteht nach (71_2)

$$(80) \qquad (qb+r)z_1+\sum_{j=2}^{m} b^{j-2}B^j z_j \equiv B \pmod p.$$

Wir haben also zu beweisen, daß das System (78) bis (80) lösbar ist. Wir zeigen sogar, daß das mit $m=3$ möglich ist.

Für $m=3$ lauten (78) bis (80) so:

$$(81) \qquad A(-a+z_1+aAz_2+a^2A^2z_3)=0,$$

$$(82) \qquad rz_1+s^2z_2\equiv s \pmod p,$$

$$(83) \qquad (qb+r)z_1+B^2z_2+bB^3z_3\equiv B \pmod p.$$

Man nehme für z_2, z_3 vorläufig beliebige ganze Zahlen und bestimme z_1 so, daß der zweite Faktor in (81) verschwindet. Dieses z_1 ist wegen $p|a$ durch p teilbar. Damit also auch (82), (83) erfüllt sind, hat man nur z_2, z_3 so zu bestimmen, daß

$$s^2z_2\equiv s \pmod p), \qquad B(Bz_2+bB^2z_3-1)\equiv 0 \pmod p)$$

gelten. Die erste Kongruenz hat für jedes s eine Lösung z_2. Die zweite läßt wegen $p\nmid b$ bei jedem z_2 und B offenbar eine Lösung z_3 zu. Wir haben die Behauptung bewiesen.

Fälle 5, 6

Wir zeigen, daß $\mathfrak{a}$ in den Fällen 5, 6 von Hilfssatz 8 zulässig ist. Da jedes Ideal $\mathfrak{a}$ des Falles 6 ein Ideal $\mathfrak{a}$ des Falles 5 enthält, so genügt es nur die auf Fall 5 bezügliche Behauptung zu beweisen.

Wir haben also zu beweisen, daß das Ideal

$$\mathfrak{a}=(x^2(x-a),p) \qquad (p\nmid a)$$

zulässig ist. Mit veränderter Bezeichnung handelt es sich um

$$\mathfrak{a}=(x^2(x-b),p) \qquad (p\nmid b).$$

Da in diesem das Ideal

$$x(x(x-b),p)$$

enthalten ist und letzteres der Spezialfall $a=0$ des Ideals (68) ist, von dem die Zulässigkeit eben bewiesen wurde, so ist die Behauptung richtig.

Fall 7

Wir zeigen, daß $\mathfrak{a}$ im Fall 7 von Hilfssatz 8 zulässig ist. Es handelt sich um

$$\mathfrak{a}=(x-a,p^n) \qquad (n\geqq 1).$$

Da $x\mathfrak{a}$ das Polynom x^2-ax enthält, so gibt es in jeder Restklasse von $xI[x]/x\mathfrak{a}$ ein Polynom von der Form

$$\varphi(x)=sx.$$

Wegen Hilfssatz 2 genügt es also zu beweisen, daß

$$\zeta(sx)\equiv sx^2 \pmod{x\mathfrak{a}} \qquad (\zeta(x)\in xI[x])$$

gelöst werden kann. Offenbar ist $\zeta(x) = ax$ eine Lösung, womit die Behauptung bewiesen ist.

Fall 8

Im Fall 8 von Hilfssatz 8 handelt es sich um das Ideal

(84) $\mathfrak{a} = (x^2(x-b) - a^2(a-b) - p^{n-1}c, p(x-a), p^n) \quad (b \not| p;\ p|a(a-b);\ n \geqq 2).$

Wir zeigen, daß dies dann und nur dann zulässig ist, wenn

(85) $$p|a,\ p|c$$

gilt, und bemerken gleich, daß sich dann (84) auch in der Form

(86) $$\mathfrak{a} = (x(x-a)(x-b), p(x-a), p^n) \qquad (p \not| b;\ p|a(a-b);\ n \geqq 2).$$

schreiben läßt.

Wegen (85_2) läßt sich nämlich das Glied $p^{n-1}c$ in (84) streichen, ferner gilt wegen (85_1) und $pa \equiv px \pmod{\mathfrak{a}}$ offenbar

$$a^2 \equiv ax \pmod{\mathfrak{a}}, \qquad a^3 \equiv ax^2 \pmod{\mathfrak{a}},$$

also

$$x^2(x-b) - a^2(a-b) \equiv x^2(x-b) - ax(x-b) \equiv x(x-a)(x-b) \pmod{\mathfrak{a}},$$

woraus die Richtigkeit der Bemerkung folgt.

Da $x\mathfrak{a}$ ein Hauptpolynom 4-ten Grades enthält, so lassen sich die Restklassen von $xI[x]\,x\mathfrak{a}$ durch die Polynome von der Form

(87) $$\varphi(x) = qx^3 + rx^2 + sx$$

repräsentieren. Wegen Hilfssatz 2 genügt es also zu beweisen, daß

(88) $$\sum_{j=1}^{m} \varphi(x)^j z_j \equiv x\varphi(x) \pmod{x\mathfrak{a}}$$

dann und nur dann für die sämtlichen $\varphi(x)$ lösbar ist, wenn (85) gilt.

Es gilt nach (84)

(89) $$\mathfrak{a} \supset (x^2(x-b) - a^2(a-b) - p^{n-1}c, p(x-a)).$$

Wegen

$$p|a^2(a-b) + p^{n-1}c$$

läßt sich also Hilfssatz 9 anwenden, so daß man $a^2(a-b) + p^{n-1}c$ für c einsetzt. Hiernach gelten mit der Bezeichnung

(90) $$A = qa^2 + ra + s, \quad B = qb^2 + rb + s$$

die Kongruenzen

(91) $$\varphi(x)^j \equiv b^{j-3}B^j x^3 + [a^2(a-b) + p^{n-1}c]\,\frac{a^{j-3}A^j - b^{j-3}B^j}{a-b}\,x \pmod{x\mathfrak{a}}\ (j \geqq 3),$$

(92) $$\varphi(x)^2 \equiv \frac{B^2 - s^2}{b}x^3 + s^2x^2 + \\ + [a^2(a-b) + p^{n-1}c]\,\frac{b(A^2 - s^2) - a(B^2 - s^2)}{ab(a-b)}\,x \pmod{x\mathfrak{a}},$$

(93) $$x\varphi(x) \equiv (qb + r)x^3 + sx^2 + q[a^2(a-b) + p^{n-1}c]x \pmod{x\mathfrak{a}}.$$

Werden diese und (87) in (88) links und rechts eingesetzt, so entsteht eine Kongruenz von der Form

$$(94)\qquad ux^3+vx^2+wx\equiv u_1x^3+v_1x^2+w_1x \pmod{x\mathfrak{a}}.$$

Bevor wir diese Einsetzung ausführen, bemerken wir, daß (94) gleichbedeutend mit

$$(95)\qquad u\equiv u_1 \pmod p,\ v\equiv v_1 \pmod p,\ ua^2+va+w\equiv u_1a^2+v_1a+w_1 \pmod{p^n}$$

ist. Es genügt nämlich diese Bemerkung nur für den Fall $u_1=v_1=w_1=0$ zu beweisen. Dann bedeutet (84) dasselbe wie

$$(96)\qquad ux^2+vx+w\in\mathfrak{a}.$$

Der Vergleich mit (32) zeigt, daß $\mathfrak{a}$ in (84) durch eine Kronecker-Henselsche Basis dargestellt ist. Nach (31) besagt also (96) die Erfüllbarkeit von

$$ux^2+vx+w=p(x-a)(\alpha x+\beta)+p^n\gamma$$

mit ganzen Zahlen α, β, γ. Diese Forderung besagt offenbar dasselbe wie

$$u\equiv 0 \pmod p,\ v\equiv 0 \pmod p,\ ua^2+va+w\equiv 0 \pmod{p^n}.$$

Da dies mit dem Fall $u_1=v_1=w_1=0$ von (95) übereinstimmt, so ist unsere Bemerkung richtig.

Setzt man nunmehr (87), (91), (92), (93) in (88) wirklich ein und zerlegt die so entstehende Kongruenz von der Form (94) gleich in drei Kongruenzen nach dem Muster (95), so erhält man

$$qz_1+\frac{B^2-s^2}{b}z_2+\sum_{j=3}^{m}b^{j-3}B^jz_j\equiv qb+r \pmod p,$$

$$rz_1+s^2z_2\equiv s \pmod p,$$

$$(qa^2+ra+s)z_1+\left\{\frac{B^2-s^2}{b}a^2+s^2a+[a^2(a-b)+p^{n-1}c]\frac{b(A^2-s^2)-a(B^2-s^2)}{ab(a-b)}\right\}z_2+$$
$$+\sum_{j=3}^{m}\left\{b^{j-3}B^ja^2+[a^2(a-b)+p^{n-1}c]\frac{a^{j-3}A^j-b^{j-3}B^j}{a-b}\right\}z_j\equiv$$
$$\equiv(qb+r)a^2+sa+q[a^2(a-b)+p^{n-1}c] \pmod{p^n}.$$

Da $p\not| b$ ist, so darf man die erste dieser Kongruenzen mit b multiplizieren. Addiere man dann noch die zweite zu der ersten. Nach leichter Umformung der dritten lauten unsere drei Kongruenzen wegen (90) so:

$$(97)\qquad (qb+r)z_1+\sum_{j=2}^{m}b^{j-2}B^jz_j\equiv B \pmod p,$$

$$(98)\qquad rz_1+s^2z_2\equiv s \pmod p,$$

$$(99)\qquad \begin{aligned}&Az_1+\left\{aA^2+p^{n-1}c\frac{b(A^2-s^2)-a(B^2-s^2)}{ab(a-b)}\right\}z_2+\\&+\sum_{j=3}^{m}\left\{a^{j-1}A^j+p^{n-1}c\frac{a^{j-3}A^j-b^{j-3}B^j}{a-b}\right\}z_j\equiv aA+p^{n-1}cq \pmod{p^n}.\end{aligned}$$

Wir haben zu beweisen, daß dieses System (97) bis (99) genau nur im Fall (85) für alle q, r, s lösbar ist.

Erstens sei $p \nmid a$. Wegen der Bedingung $p \mid a(a-b)$ in (84) gilt dann

$$a \equiv b \pmod{p}.$$

Weiter unterscheiden wir zwei Fälle, je nachdem $p \mid c$ oder $p \nmid c$ ist.

Wenn $p \mid c$ ist, so setze man

$$q = 0, \; r = 1, \; s = -a+p.$$

Nach (90) gilt dann

$$A = p, \; B = b-a+p.$$

Die mit c multiplizierten Glieder fallen in (99) wegen $p \mid c$ heraus. Wird dann mit $A (= p)$ dividiert, so entsteht wegen $n \geqq 2$ immer noch eine Kongruenz $\bmod\, p$. Somit folgen aus (97), (99) (wegen $p \mid B$) der Reihe nach

$$z_1 \equiv 0 \pmod{p}, \qquad z_1 \equiv a \pmod{p}.$$

Das ist wegen der Annahme $p \nmid a$ unmöglich, weshalb jetzt (97) bis (99) unlösbar ist.

Wenn $p \nmid c$ ist, so setze man

$$q = 0, \; r = 1, \; s = -a.$$

Jetzt gilt nach (90)

$$A = 0, \; B = b-a.$$

Wegen $A = 0$ läßt sich (99) mit $p^{n-1}c$ dividieren, wodurch eine Kongruenz $\bmod\, p$ entsteht. Wird auch $p \mid B$ berücksichtigt, so gewinnt man aus (97) bis (99), wie leicht zu sehen, der Reihe nach

$$z_1 \equiv 0 \pmod{p}, \qquad z_1 + a^2 z_2 \equiv -a \pmod{p}, \qquad z_2 \equiv 0 \pmod{p}.$$

Diese sind wegen $p \nmid a$ unmöglich, weshalb (97) bis (99) auch jetzt unlösbar ist.

Zweitens sei $p \mid a$, $p \nmid c$. Wir setzen

$$q = -1, \; r = b, \; s = a(a-b).$$

Nach (90) gilt

$$A = 0, \; B = a(a-b).$$

Man sieht wegen $p \mid s, B$, daß auf der linken Seite von (99) alle Glieder durch p^n teilbar sind, dagegen die rechte Seite durch p^n nicht teilbar ist. Wir haben bewiesen, daß (85) notwendig ist, damit (97) bis (99) für alle q, r, s lösbar ist.

Umgekehrt nehmen wir jetzt (85) d. h. $p \mid a, p \mid c$ an. Wegen des letzteren fallen in (99) die mit c multiplizierten Glieder heraus. Dann werden beide Seiten durch A teilbar, somit ist (99) gleichbedeutend mit

$$(100) \qquad \sum_{j=1}^{m} (aA)^{j-1} z_j \equiv a \left(\bmod \frac{p^n}{(A, p^n)} \right).$$

Wir haben zu beweisen, daß das System (97), (98), (100) für alle q, r, s lösbar ist. Wieder unterscheiden wir zwei Fälle.

Erstens sei $p^n | A$. Jetzt ist (100) (als Kongruenz mod 1) identisch erfüllt, somit hat man es nur mit (97), (98) zu tun. Ferner gilt wegen (90_1) und $p|a$ notwendig $p|s$, also nach (90_2) $B \equiv (qb+r)b \pmod p$. Hiernach lauten (97), (98) so:

$$(qb+r)\left[-b+\sum_{j=1}^{m} b^{2j-2}(qb+r)^{j-1}z_j\right] \equiv 0 \pmod p,$$

$$rz_1 \equiv 0 \pmod p.$$

Wegen $q \nmid b$ läßt dieses System sogar schon für $m=2$ stets eine Lösung zu.

Zweitens sei $p^n \nmid A$. Wir sehen, daß z_1 in (100) den Koeffizienten 1 hat. Hieraus folgt, daß sich z_1 bei irgendwelchen $z_2, \ldots, z_m$ so bestimmen läßt, daß (100) erfüllt ist. Dabei muß wegen $p|a$ stets $z_1 \equiv 0 \pmod p$ ausfallen. Es genügt also zu zeigen, daß das System (97), (98) stets eine Lösung mit $z_1 = 0$ zuläßt. Nach wenig Umformung handelt es sich um

$$B\left[-1+\sum_{j=2}^{m} b^{j-2}B^{j-1}z_j\right] \equiv 0 \pmod p,$$

$$s^2 z_2 \equiv s \pmod p.$$

Wegen $p \nmid b$ ist dieses System sogar schon für $m=3$ stets lösbar. Die Behauptung haben wir bewiesen.

Fall 9

Im Fall 9 von Hilfssatz 8 handelt es sich um das Ideal

(101) $$\mathfrak{a} = (x(x-b)-a(a-b)-p^{n-1}c, p(x-a), p^n) \qquad (p|a(a-b);\, n \geqq 2).$$

Wir zeigen, daß dies dann und nur dann zulässig ist, wenn

(102) $$p|a$$

und im Fall $p|b$ auch

(103) $$p|c$$

gilt, ferner bemerken wir gleich, daß die diesen Bedingungen genügenden $\mathfrak{a}$ mit den Idealen

(104) $$\mathfrak{a} = ((x-a)(x-b)+p^{n-1}bc', p\, x-a), p^n) \qquad (p|a;\; n \geqq 2)$$

zusammenfallen (wobei c' eine beliebige ganze Zahl sein kann).

Da nämlich nach (101) $pa \equiv px \pmod{\mathfrak{a}}$ ist, so gilt im Fall (102) auch $a^2 \equiv ax \pmod{\mathfrak{a}}$, somit läßt sich dann für (101)

$$\mathfrak{a} = ((x-a)(x-b)-p^{n-1}c, p(x-a), p^n)$$

schreiben. Da das Glied $-p^{n-1}c$ nur mod p^n in Betracht kommt, so darf c von vornherein auf ein volles Restsystem mod p beschränkt werden, und die Bedingung (103) besagt, daß dabei im Falle $p|b$ nur $c \equiv 0 \pmod p$ zuzulassen ist. Man sieht hieraus die Richtigkeit der Bemerkung über (104) leicht ein.

Da $x\mathfrak{a}$ ein Hauptpolynom 3-ten Grades enthält, so lassen sich die Restklassen $x/[x]/x\mathfrak{a}$ durch die Polynome von der Form

(105) $$\varphi(x) = rx^2 + sx$$

repräsentieren. Wegen Hilfssatz 2 genügt es also zu beweisen, daß

$$(106)\qquad \sum_{j=1}^{m} \varphi(x)^j z_j \equiv x\varphi(x) \pmod{x\mathfrak{a}}$$

dann und nur dann für die sämtlichen $\varphi(x)$ in (105) lösbar ist, wenn (102) und im Falle $p|b$ auch noch (103) gilt.

Andererseits gilt nach (101)

$$(107)\qquad \mathfrak{a} \supset (x(x-b)-a(a-b)-p^{n-1}c,\ p(x-a)).$$

Wegen

$$p|a(a-b)+p^{n-1}c$$

läßt sich Hilfssatz 9 anwenden, so daß man $a(a-b)+p^{n-1}c$ für c einsetzt. Hiernach gelten mit der Bezeichnung

$$(108)\qquad A=ra+s,\ B=rb+s$$

die Kongruenzen

$$(109)\qquad \varphi(x)^j \equiv b^{j-2}B^j x^2+[a(a-b)+p^{n-1}c]\frac{a^{j-2}A^j-b^{j-2}B^j}{a-b}x \pmod{x\mathfrak{a}}\ (j\geqq 2),$$

$$(110)\qquad x\varphi(x) \equiv Bx^2+r[a(a-b)+p^{n-1}c]x \pmod{x\mathfrak{a}}.$$

Werden diese und (105) in (106) eingesetzt, so entsteht eine Kongruenz von der Form

$$ux^2+vx \equiv u_1x^2+v_1x \pmod{x\mathfrak{a}}.$$

Da dies mit

$$u \equiv u_1 \pmod{p}, \qquad ua+v \equiv u_1a+v_1 \pmod{p^n}$$

gleichbedeutend ist, was man so einsieht, wie die ähnliche Behauptung über (94), so zerlegt sich (106) nach der gesagten Einsetzung in die zwei Kongruenzen

$$(111)\qquad rz_1+\sum_{j=2}^{m} b^{j-2}B^j z_j \equiv B \pmod{p},$$

$$(ra+s)z_1+\sum_{j=2}^{m}\left\{ab^{j-2}B^j+[a(a-b)+p^{n-1}c]\frac{a^{j-2}A^j-b^{j-2}B^j}{a-b}\right\}z_j \equiv$$

$$\equiv aB+r[a(a-b)+p^{n-1}c] \pmod{p^n}.$$

Die letztere Kongruenz lautet nach leichter Umformung wegen (108) so:

$$(112)\qquad Az_1+\sum_{j=2}^{m}\left\{a^{j-1}A^j+p^{n-1}c\frac{a^{j-2}A^j-b^{j-2}B^j}{a-b}\right\}z_j \equiv aA+p^{n-1}cr \pmod{p^n}.$$

Wir haben zu beweisen, daß das System (111), (112) genau dann für alle r, s lösbar ist, wenn (102) und im Falle $p|b$ auch (103) gilt.

Zu diesem Zweck nehmen wir zuerst an, daß diese Bedingungen nicht erfüllt sind. Wir unterscheiden zwei Fälle.

Erstens sei $p\nmid a$. Wegen der Bedingung $p|a(a-b)$ in (101) gilt dann

$$a \equiv b \pmod{p}.$$

Weiter unterscheiden wir zwei Fälle, je nachdem $p|c$ oder $p\nmid c$ ist.

Wenn $p|c$ ist, so setze man

$$r=1,\ s=-a+p.$$

Nach (108) gilt dann

$$A=p,\ B=b-a+p.$$

Die mit c multiplizierten Glieder fallen in (112) wegen $p|A, B$ heraus. Wird dann mit A dividiert, so entsteht wegen $n\geqq 2$ immer noch eine Kongruenz mod p. Somit folgt aus (111), (112) wegen $p|B$ der Reihe nach

$$rz_1\equiv 0 \pmod p, \qquad z_1\equiv a \pmod p.$$

Dies ist wegen der Annahme $p\nmid a$ unmöglich, weshalb jetzt das System (111), (112) unlösbar ist.

Wenn $p\nmid c$ ist, so setze man

$$r=1,\ s=-a.$$

Nach (108) gilt dann

$$A=0,\ B=b-a.$$

Man sieht, daß alle Glieder der linken Seite von (112) durch p^n teilbar sind, dagegen die rechte Seite durch p^n nicht teilbar ist. Deshalb ist das System (111), (112) auch jetzt unlösbar.

Zweitens sei $p|a, p|b, p\nmid c$. Wir setzen

$$r=1,\ s=-a+p^{n-1}.$$

Nach (108) gilt dann

$$A=p^{n-1},\ B=b-a+p^{n-1}.$$

Aus (111) folgt wegen $p|B$

$$z_1\equiv 0 \pmod p.$$

Im diesem Falle sind — wie leicht zu sehen — alle Glieder der linken Seite von (112) durch p^n teilbar, dagegen ist die rechte durch p^n nicht teilbar. Wieder ist also das System (111), (112) unlösbar. Wir haben bewiesen, daß das System (111), (112) unlösbar ist, wenn entweder (102) oder im Falle $p|b$ (103) nicht erfüllt ist.

In folgendem nehmen wir an, daß (102) d. h.

$$p|a$$

und im Falle $p|b$ auch noch (103) d. h. $p|c$ gilt. Wir haben zu beweisen, daß dann das System (111), (112) für alle r, s lösbar, oder, was auf dasselbe hinauskommt, daß $\mathfrak{a}$ zulässig ist. Wieder unterscheiden wir zwei Fälle, je nachdem nämlich $p|c$ oder $p\nmid c$ ist.

Erstens sei $p|c$. Jetzt handelt es sich offenbar um den Fall $p|c'$ von (104), d. h. um

$$\mathfrak{a}=((x-a)(x-b), p(x-a), p^n).$$

Hierfür läßt sich insbesondere im Fall $p|b$ einfach

$$\mathfrak{a}=(x(x-a), p(x-a), p^n)$$

schreiben. Gleichgültig also ob $p\nmid b$ oder $p|b$ ist, enthält $\mathfrak{a}$ das Ideal (86).

Folglich ist a zulässig, weil die Zulässigkeit von (86) schon ausgewiesen wurde.

Zweitens sei $p \nmid c$, also wegen der Annahme auch $p \nmid b$. Wir unterscheiden zwei Fälle, je nachdem $p^n | A$ oder $p^n \nmid A$ ist.

Wenn $p^n | A$ ist, so ist nach (108) $p | s, B \equiv rb \pmod{p}$. Folglich lautet jetzt (111) so:

$$(113) \qquad r\left(-b + \sum_{j=1}^{m} (b^2 r)^{j-1} z_j\right) \equiv 0 \pmod{p}.$$

Ferner lassen sich die A enthaltenden Glieder in (112) streichen. Wird dann noch mit $p^{n-1}c$ dividiert, so entsteht (wegen $a-b \equiv -b \not\equiv 0 \pmod{p}$) die mit (112) gleichberechtigte Kongruenz

$$\sum_{j=2}^{m} b^{2j-3} r^j z_j \equiv r \pmod{p}.$$

Diese letztere ist offenbar für jedes r lösbar. Zu jeder Lösung läßt sich ein z_1 so bestimmen, daß auch (113) gilt, weil z_1 im zweiten Faktor von (113) mit dem Koeffizienten 1 vorkommt. Hiernach ist jetzt das System (111), (112) auch lösbar.

Wenn $p^n \nmid A$ ist, so bestimmen wir die ganzen Zahlen d, A' mit

$$(114) \qquad A = p^{n-1-d} A' \qquad (0 \leq d \leq n-1;\ p \nmid A').$$

Offenbar sind alle Glieder von (112) durch p^{n-1-d} teilbar. Wenn wir also beliebige (ganzzahlige) Werte für $z_2, \ldots, z_m$ einsetzen, so läßt sich z_1 stets so (ganzzahlig) bestimmen, daß (112) gilt. Um $z_1 \pmod{p}$ aus (112) zu berechnen, lösen wir die Kongruenz

$$(115) \qquad A' A'' \equiv c \pmod{p}$$

mit einer ganzen Zahl A''. Dann gewinnt man aus (112) (wegen $p | a$ und $a-b \equiv -b \not\equiv 0 \pmod{p}$) offenbar

$$(116) \qquad z_1 + p^d A'' \sum_{j=2}^{m} b^{j-3} B^j z_j \equiv p^d A'' r \pmod{p},$$

wobei b^{-1} eine ganze Zahl mit $bb^{-1} \equiv 1 \pmod{p}$ bezeichnet. Es genügt die Lösbarkeit des Systems (111), (116) auszuweisen, denn hieraus folgt nach obigem auch die des Systems (111), (112).

Bezeichne M die $2 \times m$-Matrix der Koeffizienten auf der linken Seite des Systems (111), (116), ferner bezeichne M' diejenige $2 \times (m+1)$-Matrix, die aus M so entsteht, daß man auch die konstanten Glieder von (111), (116) (als letzte Spalte) hinzunimmt. Bezeichne ϱ und ϱ' den Rang mod p von M bzw. M'.[10] Wir haben nur

$$\varrho = \varrho'$$

[10] Unter dem Rang mod p einer ganzzahligen Matrix verstehen wir den gewöhnlichen Rang dieser Matrix, indem wir die Elemente im Primkörper von p Elementen deuten.

zu beweisen, denn dies kommt nach allgemeinem Satz eben der Lösbarkeit von (111), (116) gleich.

Da stets $\varrho \leqq \varrho' \leqq 2$ und offenbar $1 \leqq \varrho$ ist, so brauchen wir nur zu zeigen, daß im Fall $\varrho = 1$ auch $\varrho' = 1$ gilt. Ein Blick auf M, M' zeigt, daß dies der folgenden Behauptung äquivalent ist: Aus

$$\begin{vmatrix} r & B^2 \\ 1 & p^d A'' B^2 b^{-1} \end{vmatrix} \equiv 0 \pmod{p} \tag{117}$$

folgt

$$\begin{vmatrix} r & B \\ 1 & p^d A'' r \end{vmatrix} \equiv 0 \pmod{p}. \tag{118}$$

Zum Beweis nehmen wir (117) an, und unterscheiden zwei Fälle.

Fall $p|B$. Ist auch $p|r$, so gilt (118). Ist $p \nmid r$, so gilt nach (108_2) $p \nmid s$, folglich nach (108_1) $p \nmid A$, also nach (114) $d > 0$, woraus wieder (118) folgt.

Fall $p \nmid B$. Nach (117) muß dann $p \nmid r$, $d = 0$ sein. Aus letzterem und (114) folgt $p|A$ und weiter hieraus nach (108) auch $p|s$, $B \equiv rb \not\equiv 0 \pmod{p}$. Wegen letzteres stimmt die zweite Elementenspalte in (117) mod p mit der B-fachen der zweiten Elementenspalte in (118) überein. Folglich gilt (118) auch jetzt. Die Behauptung haben wir bewiesen.

*

Zusammenfassend haben wir bekommen, daß unter den in den sämtlichen Fällen 1 bis 9 von Hilfssatz 8 angegebenen Idealen $\mathfrak{a} = F(x)\mathfrak{P}$ eben nur die folgenden zulässig sind. Die sämtlichen $\mathfrak{a}$ in den Fällen 3 bis 7. Ferner (aus den Fällen 8, 9) die Ideale (86) und (104). Man sieht hieraus und aus der nach dem Beweis von Hilfssatz 9 gemachten Bemerkung die Richtigkeit des Satzes ein.

(Eingegangen am 31. Dezember 1952.)

ПОЛНЫЕ ИДЕАЛЬНЫЕ КОЛЬЦА В ОБЩЕМ СМЫСЛЕ, I

Л. РЭДЭИ (Сегед)

(Резюме)

В качестве теоретико-кольцового аналога известных гамилтоновых групп, автор раньше определил полные идеальные кольца, как кольца, все подмодули которых суть идеалы. Если свойство быть идеалом требуется только от подколец, то получаются полные идеальные кольца в общем смысле. Класс этих колец также допускает исчерпывающее описание. Оказывается, что все элементы этих колец являются алгебраическими со степенью $\leqq 3$, т. е. что каждый такой элемент удовлетворяет алгебраическое уравнение степени $\leqq 4$ с целыми коэффициентами, первый коэффициент которого есть 1, а постоянный член равен нулю. Настоящая первая часть дает решение проблемы для колец порождаемых одним элементам, во второй же части на основе приведенных здесь результатов будем исследовать общий случай.

WERTVERTEILUNG BEI POLYNOMEN MIT LAUTER REELLEN NULLSTELLEN UND KOEFFIZIENTEN

Von

GYULA SZ.-NAGY (Szeged), Mitglied der Akademie

1. Bezeichnet $\mathrm{Re}\, f(z)$ den reellen Teil der Funktion $f(z)$ im Punkt z, so gilt der Satz

I. *Hat das Polynom*

$$(1)\quad f(z) = (z-a_1)(z-a_2)\cdots(z-a_n), \quad a_1 \geqq a_2 \geqq \cdots \geqq a_p \geqq a_{p+1} \geqq \cdots \geqq a_n$$

lauter reelle Nullstellen, bezeichnet D_p $(p = 1, 2, \ldots, n-1)$ den bezüglich der reellen Achse symmetrischen Deltoidbereich, der die Hauptdiagonale (a_p, a_{p+1}) und bei der Ecke a_p bzw. a_{p+1} den Winkel $\frac{\pi}{p}$ bzw. $\frac{\pi}{n-p}$ besitzt, und bezeichnet D_0 bzw. D_n den bezüglich der reellen Achse symmetrischen Winkelraum von der Öffnung $\frac{\pi}{n}$ mit dem Scheitel a_1 bzw. a_n, dessen Inneres keine Nullstelle des Polynoms $f(z)$ enthält, so besteht die Ungleichung

$$(-1)^k\, \mathrm{Re}\, f(z) = \mathrm{Re}\, (-1)^k f(z) \equiv \mathrm{Re}\, f_k(z) > 0 \qquad (k = 0, 1, \ldots, n)$$

in jedem inneren Punkte des Bereiches D_k $(k = 0, 1, 2, \ldots, n)$.

Die Transformation $Z = f(x+iy) = X + iY$ bildet also den Bereich D_k der komplexen z-Ebene auf einen Bereich der komplexen Z-Ebene oberhalb bzw. unterhalb der reellen Achse ab, je nachdem k eine gerade bzw. ungerade Zahl ist.

Dieser Satz ist für die auf der reellen Achse liegenden Punkte von D_k klar. Wegen der Gleichung $\mathrm{Re}\, f(x+iy) = \mathrm{Re}\, f(x-iy)$ genügt es ihn nur für die inneren Punkte von D_k oberhalb der reellen Achse zu beweisen.

Wir bezeichnen mit ζ die Ecke von D_p oberhalb der reellen Achse, mit φ_h bzw. ψ_j $(h = 1, 2, \ldots, p;\ j = p+1, p+2, \ldots, n)$ den Winkel des Dreiecks $\zeta a_h a_j$ bei der Ecke a_h bzw. a_j. Nach dem Satz der Außenwinkel eines Dreiecks bestehen also die Ungleichungen

$$(2)\quad \frac{\pi}{2p} = \varphi_p \geqq \varphi_h, \ \frac{\pi}{2(n-p)} = \psi_{p+1} \geqq \psi_j \qquad (h = 1, 2, \ldots, p;\ j = p+1, \ldots, n).$$

Das Gleichheitszeichen besteht hier nur im Falle $a_p = a_h$ bzw. $a_{p+1} = a_j$.

Für einen beliebigen inneren Punkt z ($\operatorname{Im} z > 0$) des Deltoidbereichs D_p bestehen offenbar die Ungleichungen

$$(3)\qquad -\frac{\pi}{2p} \leqq -\varphi_h < \operatorname{arc}(a_h - z) < 0, \quad 0 < \operatorname{arc}(z - a_j) < \psi_j \leqq \frac{\pi}{2(n-p)}\,.$$

$$(h = 1, 2, \ldots, p;\ j = p+1, p+2, \ldots, n).$$

Aus (3) folgt die Ungleichung

$$(4)\qquad -\frac{\pi}{2} < \sum_{h=1}^{p} \operatorname{arc}(a_h - z) + \sum_{j=p+1}^{n} \operatorname{arc}(z - a_j) = \operatorname{arc} f_p(z) < \frac{\pi}{2}\,.$$

Damit ist der Satz I für den Deltoidbereich D_p ($p = 1, 2, \ldots, n-1$) bewiesen; im Falle $|\operatorname{arc} f(z)| < \frac{\pi}{2}$ ist nämlich $\operatorname{Re} f_p(z) > 0$.

Für einen inneren Punkt ($\operatorname{Im} z > 0$) des Winkelraumes D_0 bzw. D_n bestehen die Ungleichungen

$$\frac{\pi}{2n} > \operatorname{arc}(z - a_1) \geqq \operatorname{arc}(z - a_2) \geqq \cdots \geqq \operatorname{arc}(z - a_n) > 0$$

und

$$\frac{\pi}{2} > \sum_{k=1}^{n} \operatorname{arc}(z - a_k) = \operatorname{arc} f_0(z) > 0$$

bzw.

$$-\frac{\pi}{2n} < \operatorname{arc}(a_n - z) \leqq \operatorname{arc}(a_{n-1} - z) \leqq \cdots \leqq \operatorname{arc}(a_1 - z) < 0$$

und

$$-\frac{\pi}{2} < \sum_{k=1}^{n} \operatorname{arc}(a_k - z) = \operatorname{arc} f_n(z) < 0.$$

Damit ist der Satz I bewiesen. Die Behauptung von Satz I gilt natürlich umso eher, wenn alle Deltoidbereiche D_p mit den Winkeln $\frac{\pi}{n}$ gebildet sind.

Der Satz I wird durch den folgenden ergänzt:

II. *Der Satz I ist scharf im folgenden Sinne:*

Ist das Deltoid D_p ein Teil eines Deltoids D_p^* mit der gemeinsamen Hauptdiagonalen (a_p, a_{p+1}) $(a_p > a_{p+1})$ und haben D_p und D_p^* keine Seite gemeinsam, so gibt es ein Polynom $f(z)$ n-ten Grades von der Form (1), *dessen Realteile in zwei inneren Punkten von D_p^* entgegengesetzte Vorzeichen besitzen. Enthält ein bezüglich der reellen Achse symmetrischer Winkelraum W den Winkelraum D_0 oder D_n im Innern, so enthält er im Innern immer zwei Punkte, in denen die Realteile eines passenden Polynoms $f(z)$ entgegengesetzte Vorzeichen besitzen.*

Der Winkel α des Deltoids D_p^* bei der Ecke a_p ist größer als $\frac{\pi}{p}$. Es gibt daher Winkel β und γ, für welche

$$\alpha > \beta > \frac{\pi}{p}, \quad 0 < \gamma < \frac{\pi}{n-p} \quad \text{und} \quad \pi + \frac{\pi}{2} > p\beta - (n-p)\gamma = \omega > \pi$$

sind. Dann liegt die dritte Ecke z_0 ($\mathrm{Im}\, z_0 > 0$) des Dreiecks, das die Grundlinie (a_p, a_{p+1}) und bei der Ecke a_p bzw. a_{p+1} den Winkel β bzw. γ besitzt, innerhalb des Deltoids D_p^* und außerhalb von D_p.

Im Falle $a_1 = a_2 = \cdots = a_p > a_{p+1} = a_{p+2} = \cdots = a_n$ hat das Polynom

$$f_p(z) = (-1)^p f(z) = (a_p - z)^p (z - a_{p+1})^{n-p}$$

im Punkt z_0 einen negativen Realteil, weil

$$\operatorname{arc} f_p(z_0) = p \operatorname{arc}(a_p - z_0) + (n-p) \operatorname{arc}(z_0 - a_{p+1}) = -p\beta + (n-p)\gamma = -\omega$$

und $\pi < \omega < \pi + \frac{\pi}{2}$ sind. In den inneren Punkten von D_p ist $\mathrm{Re}\, f_p(z) > 0$.

Der Winkelraum W enthält im Innern offenbar einen Punkt z_0, für den

$$\frac{\pi}{2n} < \operatorname{arc}(z_0 - a_1) < \frac{\pi}{n} \quad \text{bzw.} \quad -\frac{\pi}{2n} > \operatorname{arc}(a_n - z_0) > -\frac{\pi}{n}$$

sind, je nachdem D_0 bzw. D_n ein Teil von W ist.

Im Falle $a_1 = a_2 = \cdots = a_n$ hat das Polynom

$$f_0(z) = (z - a_1)^n \quad \text{bzw.} \quad f_n(z) = (a_n - z)^n$$

im Punkte z_0 einen negativen Realteil, jedoch in den inneren Punkten von D_0 bzw. D_n einen positiven Realteil.

Damit ist die Ergänzung II des Satzes I bewiesen.

2. III. *Die Punkte der komplexen Ebene, in denen der Wert des Polynoms* (1) *reell bzw. rein imaginär ist, liegen auf der reellen Achse und auf der algebraischen Kurve*

$$f'(x) - \frac{y^2}{3!} f'''(x) + \frac{y^4}{5!} f^{(5)}(x) - \cdots = 0 \tag{5}$$

$(n-1)$*-ter, bzw. auf der Kurve*

$$f(x) - \frac{y^2}{2!} f''(x) + \frac{y^4}{4!} f^{(4)}(x) - \cdots = 0 \tag{6}$$

n-ter Ordnung. Diese Kurven haben $n-1$ *bzw.* n *Asymptoten, die durch den Schwerpunkt der Nullstellen von* $f(z)$ *gehen und zur reellen Achse die Winkel*

$k\frac{\pi}{n}$ $(k = 1, 2, \ldots, n-1)$ *bzw.* $k\frac{\pi}{2n}$ $(k = 1, 2, \ldots, n)$ *bilden.*

Die Gleichungen (5) und (6) der Kurven folgen unmittelbar aus der Taylorschen Formel

$$f(z) = f(x+iy) = f(x) + \sum_{k=1}^{n} \frac{(iy)^k}{k!} f^{(k)}(x) = \mathrm{Re}\, f(z) + \mathrm{Im}\, f(z).$$

Man kann annehmen, daß der Schwerpunkt der Nullstellen von $f(z)$ in den Nullpunkt fällt, d. h. $f(z)$ die Form

$$f(z) = z^n + A_2 z^{n-2} + \cdots + A_{n-2} z + A_n = 0 \qquad (A_i \text{ reell}) \tag{7}$$

besitzt. Dann folgt aus

$$f(z)=f\left(r(\cos\varphi+i\sin\varphi)\right)=r^{n}(\cos n\varphi+i\sin n\varphi)+$$
$$+A_2 r^{n-2}\left[\cos(n-2)\varphi+i\sin(n-2)\varphi\right]+\cdots+A_{n-1}r(\cos\varphi+i\sin\varphi)+A_n,$$

daß die Asymptoten der Kurve (5) bzw. (6) durch den Nullpunkt gehen und ihre Winkel der Gleichung $\frac{\sin n\varphi}{\sin\varphi}=0$ bzw. $\cos n\varphi=0$ genügen, w. z. b. w.

Wir bemerken, daß wenn die Wurzeln des Polynoms alle verschieden sind, $f(z)$ in den Punkten jedes endlichen Bogens B der Kurve (5) dasselbe Vorzeichen besitzt. Gibt es nämlich auf B zwei Punkte z_1 und z_2 mit $f(z_1)<0$ und $f(z_2)>0$, so hat $f(z)$ auf B mindestens eine Nullstelle z_0. Der Punkt z_0 liegt auf der Geraden $y=0$. Es gilt also infolge (5) außer $f(z_0)=0$ auch $f'(z_0)=0$, was der Einfachheit der Nullstellen von $f(z)$ widerspricht.

Hat das Polynom $f(z)$ die Form (7), so hat die Kurve (5) im Falle $n=3$ die Gleichung

$$f'(x)-\frac{y^2}{3!}f'''(x)=3x^2+A_2-y^2=0.$$

Sie ist also im Falle $A_2\neq 0$ eine Hyperbel, deren Scheitel in die Nullstellen der Derivierten von $f(z)$ fallen. Auf dem rechten bzw. linken Hyperbelzweig hat das Polynom $f(z)$ dritten Grades positive bzw. negative Werte.

3. Bezüglich der Wertverteilung der Polynome mit lauter reellen Nullstellen auf einer zur reellen Achse parallelen Geraden gilt der Satz

IV. *Bezeichnet $f(z)$ ein beliebiges Polynom n-ten Grades mit lauter reellen Nullstellen, ist $z_0=x_0+iy_0$ $(y_0>0)$ und bezeichnet ζ bzw. ζ' einen beliebigen inneren Punkt der Strecke*

$$S_1=\left(z_0,\ z_1=z_0+2y_0\operatorname{tg}\frac{\pi}{n}\right)\ \textit{oder}\ S_2=\left(z_0,\ z_2=z_0-2y_0\operatorname{tg}\frac{\pi}{n}\right)$$

bzw.

$$S_3=\left(z_0,\ z_3=z_0+2y_0\operatorname{tg}\frac{\pi}{2n}\right)\ \textit{oder}\ S_4=\left(z_0,\ z_4=z_0-2y_0\operatorname{tg}\frac{\pi}{2n}\right),$$

so ist das Verhältnis

$$V=f(\zeta):f(z_0)\quad \textit{bzw.}\quad V'=f(\zeta'):f(z_0)$$

keine positive bzw. keine reelle Zahl.

Während ein Punkt z die Strecke S_1 bzw. S_2 von z_0 ab beschreibt, nimmt die Funktion

$$\varphi_k(z)=\operatorname{arc}(z-a_k)\qquad (k=1,2,\ldots,n), \tag{8}$$

wo a_k die Nullstellen des Polynoms $f(z)$ bedeuten, offenbar stetig und monoton ab bzw. zu. Erscheint die Teilstrecke (z_0,ζ) von S_1 bzw. S_2 vom Punkt a_k aus unter einem Winkel ω_k' bzw. ω_k'', so sind

$$\varphi_k(z_0)-\varphi_k(\zeta)=\omega_k' \quad \text{bzw.} \quad \varphi_k(z_0)-\varphi_k(\zeta)=-\omega_k''$$

und

$$\operatorname{arc}\frac{f(z_0)}{f(\zeta)}=\sum_{k=1}^{n}[\varphi_k(z_0)-\varphi_k(\zeta)]=\sum_{k=1}^{n}\omega_k' \quad \text{bzw.} \quad \operatorname{arc}\frac{f(z_0)}{f(\zeta)}=-\sum_{k=1}^{n}\omega_k''.$$

Die Strecke S_1 bzw. S_2 erscheint vom Punkt $x_0+y_0 \operatorname{tg}\frac{\pi}{n}$ bzw. $x_0-y_0 \operatorname{tg}\frac{\pi}{n}$ der reellen Achse aus unter dem Winkel $2\frac{\pi}{n}$, von jedem anderen Punkt der reellen Achse aus unter einem Winkel $<2\frac{\pi}{n}$. Eine Teilstrecke von S_1 bzw. S_2 erscheint also von einem Punkt a_k aus unter einem Winkel $<\frac{2\pi}{n}$. Hieraus folgt, daß

$$\omega'<\frac{2\pi}{n}, \quad \omega''<\frac{2\pi}{n} \quad \text{und} \quad \operatorname{arc}\frac{f(z_0)}{f(\zeta)}\not\equiv 0 \pmod{2\pi}$$

sind. Ebenso kann man beweisen, daß

$$\operatorname{arc}\frac{f(z_0)}{f(\zeta')}\not\equiv 0 \pmod{\pi}$$

ist. Damit ist der Satz IV bewiesen. Dieser Satz gilt leicht ersichtlich auch dann, wenn y_0 negativ ist.

Der Beweis zeigt, daß $\frac{f(z_0)}{f(z_1)}$ nur dann positiv ist, wenn

$$f(z)=\left(z-x_0-y_0 \operatorname{tg}\frac{\pi}{n}\right)^n$$

ist.

Auf eine ähnliche Weise läßt sich der folgende Satz beweisen:

V. *Hat das Polynom $f(z)$ n-ten Grades lauter reelle Nullstellen a_k und bezeichnet ζ einen beliebigen inneren Punkt der Kreisscheibe K*

$$|z-z_0|=|y_0| \sin\frac{\pi}{n}, \quad z_0=x_0+iy_0 \quad (y_0\neq 0),$$

so ist das Verhältnis $V(\zeta)=f(\zeta):f(z_0)$ keine negative Zahl.

Die Strecke (z_0, ζ) erscheint von einem Punkt x der reellen Achse aus unter einem Winkel $\omega(x)<\frac{\pi}{n}$.

Sind nun

$$\omega(a_k)=\omega_k \quad \text{und} \quad \vartheta_k=\operatorname{arc}\frac{\zeta-a_k}{z_0-a_k}, \quad |\vartheta_k|<\pi \quad (k=1, 2, \ldots, n),$$

so sind $|\vartheta_k|=\omega_k$. Es kann angenommen werden, daß $\vartheta_1=\omega_1$ ist. Liegen dann die Punkte $a_1, a_2, \ldots, a_p$ bzw. $a_{p+1}, a_{p+2}, \ldots, a_q$ an der einen bzw. an der anderen Seite der Geraden g durch das Punktpaar z_0, ζ und fallen die übrigen $n-q$ Nullstellen von $f(z)$ auf g, so sind

$$\vartheta_h = \omega_h, \quad \vartheta_j = -\omega_j, \quad \omega_l = 0$$
$$(h = 1, 2, \ldots, p;\ j = p+1, \ldots, q;\ l = q+1, \ldots, n)$$

und

$$\Theta = \operatorname{arc} V(\zeta) = \sum_{k=1}^{n} \vartheta_k = \sum_{h=1}^{p} \omega_h - \sum_{j=p+1}^{q} \omega_j.$$

Daraus folgt die Ungleichung

$$|\Theta| \leqq \sum_{k=1}^{n} |\vartheta_k| = \sum_{k=1}^{n} \omega_k < n \frac{\pi}{n} = \pi.$$

Damit ist der Satz V bewiesen.

Der Beweis zeigt, daß $V(\zeta)$ für einen Punkt ζ des Randkreises von K dann und nur dann einen negativen Wert besitzt, wenn $f(z) = (z-z_0)^n$ gilt und ζ mit einem der Berührungspunkte der Tangenten von x_0 zu K zusammenfällt.

(Eingegangen am 20. Juni 1952.)

РАСПРЕДЕЛЕНИЕ ЗНАЧЕНИЙ ПОЛИНОМОВ С ДЕЙСТВИТЕЛЬНЫМИ КОЭФФИЦИЕНТАМИ И ДЕЙСТВИТЕЛЬНЫМИ КОРНЯМИ

Д. С.-НАДЬ (Сегед)

(Резюме)

Знак вещественной части полинома

$$f(z) = (z-a_1)(z-a_2)\ldots(z-a_n), \qquad a_1 \geqq a_2 \geqq \ldots \geqq a_n$$

внутри делтоида D_p $(1 \leqq p \leqq n-1)$ с главной диагоналей (a_p, a_{p+1}) и с углами $\frac{\pi}{p}$ и $\frac{\pi}{n-p}$ у вершин a_p и a_{p+1} соответственно, равен знаку $(-1)^p$. Подобное имеет место внутри областей D_0 и D_n, где D_0 и D_n суть „уголные пространства“ угла $\frac{\pi}{n}$ симметрические относительно действительной оси, с вершинами a_1 и a_n соответственно, не содержащие во внутренности корней a_k.

Связные области комплексной плоскости в которых вещественная часть $f(z)$ сохраняет знак, ограничены алгебраической кривой порядка n. В предельных точках вещественная часть $f(z)$ обращается в нуль.

Точки в которых значение $f(z)$ вещественно, размещены вдоль действительной оси и на алгебраической кривой порядка $n-1$. Если корни $f(z)$ различны, то на любой конечной дуги этой кривой $f(z)$ имеет один и тот же знак.

ON THE ZEROS OF POLYNOMIALS

By

A. RÉNYI (Budapest) and P. TURÁN (Budapest), corresponding members of the Academy

1. This paper deals with the method of D. BERNOULLI,[1] N. I. LOBATSCHEWSKY[2] and N. GRAEFFE[3] devised for the approximative solution of algebraic equations. In the usual form[4] the method asserts that if

$$f_0(x) = a_{00} + a_{10}x + \ldots + a_{n0}x^n = 0 \qquad (a_{n0} = 1) \tag{1.1}$$

is the equation to be solved, with the zeros $z_1, z_2, \ldots, z_n$ where

$$|z_n| < |z_{n-1}| < \ldots < |z_1|, \tag{1.2}$$

then we have to form the so-called Graeffe-transforms $f_\nu(x)$ defined by

$$f_\nu(x) = (-1)^n f_{\nu-1}(\sqrt{x}) f_{\nu-1}(-\sqrt{x}) \qquad (\nu = 1, 2, \ldots). \tag{1.3}$$

If

$$f_\nu(x) = a_{0\nu} + a_{1\nu}x + \ldots + a_{n\nu}x^n \qquad (a_{n\nu} = 1), \tag{1.4}$$

then the method asserts that

$$\begin{aligned} |z_1| &= \lim_{\nu\to\infty} \left|\frac{a_{n-1,\nu}}{a_{n,\nu}}\right|^{\frac{1}{2^\nu}} \\ |z_2| &= \lim_{\nu\to\infty} \left|\frac{a_{n-2,\nu}}{a_{n-1,\nu}}\right|^{\frac{1}{2^\nu}} \\ &\cdots\cdots\cdots \\ |z_n| &= \lim_{\nu\to\infty} \left|\frac{a_{0\nu}}{a_{1\nu}}\right|^{\frac{1}{2^\nu}}. \end{aligned} \tag{1.5}$$

Curiously enough the method was used until 1930 without hesitation for small ν-values and without estimation of the error, nothing said about the condition (1. 2); without this restriction the rule is false in general. Afterwards, in the papers of R. SAN JUAN,[5] A. OSTROWSKI[6] and the second-named

[1] D. BERNOULLI, *Commentationes Petropolitanae,* **3** (1728).

[2] Н. И. Лобачевский, Алгебра или вычисление конечных (Казань, 1834).

[3] N. GRAEFFE, Die Auflösung der höheren numerischen Gleichungen (Zürich, 1837).

[4] See e. g. Я. С. Безикович, Приближенные вычисления.

[5] R. SAN JUAN, Compléments à la méthode de Gräffe pour la résolution des équations algébriques, *Bull. des Sciences Math.*, **59** (1935), pp. 104—109.

[6] A. OSTROWSKI, Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent, *Acta Math.*, **72** (1940), pp. 99—257.

author[7] are contained the first modified forms of this method which work with *finite* ν's, give estimations for the errors and are valid without the restriction (1.2). Replacing (1.2) by[8]

$$|z_n| \leqq |z_{n-1}| \leqq \ldots \leqq |z_2| \leqq |z_1| \tag{1.6}$$

and restricting ourselves to the approximation of $|z_1|$, the rule of OSTROWSKI and the first rule in [7] both form the first ν Graeffe-transforms and obtain approximative values T_ν resp. T'_ν for $|z_1|$ in terms of $a_{j,\nu}$ so that

$$\left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{T_\nu} \leqq 2^{2^{-\nu}} \tag{1.7}$$

and

$$\left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{T'_\nu} \leqq 2^{2^{-\nu}}. \tag{1.8}$$

The remarkable fact in both estimations (1.7) and (1.8) is that they depend only upon n and ν, i. e. *not* upon the coefficients of $f_0(x)$ and give exactly the same bounds.

2. In [7] the author has expressed his opinion that his procedure can be imbedded in a chain of procedures (some of which give narrower bounds) *working with the* ν^{th} *Graeffe-transforms.* In this paper we shall show that this opinion was right. Indeed, we shall show the correctness of the following

Rule I. Let us form with the coefficients (1.4) of the ν^{th} Graeffe-transform $f_\nu(x)$ of $f_0(x)$, with $M = [n^2 \log(2n^2)] + 2$ the sequence $s_1, s_2, \ldots, s_M$ successively from the system of recurrent equations

$$\begin{aligned} s_1 + a_{n-1,\nu} &= 0 \\ s_2 + a_{n-1,\nu}s_1 + 2a_{n-2,\nu} &= 0 \\ &\;\;\vdots \\ s_n + a_{n-1,\nu}s_{n-1} + \ldots + na_{0,\nu} &= 0 \\ s_{n+1} + a_{n-1,\nu}s_n + \ldots + a_{0\nu}s_1 &= 0 \\ &\;\;\vdots \\ s_M + a_{n-1,\nu}s_{M-1} + \ldots + a_{0\nu}s_{M-n} &= 0. \end{aligned} \tag{2.1}$$

Then we have

$$\left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{\left(\max\limits_{j=1,2,\ldots,M} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}}} \leqq \left(\frac{1}{1-\frac{1}{n}}\right)^{2^{-\nu}}.$$

[7] P. TURÁN, On approximative solution of algebraic equations, *Publ. Math. Debrecen,* 2 (1951), pp. 26—42.

[8] Obviously (1.6) is a notation only and contains no restrictions, in contrary to (1.2).

With the notation

$$(2.2)\qquad \left(\max_{j=1,2,\ldots,M} |s_j|^{1/j}\right)^{2^{-\nu}} = T_\nu''$$

we have three approximative values for $|z_1|$, namely T_ν, T_ν' and T_ν''. Owing to the rule I, T_ν'' gives the closest approximation, but needs obviously the greatest computational work. Rule I in itself gives no scale of rules, such a scale is furnished by the more general

Rule II. With an arbitrary ε in $0<\varepsilon<1$ and

$$(2.3)\qquad N = \left[\frac{n}{\varepsilon}\log\frac{2n}{\varepsilon}\right]+2$$

form the system analogous to (2. 1) but ending with the equation

$$s_N + a_{n-1,\nu}s_{N-1} + \ldots + a_{0,\nu}s_{N-\nu} = 0.$$

Then we can again successively determine $s_1, s_2, \ldots, s_N$ and we have

$$(2.4)\qquad \left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{\left(\max\limits_{j=1,2,\ldots,N} |s_j|^{1/j}\right)^{2^{-\nu}}} \leqq \left(\frac{1}{1-\varepsilon}\right)^{2^{-\nu}}.$$

Rule I follows from rule II taking $\varepsilon = \frac{1}{n}$; hence it suffices to prove rule II only. For $\varepsilon = \frac{1}{2}$ we have

$$N = [2n \log 4n] + 2$$

and the corresponding bounds in (2. 4) became $\left(\frac{1}{n}\right)^{2^{-\nu}}$, $2^{2^{-\nu}}$. These bounds are identical with those of rule I in [7]. We had there however to form

$$T_\nu' = \left(\max_{j=1,2,\ldots,2n} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}};$$

this means that by greater computational work we obtained an approximative value, which is not better. The computational work with the approximating value T_ν' is also considerable; we emphasise again the importance of the conjecture expressed in [7] that for a suitable $c>1$, *independent of* n, we have already

$$(2.5)\qquad \left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{\left(\max\limits_{j=1,2,\ldots,n} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}}} \leqq c^{2^{-\nu}}.$$

A comparison of these with the rule II shows that probably in (2. 3) one can replace N by an N' of the form

$$N' = [c_1(\varepsilon)n];$$

the truth of this conjecture would diminish considerably the necessary calculations. Further remarks on this subject can be found in **7**.

As the example

$$z_1 = 1,\ z_2 = z_3 = \ldots = z_n = 0$$

shows, we can have for an arbitrary large integer N''

$$\frac{|z_1|}{\left(\max\limits_{j=1,2,\ldots,N''} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}}} = 1,$$

i. e. one cannot replace in (2.4) the quantity $(1-\varepsilon)^{2^{-\nu}}$ by another one, less than 1.

3. The paper [7] was one of a series of papers dealing with the various applications of a central idea.[9] This idea was to estimate from below sums of the form

$$\left|\sum_{j=1}^{n} b_j w_j^t\right| \tag{3.1}$$

where the numbers b_j and w_j are arbitrary complex numbers, by $\max\limits_{j=1,2,\ldots n} |w_j|^t$ resp. by $\min\limits_{j=1,\ldots,n} |w_j|^t$ when t is an appropriate integer from an interval $(m+1, m+n)$ (m a non-negative integer). The deepest result of this theory is the inequality

$$\max_{\substack{m+1 \leqq t \leqq m+n \\ t \text{ integer}}} |b_1 w_1^t + \ldots + b_n w_n^t| \geqq \left(\frac{n}{250(m+2n)}\right)^n \min_j |b_1 + \ldots + b_j| \tag{3.2}$$

if

$$1 = |w_1| \geqq |w_2| \geqq \ldots \geqq |w_n|. \tag{3.3}$$

The result so obtained for $m=0$ is rather weak, even in the case

$$b_1 = b_2 = \ldots = b_n = 1; \tag{3.4}$$

the main tool of the proofs for the rules in [7] was a direct approach for this case. So we obtained in [7] the inequalities

$$\max_{\substack{1 \leqq t \leqq n \\ t \text{ integer}}} |w_1^t + w_2^t + \ldots + w_n^t| \geqq \frac{\log 2}{\frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n}}, \tag{3.5}$$

$$\max_{\substack{1 \leqq t \leqq 2n \\ t \text{ integer}}} |w_1^t + \ldots + w_n^t| \geqq \frac{1}{2}. \tag{3.6}$$

The conjecture (2.5) would follow from the proof of

$$\max_{\substack{1 \leqq t \leqq n \\ t \text{ integer}}} |w_1^t + \ldots + w_n^t| \geqq c. \tag{3.7}$$

[9] For a detailed exposition see the forthcoming book of the second-named author entitled *Über eine neue Methode der Analysis und ihre Anwendungen.*

The proof for the rule II of the present paper will be based on the inequality

$$\max_{\substack{1 \leqq t \leqq N \\ t \text{ integer}}} |w_1^t + \ldots + w_n^t| \geqq 1 - \varepsilon \tag{3.8}$$

where N is defined in (2.3) and (3.3) holds. I. e. the lower estimation in (3.8) is better than in (3.5) and (3.6), but the range of t is bigger. But more generally for the general expression $\sum_{j=1}^{n} b_j w_j^t$ one can obtain for $m = 0$ an estimation, in a certain respect better than (3.2), in an important special case which may be called for a certain reason Dirichletian case. This will be given by

THEOREM I. *Putting*

$$f(t) = \sum_{j=1}^{n} b_j w_j^t,$$

we suppose that the w_j's satisfy (3.3) *and the b_j's are positive. Choosing ε satisfying the restriction*

$$0 < \varepsilon < \min\left(1, \frac{2(b_2 + b_3 + \ldots + b_n)}{b_1}\right), \tag{3.9}$$

we define

$$M_1 = \left[\frac{f(0)}{b_1 \varepsilon} \log \frac{2f(0)}{b_1 \varepsilon}\right] + 2. \tag{3.10}$$

Then we have

$$\max_{\substack{1 \leqq t \leqq M_1 \\ t \text{ integer}}} |f(t)| \geqq b_1(1 - \varepsilon).$$

For $b_1 = b_2 = \ldots = b_n = 1$ we have obviously got again (3.8).

Next we deduce from (3.8) the rule II, in the following § we prove theorem I and in the further §§ we treat the similar problem for integrals.

4. Hence we turn to the proof of rule II. The quantity s_j is obviously the j^{th} power-sum of the zeros of $f_\nu(x)$, i. e.

$$s_j = \sum_{l=1}^{n} z_l^{j 2^\nu}.$$

Hence, owing to (1.6), for *all* natural j's, we have

$$|s_j| \leqq n |z_1|^{j 2^\nu},$$

i. e. this will hold choosing $j = j_0$, where

$$|s_{j_0}|^{\frac{1}{j_0}} = \max_{j = 1, 2, \ldots, N} |s_j|^{\frac{1}{j}}.$$

Thus

$$|s_{j_0}|^{\frac{1}{j_0}} \leqq n^{\frac{1}{j_0}} |z_1|^{2^\nu} \leqq n |z_1|^{2^\nu},$$

$$\frac{|z_1|}{\left(|s_{j_0}|^{\frac{1}{j_0}}\right)^{2^{-\nu}}} \geqq n^{-2^{-\nu}},$$

which gives already the lower limitation of the rule II. To prove the upper limitation, we apply (3.8) with

$$w_j = \left(\frac{z_j}{z_1}\right)^{2^\nu} \qquad (j = 1, 2, \ldots, n).$$

Then there is an integer t_0 with $1 \leqq t_0 \leqq N$ such that

$$\frac{|s_{t_0}|}{|z_1|^{2^\nu t_0}} \geqq 1 - \varepsilon,$$

$$|z_1|^{2^\nu} \leqq \left(\frac{1}{1-\varepsilon}\right)^{\frac{1}{t_0}} |s_{t_0}|^{\frac{1}{t_0}} \leqq \frac{1}{1-\varepsilon} \max_{j=1,2,\ldots,N} |s_j|^{\frac{1}{j}},$$

i. e.

$$\frac{|z_1|}{\left(\max\limits_{1 \leqq j \leqq N} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}}} \leqq \left(\frac{1}{1-\varepsilon}\right)^{2^{-\nu}}$$

which completes the proof of the rule II.

5. Now we turn to the proof of theorem I. Without loss of generality we may suppose

$$w_1 = 1. \tag{5.1}$$

We consider for $|z| > 1$ the function

$$g(z) = \sum_{j=1}^{n} \frac{b_j}{z - w_j}. \tag{5.2}$$

Since $|w_j| \leqq 1, g(z)$ is here regular, i. e. we have from (5.2) for $|z| > 1$

$$g(z) = \sum_{\nu=0}^{\infty} \frac{f(\nu)}{z^{\nu+1}}. \tag{5.3}$$

If $R > 1$ (we shall fix the value of R only later), introducing the notation

$$\max_{\substack{1 \leqq \nu \leqq M_1 \\ \nu \text{ integer}}} |f\,\nu)| = U,$$

we obtain from (5.3), owing to the positivity of the b_j's,

$$|g(R)| \leqq \frac{f(0)}{R} + U\left(\frac{1}{R^2} + \frac{1}{R^3} + \ldots + \frac{1}{R^{M_1}}\right) + f(0)\left(\frac{1}{R^{M_1+1}} + \ldots\right) <$$

$$< \frac{f(0)}{R} + \frac{U}{R(R-1)} + \frac{f(0)}{R^{M_1}(R-1)}. \tag{5.4}$$

On the other hand we have, using also (5. 1) and (3. 3),[10]

$$|g(R)| \geqq \Re g(R) = \frac{b_1}{R-1} + \sum_{j=2}^{n} b_j \Re \frac{1}{R-w_j}. \tag{5.5}$$

Now we observe that

$$\Re \frac{1}{R-w_j} \geqq \frac{1}{R+1};$$

hence from (5. 5) and (5. 4) we get

$$\frac{b_1}{R-1} + \frac{b_2+\ldots+b_n}{R+1} \leqq \frac{b_1+\ldots+b_n}{R} + \frac{b_1+\ldots+b_n}{R^{M_1}(R-1)} + \frac{U}{R(R-1)}. \tag{5.6}$$

Choosing

$$R = 1 + \frac{b_1\varepsilon}{(b_1+\ldots+b_n) - b_1\left(1+\frac{\varepsilon}{2}\right)}, \tag{5.7}$$

the condition $R>1$ is fulfilled owing to (3. 9). Since

$$\frac{b_1}{R-1} + \frac{b_2+\ldots+b_n}{R+1} - \frac{b_1+\ldots+b_n}{R} = \frac{b_1\left(1-\frac{\varepsilon}{2}\right)}{R(R-1)},$$

(5. 6) gives

$$U \geqq b_1\left(1-\frac{\varepsilon}{2}\right) - (b_1+\ldots+b_n)R^{1-M_1}. \tag{5.8}$$

Now from (5. 7) and (3. 9)

$$\log R \geqq \frac{R-1}{R} = \frac{b_1\varepsilon}{(b_2+\ldots+b_n)+b_1\frac{\varepsilon}{2}} > \frac{b_1\varepsilon}{b_1+\ldots+b_n}$$

and using the definition of M_1 in (3. 10)

$$R^{M_1-1} \geqq \exp\left\{\left(1+\left[\frac{f(0)}{b_1\varepsilon}\log\frac{2f(0)}{b_1\varepsilon}\right]\right)\frac{b_1\varepsilon}{b_1+b_2+\ldots+b_n}\right\} \geqq$$

$$\geqq \exp\left\{\frac{f(0)}{b_1\varepsilon}\log\frac{2f(0)}{b_1\varepsilon}\cdot\frac{b_1\varepsilon}{f(0)}\right\} = \frac{2f(0)}{b_1\varepsilon};$$

putting this in (5. 8) we obtain

$$U \geqq b_1\left(1-\frac{\varepsilon}{2}\right) - \frac{b_1\varepsilon}{2} = b_1(1-\varepsilon),$$

indeed. Q. e. d.

6. The expression

$$f(t) = \sum_{j=1}^{n} b_j w_j^t$$

— in the important case, when the w_j's are of the form $w_j = e^{i\alpha_j}$ with real

[10] $\Re z$ denotes the real part of the complex number z.

α_j's and the b_j's are positive — can be brought to the form of a Fourier—Stieltjes integral $\int_{-\infty}^{\infty} e^{itx} dF(x)$ with a non-negative non-decreasing $F(x)$; if $\alpha_1 > \alpha_2 > \ldots > \alpha_n$, take simply

$$\begin{aligned}
F(x) &= 0 && \text{for } -\infty < x < \alpha_n, \\
F(x) &= b_n && \text{for } \alpha_n \leqq x < \alpha_{n-1}, \\
F(x) &= b_n + b_{n-1} && \text{for } \alpha_{n-1} \leqq x < \alpha_{n-2}, \\
&\ldots\ldots\ldots\ldots \\
F(x) &= b_n + \ldots + b_2 && \text{for } \alpha_2 \leqq x < \alpha_1, \\
F(x) &= b_n + \ldots + b_1 && \text{for } \alpha_1 \leqq x < \infty.
\end{aligned}$$

The extension of the whole theory, mentioned in **3,** to Fourier—Stieltjes integrals seems to be very desirable in particular with regard to possible applications to the study of characteristic functions in probability theory. As a first result in this trend we prove

THEOREM II. *If $F(x)$ is positive and non-decreasing on the real axis such that*

$$F(+0) - F(-0) = \Delta > 0 \tag{6.1}$$

and

$$\int_{-\infty}^{\infty} dF(x) = 1, \tag{6.2}$$

then, if

$$0 < \varepsilon < \min\left(1, \frac{2(1-\Delta)}{\Delta}\right) \tag{6.3}$$

we have

$$\max_{\substack{1 \leqq \nu \leqq 2 + \left[\frac{1}{\Delta\varepsilon} \log \frac{2}{\Delta\varepsilon}\right] \\ \nu \text{ integer}}} \left| \int_{-\infty}^{\infty} e^{i\nu x} dF(x) \right| \geqq \Delta(1-\varepsilon). \tag{6.4}$$

Since the proof is very similar to that of theorem I, it will suffice only to sketch it. We denote the expression on the left of (6.4) by V, the quantity

$$2 + \left[\frac{1}{\Delta\varepsilon} \log \frac{2}{\Delta\varepsilon}\right] \tag{6.5}$$

by L and choose

$$R = 1 + \frac{\Delta\varepsilon}{1 - \left(1 + \frac{\varepsilon}{2}\right)\Delta} (>1). \tag{6.6}$$

Then we have on one hand

$$\int_{-\infty}^{\infty} \frac{dF(x)}{R - e^{ix}} = \frac{1}{R} + \sum_{\nu=1}^{\infty} \frac{1}{R^{\nu+1}} \int_{-\infty}^{\infty} e^{i\nu x} dF(x)$$

and on the other hand

$$\int_{-\infty}^{\infty} \frac{dF(x)}{R-e^{ix}} \geqq \frac{\Delta}{R-1} + \frac{1-\Delta}{R+1}.$$

Hence as in the previous proof

$$\frac{\Delta}{R-1} + \frac{1-\Delta}{R+1} \leqq \frac{1}{R} + \frac{V}{R(R-1)} + \frac{1}{R^L(R-1)}.$$

Replacing R by its value in (6.6) we obtain

$$\frac{\Delta}{R-1} + \frac{1-\Delta}{R+1} - \frac{1}{R} = \frac{\Delta\left(1-\frac{\varepsilon}{2}\right)}{R(R-1)},$$

i. e.

(6.7) $$V \geqq \Delta\left(1-\frac{\varepsilon}{2}\right) - R^{1-L}.$$

Since

$$\log R \geqq \frac{R-1}{R} = \frac{\Delta\varepsilon}{1-\left(1-\frac{\varepsilon}{2}\right)\Delta} \geqq \Delta\varepsilon$$

$$R^{L-1} \geqq \exp\left\{\left(1+\left[\frac{1}{\Delta\varepsilon}\log\frac{2}{\Delta\varepsilon}\right]\right)\Delta\varepsilon\right\} \geqq \frac{2}{\Delta\varepsilon},$$

we obtain indeed from (6.7)

$$V \geqq \Delta(1-\varepsilon).$$

7. In all the rules mentioned above, the upper limitation is generally the better one. It would be at the same time of theoretical and practical interest to modify the procedure so as to improve the *lower* limitation at a fixed ν-value.

(*Received 5 January 1953*)

О КОРНЯХ МНОГОЧЛЕНОВ

А. РЕНЬИ и П. ТУРАН (Будапешт)

(Резюме)

В работе доказывается следующая теорема 1: Пусть $w_1, w_2, \ldots, w_n$ комплексные числа, $|w_1| \geqq |w_2| \geqq \ldots \geqq |w_n|$ и $b_1, b_2, \ldots, b_n$ положительные числа; пусть

$$f(t) = \sum_{j=1}^{n} b_j w_j^t \qquad (t = 0, 1, 2, \ldots);$$

тогда

$$\max_{t=1,2,\ldots,M_1} |f(t)| \geqq b_1(1-\varepsilon),$$

где

$$M_1 = \left[\frac{f(0)}{b_1 \varepsilon} \log \frac{2f(0)}{b_1 \varepsilon}\right] + 2 \quad \text{и} \quad 0 < \varepsilon < \min\left(1, \frac{2(f(0)-b_1)}{b_1}\right).$$

Как следствие этой теоремы получается следующий результат: пусть

$$f_0(x) = \sum_{k=0}^{n} a_{k0} x^k, \quad f_\nu(x) = (-1)^n f_{\nu-1}(\sqrt{x}) f_{\nu-1}(-\sqrt{x}) \qquad (\nu = 1, 2, \ldots),$$

$$f_\nu(x) = \sum_{k=0}^{n} a_{k\nu} x^k \qquad (a_{n\nu} = 1)$$

и определим числа s_k $(k = 1, 2, \ldots, N)$ из систем линейных уравнений

$$\sum_{j=0}^{k-1} s_{k-j} a_{n-j,\nu} + k a_{n-k,\nu} = 0 \qquad (k = 1, 2, \ldots, n)$$

и

$$\sum_{j=0}^{n} s_{k-j} a_{n-j,\nu} = 0 \qquad (k = n+1, \ldots, N),$$

тогда если z_1 обозначает найбольшой по модулю корень уравнения $f_0(x) = 0$, то имеем

$$\left(\frac{1}{n}\right)^{2^{-\nu}} \leqq \frac{|z_1|}{\left(\max\limits_{j=1,2,\ldots,N} |s_j|^{\frac{1}{j}}\right)^{2^{-\nu}}} \leqq \left(\frac{1}{1-\varepsilon}\right)^{2^{-\nu}},$$

если $0 < \varepsilon < 1$ и $N = \left[\frac{n}{\varepsilon} \log \frac{2n}{\varepsilon}\right] + 2$. Таким образом получается точный вариант метода Бернулли—Данделин—Лобачевского—Грэфе.

Если вместо суммы $\sum\limits_{j=1}^{n} b_j w_j^t$ посмотрим характеристическую функцию $f(t) = \int\limits_{-\infty}^{+\infty} e^{itx} dF(x)$ некоторой функции распределения $F(x)$, для которой $F(+0) - F(-0) = \Delta > 0$, тогда для $0 < \varepsilon < \min\left(1, \frac{2(1-\Delta)}{\Delta}\right)$ и $M = \left[\frac{1}{\Delta \varepsilon} \log \frac{2}{\Delta \varepsilon}\right] + 2$ имеем теорему 2:

$$\max_{t=1,2,\ldots,M} |f(t)| \geqq \Delta(1-\varepsilon).$$

A MOMENT PROBLEM FOR SELF-ADJOINT OPERATORS

By

BÉLA SZ.-NAGY (Szeged), corresponding member of the Academy

1. Introduction

In a recent paper[1] R. V. KADISON proved a theorem called by him "The Generalized Schwarz Inequality", and has applied it as a crucial tool in establishing some algebraic invariants for operator algebras. This theorem, in its simpler but equivalent form, reads as follows:

Let X be a compact subset of the reals, and let $C(X)$ denote the space of all continuous (real) functions $f(x)$ on X. Suppose there is given a map

$$f(x) \rightarrow B_f$$

of $C(X)$ into the set of all bounded self-adjoint operators on some Hilbert space $\mathfrak{H}$, which is linear, order preserving and of norm $\leqq 1$, i. e. such that

a) $B_{c_1 f_1 + c_2 f_2} = c_1 B_{f_1} + c_2 B_{f_2}$,

b) *if $f_1(x) \geqq f_2(x)$ on X, then* $B_{f_1} \geqq B_{f_2}$,[2]

c) $||B_f|| \leqq ||f||$.[3]

Then the operators corresponding to the functions x and x^2 satisfy the inequality

$$B_x^2 \leqq B_{x^2}. \tag{1}$$

Since this assertion refers only to the functions x and x^2, it is more natural to suppose that the map is defined originally only for the functions x^n and their linear combinations with real coefficients, i. e. for the real polynomials in x. Since these polynomials $p(x)$ are dense in $C(X)$, and since $||p_n - p_m|| < \varepsilon$ implies, by a) and c), that $||B_{p_n} - B_{p_m}|| < \varepsilon$, it is easy tho show

[1] R. V. KADISON, A generalized Schwarz inequality and algebraic invariants for operator algebras, *Annals of Math.*, **56** (1952), pp. 494—503.

[2] For two bounded self-adjoint operators A and B, $A \geqq B$ means that $(Au, u) \geqq (Bu, u)$ for all $u \in \mathfrak{H}$.

[3] For a self-adjoint operator B:

$$||B|| = \underset{||u||=1}{\text{l. u. b.}} ||Bu|| = \underset{||u||=||v||=1}{\text{l. u. b.}} |(Bu, v)| = \underset{||u||=1}{\text{l. u. b.}} |(Bu, u)|;$$

for an $f(x) \in C(X)$:

$$||f|| = \max_{x \in X} |f(x)|,$$

this maximum being finite and attained since X is compact and $f(x)$ is continuous.

that such a map may be extended by continuity to a map of the whole space $C(X)$, and that the conditions a)—c) are hereby conserved.

Observe also that it suffices to suppose c) only for the particular function $f(x)\equiv 1$, i. e. that

c′) $\|B_1\| \leqq 1$.

Indeed, for an arbitrary $f(x)$, the obvious inequality $\pm f(x) \leqq \|f\|$ implies, by conditions a) and b),

$$\pm B_f \leqq \|f\| B_1,$$

and this implies, by c′),

$$\|B_f\| \leqq \|f\| \|B_1\| \leqq \|f\|.$$

Thus, writing A_k instead of B_{x^k}, we can formulate KADISON's theorem in the following more general form:

THEOREM 1. *Let X be a compact subset of the reals, and let $A_0, A_1, A_2, \ldots$ be a sequence of bounded self-adjoint operators on the Hilbert space $\mathfrak{H}$, satisfying the following conditions:*

$$(\alpha_X) \begin{cases} \textit{if } c_0+c_1x+c_2x^2+\ldots+c_nx^n \geqq 0 \textit{ on the set } X, \\ \textit{then } c_0A_0+c_1A_1+c_2A_2+\ldots+c_nA_n \geqq O, \end{cases}$$

(β) $\|A_0\| \leqq 1$.

Then we have the inequality

$$A_1^2 \leqq A_2. \tag{2}$$

In the present Note we intend to prove this theorem by a method, entirely different of KADISON's, which is powerful enough not only to yield the inequality (2), but also to characterize in a constructive manner all the sequences $\{A_k\}$ satisfying the above conditions, or, more generally, the condition (α_X). As a matter of fact, condition (β) serves only for a normalization; if a sequence $\{A_k\}$ satisfies only (α_X) and if $A_0 \neq O$, then the sequence $\left\{\frac{1}{\|A_0\|} A_k\right\}$ satisfies both conditions. Thus, if we drop condition (β), we get instead of (2) the inequality

$$A_1^2 \leqq \|A_0\| A_2, \tag{3}$$

at least if $A_0 \neq O$. Since $\pm x \leqq \|x\|$, we have, by (α_X), $\pm A_1 \leqq \|x\| A_0$, thus $A_0 = O$ implies $A_1 = O$, i. e. (3) holds also if $A_0 = O$.

2. Proof of the inequality

First we observe that (β) implies

$$A_0 \leqq I. \tag{4}$$

Now choose an arbitrary $u \in \mathfrak{H}$ and consider the moment problem

$$\int_{-\infty}^{\infty} x^k \, dm(x) = \mu_k \quad \text{with} \quad \mu_k = (A_k u, u) \qquad (k=0,1,2,\ldots).$$

Since, by (α_X),

$$c_0\mu_0 + c_1\mu_1 + \ldots + c_n\mu_n = ((c_0 A_0 + c_1 A_1 + \ldots + c_n A_n)u, u) \geqq 0$$

whenever

$$c_0 + c_1 x + \ldots + c_n x^n \geqq 0 \quad \text{on} \quad X,$$

our moment problem has a solution $m(x)$, and this is a non-decreasing function whose points of increase lie in the set X, i. e. it is constant on every open interval contiguous to X. If $[a, b]$ is the least segment containing the compact set X, then $m(x)$ is constant in particular on the half-lines $x < a$ and $x > b$. Moreover, this solution is fully determined if we require that $m(x) = 0$ for $x < a$ and that $m(x)$ be continuous from the right. Let us denote the thus normalized solution, showing its dependence on the element $u \in \mathfrak{H}$, by $m(u; x)$.

To any couple u, v of elements of $\mathfrak{H}$ associate now the following function of x:

$$m(u, v; x) = \frac{1}{4}[m(u+v; x) - m(u-v; x) + i\, m(u+iv; x) - i\, m(u-iv; x)].$$

Since there is an analogous relation between the bilinear form $(A_k u, v)$ and the quadratic form $(A_k u, u)$, we obtain the relations

$$\int_{-\infty}^{\infty} x^k \, dm(u, v; x) = (A_k u, v) \qquad (k = 0, 1, \ldots). \tag{5}$$

The functions $m(u, v; x)$ are also normalized in the same sense as the functions $m(u; x)$, thus they are fully determined by the elements u, v; we have in particular $m(u, u; x) = m(u; x)$. Since the right-hand side of (5) is, for each k, a (hermitian) symmetric bilinear form in u, v, and since the solution $m(u, v; x)$ is uniquely determined, it follows that, for each value of x, $m(u, v; x)$ is also a symmetric bilinear form in u, v. Moreover, $m(u; x)$ being non-decreasing, we have, by (4),

$$0 \leqq m(u; x) \leqq m(u; \infty) = \int_{-\infty}^{\infty} dm(u; x) = (A_0 u, u) \leqq (u, u),$$

thus the quadratic form corresponding to the bilinear form $m(u, v; x)$ has, on the unit sphere $\|u\| = 1$, the bounds 0 and 1.

It follows that there exists, for each value of x, a bounded self-adjoint operator $F(x)$ on $\mathfrak{H}$ such that

$$m(u, v; x) = (F(x)u, v) \quad \text{for all } u, v \in \mathfrak{H};$$

and we have

$$F(x) \leqq F(x') \;\text{ for } x < x', \quad F(x+0) = F(x), \quad F(-\infty) = O, \quad F(\infty) = A_0.$$

The relations (5) can then be written, symbolically, in the form

$$A_k = \int_{-\infty}^{\infty} x^k \, dF(x) \qquad (k = 0, 1, \ldots). \tag{6}$$

Put

$$\bar{F}(x)=\begin{cases} F(x) & \text{if } x<0, \\ F(x)+I-A_0 & \text{if } x \geqq 0. \end{cases}$$

As $I-A \geqq O$ by (4), we have

(7) $\bar{F}(x) \leqq \bar{F}(x')$ for $x<x'$, $\bar{F}(x+0)=\bar{F}(x)$, $\bar{F}(-\infty)=O$, $\bar{F}(\infty)=I$,

i. e. the family of operators $\{\bar{F}(x)\}$ enjoys the properties of a resolution of the identity with the exception that it does not consist in general of projections; as a matter of fact, we know only that $\bar{F}(x)$ is, for each value of x, a self-adjoint operator on $\mathfrak{H}$. Since the difference $\bar{F}(x)-F(x)$ is constant on the half-lines $x<0$ and $x \geqq 0$, and since x^k vanishes at the point $x=0$ for $k \geqq 1$, the equations (6) remain valid, for $k \geqq 1$, if we replace $F(x)$ by $\bar{F}(x)$, i. e. we have

$$A_k=\int_{-\infty}^{\infty} x^k\, d\bar{F}(x) \qquad (k=1,2,\ldots). \tag{8}$$

Now we recall a theorem of M. A. NEUMARK[4] asserting that each one parameter family $\{\bar{F}(x)\}$ of self-adjoint operators enjoying the properties (7) may be represented in the following form:

$$\bar{F}(x)=P_+ E_+(x) \tag{9}$$

where $\{E_+(x)\}$ is a resolution of the identity (in the proper sense, i. e. consisting of projections) in a suitably defined Hilbert space $\mathfrak{H}_+$ containing the original space $\mathfrak{H}$ as a subspace, and P_+ is the projection on the subspace $\mathfrak{H}$.[5] Moreover, it may be required that $E_+(x)$, as a function of x, have the same points of increase as $\bar{F}(x)$.

Using a representation of this type (9) in our actual problem, it results from the equations (8) that

$$A_k=P_+\int_{-\infty}^{\infty} x^k\, dE_+(x) \qquad (k=1,2,\ldots), \tag{10}$$

or, putting

$$A_+=\int_{-\infty}^{\infty} x\, dE_+(x),$$

$$A_k=P_+A_+^k \qquad (k=1,2,\ldots). \tag{11}$$

A_+ is a self-adjoint operator on $\mathfrak{H}_+$, whose spectrum $\sigma(A_+)$ consists of the points of increase of the function $\bar{F}(x)$, i. e. of the points of increase of the

[4] М. А. Наймарк, Спектральные функции симметрического оператора, Известия Акад. Наук СССР, сер. мат., **4** (1940), pp. 277—309 (English summary pp. 309—318); Об одном представлении аддитивных операторных функций множеств, Доклады Акад. Наук СССР, **41** (1943), pp. 373—375.

[5] Of course, equation (9) is to be understood in the sense that the operators are equal when applied to the elements of the original space $\mathfrak{H}$.

function $F(x)$ and, in the case $A_0 \neq I$, also of the point $x=0$. Thus

$$\sigma(A_+) \subseteq X \quad \text{if} \quad A_1 = I, \quad \sigma(A_+) \subseteq X + \{0\} \quad \text{if} \quad A_0 \neq I.$$

It should be remarked that these informations on the spectrum of A_+ are of importance only in the problem dealt with in the following section; for our actual task to prove the "Generalized Schwarz Inequality" they are not necessary.

Indeed, inequality (2) results immediately from the representation (11) since for an arbitrary $u \in \mathfrak{H}$ we have

$$(A_2 u, u) = (P_+ A_+^2 u, u) = (A_+^2 u, u) = \|A_+ u\|^2 \geqq$$
$$\geqq \|P_+ A_+ u\|^2 = (A_+ P_+ A_+ u, u) = (P_+ A_+ P_+ A_+ u, u) = (A_1^2 u, u).$$

Moreover, we see that $(A_2 u, u) = (A_1^2 u, u)$ only if $P_+ A_+ u = A_+ u$, i. e. if $A_+ u \in \mathfrak{H}$. Thus we have $A_2 = A_1^2$ only if A_+ maps $\mathfrak{H}$ into itself. According to (11) we have in this case $A_1 = A_+$ (in $\mathfrak{H}$), thus $A_k = A_1^k$ $(k=1, 2, \ldots)$.

So we have proved the inequality (2) and, moreover, the following completion of Theorem 1:

In order that under the conditions of Theorem 1 *we have*

$$A_1^2 = A_2,$$

it is necessary and sufficient that

$$A_k = A_1^k \qquad (k = 1, 2, \ldots).$$

3. Construction of all sequences $\{A_k\}$

If $A_0 = I$, then the representation (11) holds also for $k=0$, and, in this case, we have $\sigma(A_+) \subseteq X$.

Conversely, any sequence of operators A_k on the Hilbert space $\mathfrak{H}$, which may be obtained by means of the formula (11) from an arbitrary self-adjoint operator A_+ with $\sigma(A_+) \subseteq X$, on an arbitrary Hilbert space $\mathfrak{H}_+ \supseteq \mathfrak{H}$, satisfies the conditions of Theorem 1; indeed we have $A_0 = I$ and, for an arbitrary polynomial $p(x) = c_0 + c_1 x + \ldots + c_n x^n$ which is $\geqq 0$ on X thus a fortiori on $\sigma(A_+)$,

$$((c_0 A_0 + \ldots + c_n A_n)u, u) = (P_+ p(A_+)u, u) = (p(A_+)u, u) =$$
$$= \int_X p(x)\, d(E_+(x)u, u) \geqq 0.$$

So we have obtained a kind of "parametric representation" of all sequences $\{A_k\}$ of self-adjoint operators on $\mathfrak{H}$ which satisfy condition (α_X) and for which $A_0 = I$.

The problem of finding all sequences $\{A_k\}$ satisfying (α_X) and beginning with an arbitrary A_0 (we do not suppose even (β)) may be solved as follows. We have, at any case, as an immediate consequence of (α_X),

$$A_0 \geqq O.$$

Suppose first that A_0 is definitely positive, i. e. $A_0 u = 0$ only for $u = 0$, and let us denote by R the positive square-root $A_0^{\frac{1}{2}}$ of A_0. Since

$$\|Ru\|^2 = (R^2 u, u) = (A_0 u, u) = 0$$

only for $u = 0$, it follows that R^{-1} exists as an (in general unbounded) self-adjoint operator with domain $\mathfrak{D}$ dense in $\mathfrak{H}$.

Let $M_k = \|x^k\| = \max_{x \in X} |x^k|$; since $\pm x^k \leqq M_k$ on X, (α_X) implies $\pm A_k \leqq M_k A_0$, i. e.

$$(12) \qquad |(A_k u, u)| \leqq M_k (A_0 u, u)$$

for all $u \in \mathfrak{H}$. Hence

$$(13) \qquad |(A_k R^{-1} v, R^{-1} v)| \leqq M_k (A_0 R^{-1} v, R^{-1} v) = M_k (R^2 R^{-1} v, R^{-1} v) = M_k (v, v)$$

for all $v \in \mathfrak{D}$. Consider the symmetric bilinear form

$$(v|w)_k = (A_k R^{-1} v, R^{-1} w)$$

defined on the linear manifold $\mathfrak{D}$. It is well known that $|(v|w)_k|$ and $|(v|v)_k|$ have, under the conditions $\|v\| = 1$, $\|w\| = 1$, the same lowest upper bounds. Since $|(v|v)_k| \leqq M_k$ for $\|v\| = 1$ by (13), we have also $|(v|w)_k| \leqq M_k$ for $\|v\| = 1$, $\|w\| = 1$; thus

$$|(v|w)_k| \leqq M_k \|v\| \, \|w\| \qquad \text{for all } v, w \in \mathfrak{D}.$$

For a fixed v, $(v|w)_k$ is therefore a bounded conjugate-linear form in w on $\mathfrak{D}$. Since R^{-1} is self-adjoint, it follows that $A_k R^{-1} v$ is in the domain of R^{-1}; moreover we have, by (13),

$$|(R^{-1} A_k R^{-1} v, v)| \leqq M_k (v, v).$$

Thus the symmetric operators

$$B_k = R^{-1} A_k R^{-1} \qquad (k = 0, 1, \ldots)$$

have all the domain $\mathfrak{D}$ and are bounded on $\mathfrak{D}$. Then they may be extended to the whole space $\mathfrak{H}$ by continuity; B_0 has, in particular, the extension I. The (extended) operators B_k satisfy the condition (α_X), for, if $c_0 + c_1 x + \ldots + c_n x^n \geqq 0$ on X, then

$$((c_0 B_0 + c_1 B_1 + \ldots + c_n B_n) v, v) = ((c_0 A_0 + c_1 A_1 + \ldots + c_n A_n) R^{-1} v, R^{-1} v) \geqq 0$$

for $v \in \mathfrak{D}$, which implies $c_0 I + c_1 B_1 + \ldots + c_n B_n \geqq O$ on $\mathfrak{H}$ by continuity. Applying our above results to the sequence $\{B_k\}$, we see that there is a representation

$$B_k = P_+ A_+^k \qquad (k = 0, 1, \ldots)$$

with a self-adjoint A_+ on some Hilbert space $\mathfrak{H}_+ \supseteqq \mathfrak{H}$, and with $\sigma(A_+) \subseteqq X$.

Since $Ru \in \mathfrak{D}$ for an arbitrary $u \in \mathfrak{H}$, we conclude that

$$R^{-1} A_k u = R^{-1} A_k R^{-1} R u = B_k R u = P_+ A_+^k R u,$$

i. e.

$$(14) \qquad A_k = A_0^{\frac{1}{2}} P_+ A_+^k A_0^{\frac{1}{2}} \qquad (k = 0, 1, \ldots).$$

It is obvious that, conversely, all sequences $\{A_k\}$ which may be obtained in this way from the given A_0 and from an arbitrarily chosen A_+ with $\sigma(A_+)\subseteq X$, satisfy condition (α_X).

At last, we get rid of the assumption that A_0 be definitely positive. Let $\mathfrak{M}$ be the subspace of $\mathfrak{H}$ formed by those elements u for which $A_0 u=0$, and let $\mathfrak{N}$ be the orthogonal complement of $\mathfrak{M}$.

It follows from (12) that $A_0 u=0$ implies $(A_k u, u)=0$, hence also $((M_k A_0-A_k)u, u)=0$; since $M_k A_0-A_k\geqq O$, this implies $(M_k A_0-A_k)u=0$, thus $A_k u=0$. We have thus the result that the subspace $\mathfrak{M}$ is mapped on the element 0 by each operator A_k.

If $v\in\mathfrak{N}$, then we have for all $u\in\mathfrak{M}$

$$(A_k v, u)=(v, A_k u)=(v, 0)=0,$$

thus $A_k v\in\mathfrak{N}$, i. e. the subspace $\mathfrak{N}$ is mapped into itself by each operator A_k.

Summing up: $\mathfrak{M}$ and $\mathfrak{N}$ reduce each A_k, and $A_k\mathfrak{M}=(0)$ for $k=0, 1, \ldots$.

In $\mathfrak{N}$, A_0 is positively definite, thus we have, in $\mathfrak{N}$, a representation of the type (14) with a self-adjoint A_+ on a Hilbert space $\mathfrak{H}_+\supseteq\mathfrak{N}$ and with $\sigma(A_+)\subseteq X$. We may suppose that even $\mathfrak{H}_+\supseteq\mathfrak{H}$. In the contrary case, we should only replace $\mathfrak{H}_+$ by $\mathfrak{H}'_+=\mathfrak{H}_+\oplus\mathfrak{M}$ and, by choosing an arbitrary self-adjoint operator C on $\mathfrak{M}$ with $\sigma(C)\subseteq X$, we should consider the operator $A'_+=A_+\oplus C$ on $\mathfrak{H}'_+$, whose spectrum is also contained in X; indeed, if P'_+ and P_+ denote the operators, on $\mathfrak{H}'_+$ and on $\mathfrak{H}_+$, of the projection onto the subspace $\mathfrak{N}$, then we have for $v\in\mathfrak{N}$: $P'_+A'^k_+v=P_+A^k_+v$.

We conclude by observing that the representation (14) is then valid not only in $\mathfrak{N}$, but also in $\mathfrak{M}$, and, consequently, in the whole space $\mathfrak{H}$. As a matter of fact, we have for $u\in\mathfrak{M}$ on the one hand $A_k u=0$, and on the other hand $\|A_0^{\frac{1}{2}}u\|^2=(A_0 u, u)=0$, $\|A_0^{\frac{1}{2}}u\|=0$, thus

$$A_0^{\frac{1}{2}}P_+A^k_+A_0^{\frac{1}{2}}u=0 \qquad (k=0, 1, \ldots).$$

It is easily seen that, conversely, each sequence $\{A_k\}$ of operators on the Hilbert space $\mathfrak{H}$ which may be obtained in the form (14) with the aid of some self-adjoint operator A_+ with $\sigma(A_+)\subseteq X$, on a Hilbert space $\mathfrak{H}_+\supseteq\mathfrak{H}$, consists of bounded self-adjoint operators satisfying the condition (α_X).

Summing up, we have the following

THEOREM 2. *We obtain all the sequences $\{A_k\}$ of bounded self-adjoint operators on a Hilbert space $\mathfrak{H}$, satisfying the condition (α_X) of Theorem 1, in the following way. We choose an arbitrary A_0 on $\mathfrak{H}$, $A_0\geqq O$, an arbitrary Hilbert space $\mathfrak{H}_+$ containing $\mathfrak{H}$ as a (non necessarily proper) subspace, and an arbitrary self-adjoint operator A_+ on $\mathfrak{H}_+$ whose spectrum is contained in the* *set X. Denoting by P_+ the orthogonal projection on $\mathfrak{H}$, we put*

$$A_k = A_0^{\frac{1}{2}} P_+ A_+^k A_0^{\frac{1}{2}} \qquad (k=1,2,\ldots), \tag{15}$$

this formula being evidently valid also for $k=0$.

Let us consider in particular the case $X=[0,1]$. We have

$$p_{m,n}(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} x^{m+k} = x^m(1-x)^n \geqq 0 \quad \text{on } [0,1]$$

$(m, n = 0, 1, \ldots)$, and by a theorem of S. BERNSTEIN,[6] each polynomial which is >0 on $[0,1]$, may be represented as a linear combination of the polynomials $p_{m,n}(x)$ with real non-negative coefficients. Thus, in this case, if we require the condition (α_X) only for the polynomials $p_{m,n}(x)$, then it follows also for all polynomials $p(x)$ which are >0 on $[0,1]$. Moreover, it follows then also for those polynomials $p(x)=c_0+c_1x+\cdots+c_nx^n$ which are $\geqq 0$ on $[0,1]$, for, ε being an arbitrary positive number, $p(x)+\varepsilon>0$ on $[0,1]$ implies $c_0A_0+c_1A_1+\cdots+c_nA_n+\varepsilon A_0 \geqq O$, thus, letting $\varepsilon\to 0$, $c_0A_0+c_1A_1+\cdots$ $\cdots+c_nA_n \geqq O$. So we get the following

COROLLARY. *In order that a sequence $\{A_k\}$ of bounded self-adjoint operators on a Hilbert space $\mathfrak{H}$ be totally monotic, i. e.*

$$A_m - \binom{n}{1} A_{m+1} + \binom{n}{2} A_{m+2} - \ldots + (-1)^n A_{m+n} \geqq O \qquad (m, n = 0, 1, 2, \ldots),$$

it is necessary and sufficient that we have a representation (15) *with a self-adjoint operator A_+ whose spectrum lies in* $[0,1]$.

BOLYAI-INSTITUTE
UNIVERSITY OF SZEGED

(*Received 27 January 1953*)

[6] С. Н. Бернштейн, О найлучшем приближении непрерывных функций посредством многочленов данной степени, Сообщ. Харьковского Матем. Общества, серия 2, **13** (1912), pp. 49—194, in particular § 67; or С. Н. Бернштейн, Собрание сочинений. I (1952), pp. 81—83. See also F. HAUSDORFF, Summationsmethoden und Momentfolgen. I, *Math. Zeitschrift*, **9** (1921), pp. 74—109.

ПРОБЛЕМА МОМЕНТОВ ДЛЯ САМОСОПРЯЖЕННЫХ ОПЕРАТОРОВ

Б. С.-НАДЬ (Сегед)

(Резюме)

Кадисон[1] недавно доказал теорему, так называемую „обобщенное неравенство Шварца“. Эта теорема эквивалентна с следующей:

Теорема 1. Пусть X компактное множество точек на действительной оси, и $A_0, A_1, A_2, \ldots$ — ограниченные самосопряженные операторы гильбертова пространства $\mathfrak{H}$, удовлетворяющие следующим условиям: (α_X) если $c_0+c_1x+\ldots+c_nx^n \geqq 0$ на X, то $c_0A_0+c_1A_1+\ldots+c_nA_n \geqq O$, $(\beta)\,\|A_0\| \leqq 1$. Тогда $A_1^2 \leqq A_2$.

В настоящей статье дается новое доказательство этой же теоремы, опирающееся на известную теорему М. А. Наймарка о представлении всякого обобщенного разложения единицы $F(x)$ пространства $\mathfrak{H}$ в виде $F(x)=P_+E_+(x)$, где $E_+(x)$ обычное (ортогональное) разложение единицы некоторого пространства $\mathfrak{H}_+$, содержащее $\mathfrak{H}$ в качестве подпространства, а P_+ — оператор проектирования на $\mathfrak{H}$.

Этим методом доказывается и следующая

Теорема 2. Мы получаем все последовательности $A_0, A_1, A_2, \ldots$ удовлетворяющие условию (α_X) следующим образом: Мы выберем произвольный ограниченный самосопряженный оператор $A_0 \geqq O$ в $\mathfrak{H}$, и произвольный самосопряженный оператор A_+ с спектром $\sigma(A_+) \subseteq X$ в произвольном гильбертовом пространстве $\mathfrak{H}_+ \supseteq \mathfrak{H}$, и положим

$$(*)\qquad A_k = A_0^{\frac{1}{2}} P_+ A_+^k A_0^{\frac{1}{2}} \qquad (k=0,1,2,\ldots),$$

где P_+ — оператор проектирования на $\mathfrak{H}$.

Следствие. Для того, чтобы последовательность $A_0, A_1, A_2, \ldots$ ограниченных самосопряженных операторов была абсолютно монотонна, т. е.

$$A_m - \binom{n}{1}A_{m+1} + \binom{n}{2}A_{m+2} - \ldots + (-1)^n A_{m+n} \geqq O \qquad (m, n = 0, 1, 2, \ldots),$$

необходимо и достаточно, что имеет место представление (*) с $\sigma(A_+) \subseteq [0, 1]$.

REMARK ON R-EQUIVALENT SPACES

By

TUDOR GANEA (Bucarest)

(*Presented by* O. VARGA)

1. The following problem has been submitted by K. BORSUK, in 1946 [1, Problem 4]:

"Two spaces A and B are said to be R-equivalent, provided that each of them is homeomorphic to a retract of the other. Are the homology and cohomology groups of A and B isomorphic? Assuming in addition that A and B are absolute neighborhood retracts, are the homotopy groups of A and B isomorphic? Are A and B of the same homotopy type?"

The purpose of this note is to point out that an affirmative answer to this problem can be given under some additional restrictions. We shall namely prove that A and B are of the same homotopy type, provided that a certain composite homeomorphism, determined by the R-equivalence, has equicontinuous positive powers.

2. Let A and B be two metric, compact, R-equivalent, absolute neighborhood retracts. Then there exist homeomorphisms

$$f\colon\ B \to A_1 \quad \text{and} \quad g\colon\ A \to B_1,$$

where $A_1 = f(B)$ and $B_1 = g(A)$ are retracts of A and B, under the retractions

$$r\colon\ A \to A_1 \quad \text{and} \quad s\colon\ B \to B_1,$$

respectively. Let $\varphi = fg$ and $\psi = gf$; these are homeomorphisms of A into A and B into B, respectively. Let further

$$A_0 = A, \quad B_0 = B, \quad A_{n+1} = f(B_n), \quad B_{n+1} = g(A_n);$$

it easily follows that

$$A_{n+1} \subset A_n, \quad B_{n+1} \subset B_n \quad \text{and} \quad A_{2n} = \varphi^n(A), \quad B_{2n} = \psi^n(B).$$

Thus A_{2n} and A_{2n+1} are homeomorphic to A and B, respectively. Denote by $f_n = f \mid B_n$, $g_n = g \mid A_n$ the corresponding partial maps, and let

$$r_0 = r, \quad s_0 = s, \quad r_n = f_n s_{n-1} f_{n-1}^{-1}, \quad s_n = g_n r_{n-1} g_{n-1}^{-1};$$

then $\varrho_n = r_{n-1} \ldots r_0$ is a retraction of A onto A_n. Finally, let $D = \bigcap A_n$; this is a compact subset of A and, since $\{A_n\}$ is decreasing, $D = \bigcap \varphi^n(A)$.

If $d \in D$, then $d = \varphi^{m+1}(a_m) = \varphi\varphi^m(a_m)$ for each $m \geqq 0$. Since φ is a homeomorphism, $a = \varphi^{-1}(d)$ is a definite point in A; it follows that

$$a = \varphi^m(a_m) \in \varphi^m(A)$$

holds for each m, hence $a \in D$, $d = \varphi(a) \in \varphi(D)$ and $D \subset \varphi(D)$. Finally, we obtain

$$(*) \qquad D \subset \bigcap \varphi^n(D).$$

3. THEOREM. *If the homeomorphism φ of A into itself has equicontinuous positive powers, then A and B are of the same homotopy type.*

PROOF. A well-known theorem asserts that the sequence $\{\varphi^n\}$ contains a uniformly convergent subsequence $\{\varphi^{n_i}\}$; then $\Phi = \lim_i \varphi^{n_i}$ is a continuous map of A into itself.

If $y = \Phi(a)$, $a \in A$, then

$$\varphi^{n_j}(a) \in \varphi^{n_j}(A) \subset \varphi^{n_i}(A) \qquad \text{if } j \geqq i,$$

hence, since $\varphi^{n_i}(A)$ is closed in A, for each i we have

$$y = \lim_j \varphi^{n_j}(a) \in \varphi^{n_i}(A).$$

Since $\{A_n\}$ is decreasing, it follows that

$$\Phi(A) \subset \bigcap \varphi^{n_i}(A) = D.$$

Any compact subset $X \subset A$ satisfies $\bigcap \varphi^n(X) \subset \Phi(X)$. For, if y is in the intersection, then $y = \varphi^{n_i}(x_i)$, $x_i \in X$, for each i. The sequence $\{x_i\}$ contains a convergent subsequence $\{x_{i_k}\}$ and letting $x = \lim_k x_{i_k} \in X$, it follows that

$$y = \lim_k \varphi^{n_{i_k}}(x_{i_k}) = \Phi(x) \in \Phi(X).$$

Taking $X = A$, we obtain $D \subset \Phi(A)$, hence

$$\Phi(A) = D.$$

Taking $X = D$, we obtain $\bigcap \varphi^n(D) \subset \Phi(D)$, hence $D \subset \Phi(D)$ by (*), and $\Phi(D) \subset \Phi(A) = D$ implies finally

$$\Phi(D) = D.$$

Now, by [3, p. 76], since the φ^n are equicontinuous, there exists an equivalent metric on A, under which every φ^n, hence also Φ, is a shortening. The compacity of D and $\Phi(D) = D$ imply then, by [4, Satz Ib], that the partial map $\tau = \Phi | D$ is an isometry, hence a homeomorphism of D onto itself. The map $\tau^{-1}\Phi$ is then a retraction of A onto D, and D is, like A, an absolute neighborhood retract.

As a consequence, by [1, Lemme (L) & 5, Lemma 8.1], there exists an open set V with $D \subset V \subset A$, and a deformation $h\colon A \times I \to A$ such that:

(1) $$h(a,0)=a \qquad \text{for } a\in A,$$

(2) $$h(d,t)=d \qquad \text{for } d\in D, \quad t\in I,$$

(3) $$h(x,1)\in D \qquad \text{for } x\in V.$$

ϱ_n is a retraction of A onto A_n and, if N is large enough, $D\subset A_n\subset V$ for $n\geqq N$. Let $\eta=\varrho_n h|A_n\times I$, and $\delta(x)=\eta(x,1)$ for $x\in A_n$, $n\geqq N$. Then:

$$\begin{array}{rl} \eta(x,t)\in A_n & \text{for } x\in A_n,\ t\in I; \\ \eta(x,0)=x & \text{for } x\in A_n, \text{ by (1)}; \\ \eta(d,t)=d & \text{for } d\in D,\ t\in I, \text{ by (2)}; \\ \delta(d)=d & \text{for } d\in D; \\ \delta(x)=\eta(x,1)\in D & \text{for } x\in A_n, \text{ by (3)}. \end{array}$$

Thus we have proved that $\delta\colon A_n\to D$ is a deformation retraction, so that A_n and D are of the same homotopy type, if $n\geqq N$. Taking $n\geqq N$, first even and then odd, it follows that A and B are of the same homotopy type, and the proof is complete.

(Received 11 October 1952)

References

1. K. Borsuk, Quelques rétractes singuliers, *Fundamenta Math.*, **24** (1935), pp. 249—258.
2. S. Eilenberg, On the problems of topology, *Annals of Math.*, **50** (1949), pp. 247—260.
3. ——— Sur les groupes compacts d'homéomorphies, *Fundamenta Math.*, **28** (1937), pp. 75—80.
4. H. Freudenthal und W. Hurewicz, Dehnungen, Verkürzungen, Isometrien, *Fundamenta Math.*, **26** (1936), pp. 120—122.
5. O. Hanner, Some theorems on absolute neighborhood retracts, *Arkiv för Matematik*, **1** (1951), pp. 389—408.

ЗАМЕЧАНИЕ ОБ R-ЭКВИВАЛЕНТНЫХ ПРОСТРАНТСВАХ

Т. ГАНЕА (Бухарест)

(Р е з ю м е)

Пусть A и B две компактные, метрические абсолютные ретракты окрестностей, пусть далее g и f являются гомеоморфизмами A в B и B в A соответственно. Если $g(A)$ и $f(B)$ суть ретракты от B и A соответственно, и если гомеоморфизм $\varphi=fg$ от A в самое себя имеет равнонепрерывные положительные степени, то A и B обладают одним и тем же гомотопным типом.

RESTGLIED EINES TAUBERSCHEN SATZES, II

Von
GÉZA FREUD (Budapest)
(*Vorgelegt von* P. TURÁN)

I. Einleitung

In einer vorigen Arbeit (G. FREUD [2], im folgenden als „erste Mitteilung" zitiert), haben wir folgendes bewiesen: Es sei $\tau(t)$ eine in $0 \leqq t < \infty$ nicht abnehmende Funktion, ferner $f(t) \geqq 0$, $\alpha > 0$ und S von s unabhängige Konstanten, schließlich sei

$$(1) \qquad F(s) = \int_{t=0}^{\infty} f(t) e^{-st} d\tau = S \frac{\Gamma(\alpha+1)}{s^{\alpha}} [1 + r(s)] \qquad \text{für } s > 0$$

mit

$$(2) \qquad |r(s)| < R(s),$$

wobei $R(s)$ eine für $S > 0$ definierte, monoton zunehmende Funktion ist, für welche $R(0) = 0$ und

$$(3) \qquad R(ks) < e^{c_1 k} R(s)$$

erfüllt ist; dann gilt für $\beta \geqq 0$

$$(4) \qquad \int_{t=0}^{x} t^{\beta} f(t)\, d\tau = \frac{\alpha}{\alpha+\beta} S x^{\alpha+\beta} [1 + \varrho(x)],$$

wobei

$$(5) \qquad |\varrho(x)| < \frac{c_2}{\log \dfrac{1}{R(1/x)}}.$$

In der vorliegenden Arbeit werden wir zeigen:[1]

SATZ. *Wenn wir die linke Seite von* (4) *m-mal nach der oberen Grenze integrieren, d. h. wenn wir das Cesàrosche Integralmittel von* (4) *bilden* (vgl.

[1] А. Г. Постников [8] und J. KOREVAAR [6] erhielten ähnliche Sätze, aber ihre Restglieder sind schwächer als (5).

G. H. HARDY [4]), *dann kann die Abschätzung* (5) *zu*

$$\text{(5a)} \qquad O\left\{\left(\log \frac{1}{R(1/x)}\right)^{-m-1}\right\}$$

verschärft werden. [2]

Unser Beweis gründet sich, wie auch in der ersten Mitteilung, auf eine Schlußweise von J. KARAMATA; im weiteren brauchen wir eine Verallgemeinerung unseres Approximationssatzes, die wir in Teil II der ersten Mitteilung bewiesen haben. Diese Verallgemeinerung wird in Teil II dieser Arbeit behandelt. Teil III enthält den Beweis unseres Satzes für gerades m, und in Teil IV wird hieraus die Gültigkeit des Satzes für ungerades m gefolgert. Endlich übertragen wir in Teil V unseren Satz auf Taubersche Reihensätze vom Typ $A \to (C, k)$. In der folgenden dritten Mitteilung wollen wir uns mit Tauberschen Sätzen vom Typ $A \to K$ beschäftigen. [3]

II. Verallgemeinerung des Approximationssatzes

HILFSSATZ. Es sei $\nu \geqq 0$ eine ganze Zahl, $g(x)$ sei im Intervall $(0, 1)$ 2ν-mal differenzierbar, $g^{(2\nu)}(x)$ sei in $(0, 1)$ stückweise stetig und genüge in jedem Stetigkeitsintervall der Lipschitz'schen Bedingung

$$\text{(6)} \qquad |g^{(2\nu)}(x_1) - g^{(2\nu)}(x_2)| < c_3 |x_1 - x_2|.$$

Hier gehören x_1 und x_2 zu demselben Stetigkeitsintervall von $g^{(2\nu)}(x)$, und wenn z. B. x_1 eine Sprungstelle von $g^{(2\nu)}(x)$ ist, dann setze man $g^{(2\nu)}(x_1) = g^{(2\nu)}(x_1 + 0)$ für $x_2 > x_1$ und $g^{(2\nu)}(x_1) = g^{(2\nu)}(x_1 - 0)$ für $x_2 < x_1$.

Dann gibt es für jedes ganzzahlige $n > n_0$ Polynome $\varphi_n(x)$ und $\Phi_n(x)$ vom niedrigeren Grade als $2n$, für welche

$$\text{(7)} \qquad \varphi_n(x) \leqq g(x) \leqq \Phi_n(x)$$

und

$$\text{(8)} \qquad \int_0^1 (\log 1/x)^{\alpha-1} [\Phi_n(x) - \varphi_n(x)] dx < \frac{c_4}{n^{2\nu+1}},$$

wobei n_0 und c_4 nur von α und der Wahl von $g(x)$ abhängig sind.

BEWEIS. Es sei ξ eine gegebene Stelle in $(0, 1)$, $t_0(x), t_1(x), \ldots, t_n(x), \ldots$ die dem Grade nach aufsteigende Folge der in $(0, 1)$ mit der Gewichtsfunktion $(\log 1/x)^{\alpha-1}$ orthogonalen Polynome; und die Wurzel von $t_n(x)$ seien $0 < x_{1n} < x_{2n} < \cdots < x_{nn} < 1$. Die ganze Zahl $r = r(n)$ sei durch

$$\text{(9)} \qquad x_{rn} \leqq \xi < x_{r+1,\, n}$$

definiert.

[2] Aus einem Beispiel von J. KOREVAAR [6] folgt, daß die Größenordnung der Restglieder (5) und (5a) nicht weiter verschärft werden kann.

[3] $A \to (C, k)$ bedeutet: aus A-Summierbarkeit folgt (C, k)-Summierbarkeit. $A \to K$ bedeutet: aus A-Summierbarkeit folgt Konvergenz.

Wir benützen die Markov—Stieltjes'sche Polynomkonstruktion in einer etwas abgeänderten Form (vgl. [7], [9] und Teil II der ersten Mitteilung). $P_{n\xi}(x), p_{n\xi}(x)$ seien Polynome vom höchstens $2n-2\nu-2$-ten Grade, die durch die Bedingungen

$$(10)\quad \begin{aligned} &P_{n\xi}(x_{1n}) = P_{n\xi}(x_{2n}) = \cdots = P_{n\xi}(x_{rn}) = P_{n\xi}(x_{r+1,\,n}) = 1,\\ &P_{n\xi}(x_{r+\nu+2,\,n}) = \cdots = P_{n\xi}(x_{nn}) = 0,\\ &P'_{n\xi}(x_{1n}) = P'_{n\xi}(x_{2n}) = \cdots = P'_{n\xi}(x_{rn}) = 0,\\ &P'_{n\xi}(x_{r+\nu+2,\,n}) = \cdots = P'_{n\xi}(x_{nn}) = 0 \end{aligned}$$

bzw.

$$(11)\quad \begin{aligned} &p_{n\xi}(x_{1n}) = p_{n\xi}(x_{2n}) = \cdots = p_{n\xi}(x_{r-\nu,\,n}) = 1,\\ &p_{n\xi}(x_{rn}) = p_{n\xi}(x_{r+1,\,n}) = \cdots = p_{n\xi}(x_{nn}) = 0,\\ &p'_{n\xi}(x_{1n}) = p'_{n\xi}(x_{2n}) = \cdots = p'_{n\xi}(x_{r-\nu,\,n}) = 0,\\ &p'_{n\xi}(x_{r+1,\,n}) = \cdots = p'_{n\xi}(x_{nn}) = 0 \end{aligned}$$

definiert sind. Aus der Diskussion der Extremalstellen und des Verlaufes dieser Polynome erhält man

$$(12)\quad P_{n\xi}(x) \geqq \begin{cases} 1 & \text{für } 0 \leqq x \leqq x_{r+1,\,n},\\ 0 & \text{für } x_{r+1,\,n} \leqq x \leqq 1 \end{cases}$$

und

$$(13)\quad p_{n\xi}(x) \leqq \begin{cases} 1 & \text{für } 0 \leqq x \leqq x_{rn},\\ 0 & \text{für } x_{rn} \leqq x \leqq 1, \end{cases}$$

ferner

$$(14)\quad \begin{aligned} &0 \leqq P_{n\xi}(x) \leqq 1 \quad \text{für } x_{r+1,\,n} \leqq x \leqq x_{r+\nu+2,\,n},\\ &0 \leqq p_{n\xi}(x) \leqq 1 \quad \text{für } x_{r-\nu,\,n} \leqq x \leqq x_{rn}. \end{aligned}$$

Aus (9), (12), (13) folgt, daß die Polynome $(x-\xi)^{2\nu}P_n(x), (x-\xi)^{2\nu}p_n(x)$, deren Grad niedriger als $2n-1$ ist, der Ungleichung

$$(15)\quad (x-\xi)^{2\nu}p_{n\xi}(x) \leqq f_\xi(x) \leqq (x-\xi)^{2\nu}P_{n\xi}(x)$$

genügen, wobei

$$(16)\quad f_\xi(x) = \begin{cases} (x-\xi)^{2\nu} & \text{für } 0 \leqq x \leqq \xi,\\ 0 & \text{für } \xi < x \leqq 1. \end{cases}$$

Nach der Gauß—Jacobischen Formel der mechanischen Quadratur ist

$$(17)\quad \begin{aligned} &\int_0^1 (\log 1/x)^{\alpha-1}(x-\xi)^{2\nu}[P_{n\xi}(x) - p_{n\xi}(x)]\,dx =\\ &= \sum_{k=r-\nu+1}^{r+\nu+1} (x-\xi)^{2\nu}[P_{n\xi}(x_{kn}) - p_{n\xi}(x_{kn})]\,\lambda_{kn}, \end{aligned}$$

wo $\{\lambda_{kn}\}$ die Cotes'schen Zahlen der mechanischen Quadratur über $(0, 1)$ mit der Gewichtsfunktion $(\log 1/x)^{\alpha-1}$ und den Grundpunkten $\{x_{kn}\}$ bedeuten. Nach dem Anhange der ersten Mitteilung gilt

$$(18)\quad 0 < \lambda_{kn} < \frac{c_5}{n}\,.$$

Aus einem Satz von ERDŐS und TURÁN [1] folgt nun

$$(19) \qquad |x_k-\xi| < \frac{c_6\nu}{n} = \frac{c_7}{n} \quad \text{für} \quad r-\nu+1 \leqq k \leqq r+\nu+1.$$

Da nach (12), (13) und (14) für $r-\nu+1 \leqq k \leqq r+\nu+1$, $0 \leqq P_{n\xi}(x_{kn}) - p_{n\xi}(x_{kn}) \leqq 1$ ist, erhalten wir aus (17), (18) und (19)

$$(20) \qquad 0 < \int_0^1 (\log 1/x)^{\alpha-1}(x-\xi)^{2\nu}[P_{n\xi}(x)-p_{n\xi}(x)]\,dx < \frac{c_8}{n^{2\nu+1}}.$$

Nun kann die Funktion $g(x)$ in der Form

$$(21) \qquad g(x) = h(x) + H(x)$$

zerlegt werden, wobei

$$(22) \qquad H(x) = \sum_{\xi_k} \frac{g^{(2\nu)}(\xi_k-0)-g^{(2\nu)}(\xi_k+0)}{(2\nu)!} f_{\xi_k}(x)$$

(die Summation wird über alle Sprungstellen ξ_k von $g^{(2\nu)}(x)$ erstreckt), $h(x)$ ist 2ν-mal differenzierbar, und $h^{(2\nu)}(x)$ genügt gleichmäßig in (0, 1) der Lipschitz'schen Bedingung

$$(23) \qquad |h^{(2\nu)}(x_1)-h^{(2\nu)}(x_2)| < c_3|x_1-x_2|.$$

Wir setzen

$$(24) \qquad \begin{aligned} \vartheta_n(x) &= \sum_{\xi_k} \frac{g^{(2\nu)}(\xi_k-0)-g^{(2\nu)}(\xi_k+0)}{(2\nu)!}(x-\xi)^{2\nu} p^*_{n\xi_k}(x), \\ \theta_n(x) &= \sum_{\xi_k} \frac{g^{(2\nu)}(\xi_k-0)-g^{(2\nu)}(\xi_k+0)}{(2\nu)!}(x-\xi)^{2\nu} P^*_{n\xi_k}(x) \end{aligned}$$

mit

$$(25) \qquad \begin{aligned} p^*_{n\xi}(x) &= \begin{cases} p_{n\xi}(x) & \text{für } g^{(2\nu)}(\xi-0)-g^{(2\nu)}(\xi+0) > 0, \\ P_{n\xi}(x) & \text{für } g^{(2\nu)}(\xi-0)-g^{(2\nu)}(\xi+0) < 0, \end{cases} \\ P^*_{n\xi}(x) &= \begin{cases} P_{n\xi}(x) & \text{für } g^{(2\nu)}(\xi-0)-g^{(2\nu)}(\xi+0) > 0, \\ p_{n\xi}(x) & \text{für } g^{(2\nu)}(\xi-0)-g^{(2\nu)}(\xi+0) < 0. \end{cases} \end{aligned}$$

Somit sind $\vartheta_n(x)$ und $\theta_n(x)$ Polynome vom höchstens $2n-2$-ten Grade, und

$$(26) \qquad \vartheta_n(x) \leqq H(x) \leqq \theta_n(x),$$

$$(27) \qquad \int_0^1 (\log 1/x)^{\alpha-1}[\theta_n(x)-\vartheta_n(x)]\,dx < \frac{c_9}{n^{2\nu+1}}.$$

Aus (23) folgt nach einem Satz von D. JACKSON [3], daß es ein Polynom von höchstens n-tem Grade $\chi_n(x)$ gibt, für welches im ganzen Intervall (0, 1)

$$(28) \qquad |h(x)-\chi_n(x)| \leqq \frac{c_{10}}{n^{2\nu+1}}$$

erfüllt ist.

Setzen wir endlich

$$
(29)\qquad \begin{aligned}
\varphi_n(x) &= -\frac{c_{10}}{n^{2\nu+1}} + \chi_n(x) + \vartheta_n(x),\\
\Phi_n(x) &= \frac{c_{10}}{n^{2\nu+1}} + \chi_n(x) + \theta_n(x),
\end{aligned}
$$

dann ist der Grad von $\varphi_n(x)$ bzw. $\Phi_n(x)$ höchstens $2n-2$, und es gilt wegen (21), (26), (27), (28)

$$
(30)\qquad \varphi_n(x) \leqq g(x) \leqq \Phi_n(x),
$$

$$
(31)\qquad \int_0^1 (\log 1/x)^{\alpha-1}[\Phi_n(x) - \varphi_n(x)]\,dx < \frac{c_{11}}{n^{2\nu+1}},
$$

w. z. b. w.

Im weiteren sei $\beta \geqq 0$ und

$$
(32)\qquad g(x) = \begin{cases} x^{-1}(\log 1/x)^{\beta}(1-\log 1/x)^{2\nu} & \text{für } e^{-1} \leqq x \leqq 1,\\ 0 & \text{für } 0 \leqq x \leqq e^{-1}. \end{cases}
$$

Aus (31) folgt — mit derselben Schlußweise, wie in Teil III der ersten Mitteilung —

$$
(33)\qquad |\Phi_n(x)| \leqq c_{12} n^{c_{13}}, \quad |\varphi_n(x)| \leqq c_{12} n^{c_{13}},
$$

und wenn

$$
(34)\qquad \Phi_n(x) = \sum_{k=0}^{2n-2} B_k x^k, \quad \varphi_n(x) = \sum_{k=0}^{2n-2} b_k x^k,
$$

dann wird (vgl. Gleichung (46) der ersten Mitteilung)

$$
(35)\qquad \sum_{k=0}^{2n-2} |B_k| < c_{13} e^{c_{14} n}, \quad \sum_{k=0}^{2n-2} |b_k| < c_{13} e^{c_{14} n},
$$

und hieraus ist leicht ersichtlich, daß

$$
(36)\qquad \sum_{k=0}^{2n-2} \frac{|B_k| e^{c_1 k}}{(k+1)^{\alpha}} < c_{15} e^{c_{16} n}, \quad \sum_{k=0}^{2n-2} \frac{|b_k| e^{c_1 k}}{(k+1)^{\alpha}} < c_1\, e^{c_{16} n}.
$$

III. Beweis des Hauptsatzes für gerades m

Wir erhalten nach dem Vorbild von Gleichung (49) der ersten Mitteilung[4] mit der Methode von KARAMATA:

$$
(37)\qquad \int_{t=0}^{\infty} f(t)\, \Phi_n(e^{-st})\, e^{-st}\, d\tau = S \frac{\alpha}{s^{\alpha}} \int_0^{\infty} t^{\alpha-1}\, \Phi_n(e^{-t})\, e^{-t}\, dt + S \frac{\Gamma(\alpha+1)}{s^{\alpha}} \sum_{k=0}^{2n-2} B_k \frac{r(ks)}{(k+1)^{\alpha}}.
$$

[4] In jener Mitteilung ist im Integral an der rechten Seite ein Faktor e^{-t} versehentlich weggeblieben.

Wegen (30), (31) und (32) ist

$$\int_0^\infty t^{\alpha-1}\Phi_n(e^{-t})e^{-t}dt \leqq \int_0^\infty t^{\alpha-1}g(e^{-t})e^{-t}dt + \int_0^\infty t^{\alpha-1}[\Phi_n(e^{-t})-\varphi_n(e^{-t})]\,e^{-t}dt =$$

$$(38)\qquad = \int_0^1 (1-t)^{2\nu}t^{\alpha+\beta-1}dt + \int_0^1 (\log 1/x)^{\alpha-1}[\Phi_n(x)-\varphi_n(x)]\,dx <$$

$$< \frac{(2\nu)!}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)} + \frac{c_{11}}{n^{2\nu+1}}.$$

Andererseits folgt aus (2), (3) und (36)

$$(39)\qquad \left|\sum_{k=0}^{2n-2} B_k \frac{r(ks)}{(k+1)^\alpha}\right| < R(s)\sum_{k=0}^{2n-2}\frac{|B_k|\,e^{c_1 k}}{(k+1)^\alpha} < c_{15}e^{c_{16}n}R(s).$$

Aus (30), (37), (38) und (39) erhalten wir

$$s^{\beta+2\nu}\int_{t=0}^{1/s} t^\beta f(t)(1/s-t)^{2\nu}d\tau = \int_{t=0}^\infty f(t)g(e^{-st})\,e^{-st}d\tau < \int_{t=0}^\infty f(t)\,\Phi_n(e^{-st})e^{-st}\,d\tau <$$

$$(40)\qquad < \frac{S\alpha}{s^\alpha}\left[\frac{(2\nu)!}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)} + \frac{c_{11}}{n^{2\nu+1}} + c_{17}e^{c_{16}n}R(s)\right].$$

Setzen wir

$$s = 1/x, \qquad n = \left[\frac{1}{2c_{16}}\log\frac{1}{R(1/x)}\right] + 1,$$

dann erhalten wir aus (40)

$$(41)\qquad \frac{1}{(2\nu)!}\int_{t=0}^{x} t^\beta f(t)(x-t)^{2\nu}d\tau < \frac{\alpha S x^{\alpha+\beta+2\nu}}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)}[1+\omega(x)],$$

wo

$$(42)\qquad \omega(x) < c_{18}\left(\log\frac{1}{R(1/x)}\right)^{-2\nu-1} + c_{19}\{R(1/x)\}^{1/2} < c_{20}\left(\log\frac{1}{R(1/x)}\right)^{-2\nu-1}$$

für $x > x_0$.

Rechnen wir statt $\Phi_n(x)$ mit $\varphi_n(x)$, so erhalten wir eine ähnliche untere Abschätzung. Beide Abschätzungen zusammen ergeben

$$(43)\qquad \frac{1}{(2\nu)!}\int_{t=0}^{x} t^\beta f(t)(x-t)^{2\nu}d\tau =$$

$$= \frac{\alpha S x^{\alpha+\beta+2\nu}}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)}\left[1 + O\left\{\left(\log\frac{1}{R\left(\frac{1}{x}\right)}\right)^{-2\nu-1}\right\}\right].$$

Bekanntlich ist

$$(44)\qquad \frac{1}{(2\nu)!}\int_{t=0}^{x} t^\beta f(t)(x-t)^{2\nu}d\tau = \int_{u_{2\nu}=0}^{x}\cdots\int_{u_1=0}^{u_2}\int_{t=0}^{u_1} t^\beta f(t)d\tau\, du_1 du_2\cdots du_{2\nu}.$$

Damit ist unser Satz für gerades $m = 2\nu$ bewiesen.

IV. Beweis des Hauptsatzes für ungerades m

HILFSSATZ. $F(u)$ sei eine in $(0, \infty)$ definierte nicht abnehmende positive differenzierbare Funktion, $\nu \geqq 1$ und

$$\int_0^x F(u)\,du = Ax^\gamma + x^\gamma O(\{K(x)\}^{-2\nu-1}), \tag{45}$$

ferner

$$F'(x) = A\gamma(\gamma-1)x^{\gamma-2} + x^{\gamma-2} O(\{K(x)\}^{-2\nu+1}). \tag{46}$$

$K(x)$ ist eine monoton gegen ∞ strebende, positive Funktion, die der Bedingung

$$K(2x) = O\{K(x)\} \tag{47}$$

genügt. Wir behaupten, daß aus (45), (46) und (47)

$$F(x) = A\gamma x^{\gamma-1} + x^{\gamma-1} O(\{K(x)\}^{-2\nu}) \tag{48}$$

folgt.[5]

BEWEIS. x sei so groß, daß $K(x) > 1$, und es sei $x < x_1 < 2x$. Dann wird infolge (46)

$$\begin{aligned}\int_{x_1}^{x} F(u)du &= (x_1-x)F(x) + \int_x^{x_1} du \int_x^u F'(v)\,dv = \\ &= (x_1-x)F(x) + A(x_1^\gamma - x^\gamma) - A\gamma x^{\gamma-1}(x_1-x) + (x_1-x)^2 x^{\gamma-2} O(\{K(x)\}^{-2\nu+1}).\end{aligned} \tag{49}$$

Beim letzten Glied haben wir berücksichtigt, daß $x_1 < 2x$ und somit $K(x_1) = O\{K(x)\}$. Aus (45) erhalten wir

$$\int_x^{x_1} F(u)\,du = A(x_1^\gamma - x^\gamma) + x^\gamma O(\{K(x)\}^{-2\nu-1}). \tag{50}$$

Aus (49) und (50) folgt

$$F(x) = A\gamma x^{\gamma-1} + (x_1-x)x^{\gamma-2} O(\{K(x)\}^{-2\nu+1}) + (x_1-x)^{-1} x^\gamma O(\{K(x)\}^{-2\nu-1}). \tag{51}$$

Setzen wir in (51) $x_1 - x = \dfrac{x}{K(x)}$, so ergibt sich (48), w. z. b. w.

Wir setzen jetzt

$$F(x) = \frac{1}{(2\nu-1)!} \int_{t=0}^{x} t^\beta f(t)(x-t)^{2\nu-1}\,d\tau\,;$$

dann ist (45) und (46) mit

$$A = \frac{\alpha S}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)}, \quad \gamma = \alpha+\beta+2\nu, \quad K(x) = \log\frac{1}{R(1/x)}$$

wegen (43) und (44) erfüllt. (47) ist infolge (3) befriedigt. Also ergibt sich

[5] Dieser Hilfssatz ist eine Weiterentwickelung einer Schlußweise von J. E. LITTLEWOOD (*Proc. London Math. Soc.*, 9 (1911), S. 434—448).

aus unserem Hilfssatz

$$\frac{1}{(2\nu-1)!}\int\limits_{t=0}^{x} t^{\beta} f(t)(x-t)^{2\nu-1}\,d\tau = \tag{52}$$

$$= \frac{\alpha S x^{\alpha+\beta+2\nu-1}}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+2\nu)}\left[1+O\left(\left\{\log\frac{1}{R(1/x)}\right\}^{-2\nu}\right)\right].$$

Somit ist unser Satz sowohl für gerades m (Gl. (43), $m=2\nu$) als auch für ungerades m (Gl. (52), $m=2\nu-1$) bewiesen.

V. Anwendung auf Taubersche Sätze vom Typ $A\to(C, m)$

Es sei $\alpha=1$, $\tau(t)=[t+1]$, $f(\nu)=s_\nu$. Dann lautet unsere Bedingung (1) folgendermaßen:

$$\sum_{\nu=0}^{\infty} s_\nu e^{-\nu s} = \frac{S}{s}[1+O(R(s))]. \tag{53}$$

Im weiteren sei $s=O\{R(s)\}$. Dann kann $s_\nu \geqq 0$ durch die schwächere Bedingung $s_\nu > -K$ ersetzt werden. (S. die Einleitung der ersten Mitteilung.) Nach (43) und (52) gilt also für jedes ganzzahlige $m\geqq 0$, wenn wir in diese Gleichungen $x=n+1$ für $m\geqq 1$ und $x=n+1/2$ für $m=0$ setzen:

$$\sum_{r=0}^{n}(n+1-r)^m s_r = \frac{n^{m+1}}{m+1} S\left[1+O\left(\left\{\log\frac{1}{R(1/n)}\right\}^{-m-1}\right)\right]. \tag{54}$$

Die Cesàroschen Summen k-ter Ordnung sind durch

$$C_n^k = \sum_{r=0}^{n}\binom{n+k-1-r}{k-1} s_r \tag{55}$$

definiert. $\binom{n+k-1-r}{k-1}$ ist ein Polynom $k-1$-ten Grades von $n+1-r$, also gilt

$$\binom{n+k-1-r}{k-1} = \frac{1}{(k-1)!}(n+1-r)^{k-1} + \sum_{m=0}^{k-2}\alpha_{mk}(n+1-r)^m. \tag{56}$$

Aus (54) und (56) folgt unter Beachtung von (55):

$$C_n^k = \frac{n^k}{k!} S\left[1+O\left(\left\{\log\frac{1}{R(1/n)}\right\}^{-k}\right)\right],$$

und endlich für die Cesàroschen Mittel k-ter Ordnung

$$\sigma_n^k = \frac{C_n^k}{\binom{n+k}{k}} = S + O\left(\left\{\log\frac{1}{R(1/n)}\right\}^{-k}\right). \tag{57}$$

(Eingegangen am 10. Mai 1952.)

Literaturverzeichnis

[1] P. Erdős—P. Turán, On interpolation, III. *Annals of Math.*, **41** (1940), Theorem VIII, S. 538.

[2] G. Freud, Restglied eines Tauberschen Satzes, I. *Acta Math. Acad. Sci. Hung.*, **2** (1951), S. 299—308.

[3] D. Jackson, *The theory of approximation*, Amer. Math. Soc. Coll. Publ., Vol. XI, (New York, 1930), Theorem VII, S. 17.

[4] G. H. Hardy, *Divergent series* (Oxford, 1949), § 5.14, S. 110.

[5] J. Karamata, Über die Hardy–Littlewoodsche Umkehrung des Abelschen Stetigkeitssatzes, *Math. Zeitschrift*, **32** (1930), S. 519—520.

[6] J. Korevaar, An estimate of the error in Tauberian Theorems for power series, *Duke Math. Journ.*, **18** (1951), S. 731—733.

[7] А. А. Марков, Доказательство некоторых неравенств П. Л. Чебышева, Избранные Труды (Москва, 1948), S. 20—21.

[8] А. Г. Постников, Остаточный член в тауберовой теореме Харди и Литтльвуда, Доклады А. Н. СССР, **77** (1951), S. 193—196.

[9] Th. J. Stieltjes, Quelques recherches sur la théorie des quadratur dites mécaniques, *Oeuvres Complétes* (Groningen, 1914), Bd. 1, S. 377—394.

ОБ ОСТАТОЧНОМ ЧЛЕНЕ НЕКОТОРОЙ ТЕОРЕМЫ ТИПА ТАУБЕРА, II

Г. ФРАЙД (Будапешт)

(Резюме)

Пусть $\tau(t)$ монотонно возрастающая функция $(0 \leqq t < \infty)$ и $f(t) \geqq 0$, если $t \geqq 0$, и интеграл Лебега—Стилтьеса

$$F(s) = \int_{t=0}^{\infty} f(t) e^{-st} d\tau(t)$$

сходится при любом $s > 0$. Пусть при $s \to +0$

$$F(s) = S \frac{\Gamma(\alpha+1)}{s^{\alpha}} [1 + r(s)], \qquad |r(s)| < R(s),$$

где $R(s)$ монотонно возрастает, $R(0) = 0$ и $R(ks) < \exp(c_1 k) . R(s)$, где c_1 не зависит от k и s. Доказывается, что в этом случае при $x \to \infty$

$$\int_{t=0}^{x} t^{\beta} f(t) (x-t)^m d\tau(t) = \frac{m!\, \alpha S x^{\alpha+\beta+m}}{(\alpha+\beta)(\alpha+\beta+1)\dots(\alpha+\beta+m)} \left[1 + O\left\{\left(\log \frac{1}{R(1/x)}\right)^{-m-1}\right\}\right].$$

Из этого следует:

Пусть $s_\nu \geqq -K$, $s = O\{R(s)\}$, $R(0) = 0$, $R(ks) < \exp(c_1 k) . R(s)$ и

$$\sum_{\nu=0}^{\infty} s_\nu e^{-\nu s} = \frac{S}{s} [1 + O(R(s))], \qquad s \to +0.$$

В этом случае Чезаровские средние

$$\sigma_n^{(k)} = \frac{1}{\binom{n+k}{k}} \sum_{r=0}^{n} \binom{n+k-1-r}{k-1} s_r = S + O\left\{\left(\log \frac{1}{R(1/n)}\right)^{-k}\right\}.$$

BEMERKUNGEN ÜBER DIE MULTIPLIKATION VON VEKTOREN UND QUATERNIONEN

Von

J. ACZÉL (Debrecen)

(Vorgelegt von G. HAJÓS)

I. In dieser Arbeit möchte ich ein Paar methodische Bemerkungen über eine von matematischem Standpunkt aus natürlich erscheinende Begründung der Multiplikationsregeln von Vektoren und Quaternionen machen.

Ich bin dessen bewußt, daß mehrere dieser Betrachtungen in ihren Einzelheiten wohl längst bekannt[1] sein mögen. Hier möchte ich nur auf die einheitlichen Methoden Gewicht legen, mit denen im folgenden gearbeitet wird. Wir werden zwei, von einander unabhängige Wege folgen. Der eine ist rein elementargeometrisch, der andere gehört der Theorie der Funktionalgleichungen zu. (Die üblichen Methoden erzielen einige der hier folgenden Resultate auf rein algebraischem Wege[1]). Es sei hier noch bemerkt, daß auch noch ein trigonometrisches Verfahren vom Verf. gefunden wurde, das auf den Sinus- und Cosinus-Sätzen beruht. In dem hier vorgelegten geometrischen Gedankengang (II a) stammt ein Teil von G. HAJÓS und ein größerer Teil von T. SZELE her. Diese Betrachtungen wurden hier mit liebenswürdiger Erlaubnis der Verfasser mitgeteilt.[2]

Wir bemerken, daß auch die Addition von Vektoren und Quaternionen bekanntlich in ganz ähnlicher Weise begründet werden kann.[3] Ähnliches gilt auch für die Multiplikation mit Skalaren. Diese Operationen werden bei der Definition der behandelten Multiplikationen als gegeben vorausgesetzt.

Im folgenden werden die Quaternionen als formale Summen von skalaren Größen und Raumvektoren aufgefaßt. Die Skalaren bzw. die Vektoren

[1] S. z. B.: G. FROBENIUS, Über lineare Substitutionen und bilineare Formen, *Journal für reine und angew. Math.*, **84** (1897), S. 1—63, und D. E. RICHMOND, Complex numbers and vector algebra, *Amer. Math. Monthly*, **58** (1951), S. 622—628.

[2] Das hier mitgeteilte bildet nur einen kleinen Teil der Untersuchungen von G. HAJÓS und T. SZELE über diese Fragen. — Auch in der Anfertigung dieser Arbeit verdanke ich ihnen mehrere wertvolle Ratschläge.

[3] Siehe z. B.: E. PICARD, *Leçons sur quelques équations fonctionelles avec des applications à divers problêmes d'analyse et de physique mathématique* (Paris, Gauthier-Villars, 1928, 1950); I. Ch. III. 1, 2.

werden mit kursiven bzw. fetten Typen bezeichnet. Insbesondere bedeutet $\mathbf{e}$ stets Einheitsvektoren, d. h. Vektoren vom Betrag 1.

Unser Verfahren wird darin bestehen, daß wir gewisse Postulate über die Multiplikation annehmen, aus denen dann die üblichen Multiplikationsregeln gewonnen werden. *Das grundlegende Postulat wird die Distributivität der Multiplikation bezüglich der Addition sein.*

II. Wir beginnen mit der Multiplikation von Vektoren. Wir nehmen zweierlei Multiplikationen

$$\mathbf{a}\cdot\mathbf{b} \quad \text{und} \quad \mathbf{a}\times\mathbf{b}$$

an, von denen die erste zu einem Skalar, die zweite zu einem Vektor führt.

SATZ 1. *Voraussetzungen:*

1) *In dem Raume ist keine Richtung ausgezeichnet. Darunter wollen wir verstehen, daß die betrachteten Multiplikationen drehungsautomorph sind (d. h. bei einer Drehung des Raumes bleibt das skalare bzw. vektorielle Produkt ungeändert bzw. ebenso gedreht).*

2) *Beide Multiplikationen sind (rechtsseitig) distributiv:*

$$(\mathbf{a}+\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c},$$
$$(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}.$$

3) *Skalaren lassen sich aus beiden Multiplikationen ausheben:*

$$(c\mathbf{a})\cdot\mathbf{b}=c(\mathbf{a}\cdot\mathbf{b})=\mathbf{a}\cdot(c\mathbf{b}),$$
$$(c\mathbf{a})\times\mathbf{b}=c(\mathbf{a}\times\mathbf{b})=\mathbf{a}\times(c\mathbf{b}).$$

Behauptung: $\mathbf{a}\cdot\mathbf{b}$ *bzw.* $\mathbf{a}\times\mathbf{b}$ *ist von einer multiplikativen Konstanten abgesehen das übliche skalare bzw. vektorielle Produkt.*

Bemerkung: Für rationale Skalaren würde 3) *schon aus der beiderseitigen Distributivität folgen;* 2) *und* 3) *zusammen bedeuten die Linearität der beiden Produkte.*

Aus den Voraussetzungen 1) und 3) folgt

$$\text{(1)} \qquad \mathbf{a}\cdot\mathbf{b}=0, \qquad \text{falls } \mathbf{a}\perp\mathbf{b},$$
$$\text{(2)} \qquad \mathbf{a}\times\mathbf{b}=0, \qquad \text{falls } \mathbf{a}\parallel\mathbf{b}.$$

Um (1) bzw. (2) zu beweisen, genügt es laut 3) das Verschwinden des skalaren bzw. vektoriellen Produktes zweier orthogonalen bzw. gleichen *Einheits*vektoren zu zeigen.

Es sei zuerst $\mathbf{e}_1\perp\mathbf{e}_2$. Dieses Paar geht mittels einer Drehung mit π um $\mathbf{e}_2$ in das Paar $-\mathbf{e}_1, \mathbf{e}_2$ über. Also muß nach 1) und 3)

$$\mathbf{e}_1\cdot\mathbf{e}_2=(-\mathbf{e}_1)\cdot\mathbf{e}_2=-\mathbf{e}_1\cdot\mathbf{e}_2, \quad \text{d. h.} \quad \mathbf{e}_1\cdot\mathbf{e}_2=0$$

sein.

Andererseits bilden wir das Produkt $\mathbf{e}\times\mathbf{e}$. Aus 1) ist es klar, daß $\mathbf{e}\times\mathbf{e}$ nur ein Vektor von der Richtung $\mathbf{e}$ (oder $-\mathbf{e}$) und mit einem von der Richtung von $\mathbf{e}$ unabhängigen absoluten Betrage sein kann, d. h.

$$\mathbf{e} \times \mathbf{e} = k\mathbf{e}.$$

Da aber $\mathbf{e}$ durch Drehung mit π in $-\mathbf{e}$ übergeht, muß nach 1) und 3)

$$-k\mathbf{e} = k(-\mathbf{e}) = (-\mathbf{e}) \times (-\mathbf{e}) = \mathbf{e} \times \mathbf{e} = k\mathbf{e},$$

also

$$k = 0 \quad \text{und} \quad \mathbf{e} \times \mathbf{e} = 0$$

sein.

Aus 1) können wir auch die Richtung des vektoriellen Produktes zweier nicht-parallelen Vektoren ablesen: Bezeichnet nämlich $\mathbf{e}$ einen auf die Ebene des Vektorenpaares $\mathbf{a}, \mathbf{b}$ orthogonalen Einheitsvektor, so läßt sich der Vektor $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ jedenfalls in der Form

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = a\mathbf{a} + b\mathbf{b} + c\mathbf{e}$$

schreiben. Das Vektorentrippel $\mathbf{a}, \mathbf{b}, \mathbf{e}$ geht aber durch eine Drehung mit π um $\mathbf{e}$ in das Trippel $-\mathbf{a}, -\mathbf{b}, \mathbf{e}$ über, also ist

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = (-\mathbf{a}) \times (-\mathbf{b}) = a(-\mathbf{a}) + b(-\mathbf{b}) + c\mathbf{e} = -a\mathbf{a} - b\mathbf{b} + c\mathbf{e}.$$

Hieraus folgt, daß $a = b = 0$ und $\mathbf{a} \times \mathbf{b}$ von der Richtung $\mathbf{e}$, d. h. orthogonal auf der Ebene $(\mathbf{a}, \mathbf{b})$ ist.

Nach diesen Vorbereitungen führen wir den hier beginnenden eigentlichen *Beweis* — wie in der Einleitung schon erwähnt — auf zwei verschiedenen Wegen: *a)* elementar-geometrisch und *b)* durch Lösung einer Funktionalgleichung.

a) Beim skalaren Produkt setzen wir $(\mathbf{a}, \mathbf{b}) \sphericalangle < \frac{\pi}{2}$ voraus. Der Fall $(\mathbf{a}, \mathbf{b}) \sphericalangle > \frac{\pi}{2}$ läßt sich ähnlich erledigen. Offenbar muß wegen 1)

$$\mathbf{e} \cdot \mathbf{e} = k_1$$

eine von $\mathbf{e}$ unabhängige (skalare) Konstante sein, und daraus folgt schon mittels 3):

$$\mathbf{a} \cdot \mathbf{b} = k_1 |\mathbf{a}| \cdot |\mathbf{b}|, \quad \text{falls } \mathbf{a} \uparrow\uparrow \mathbf{b}. \tag{3}$$

Nunmehr finden wir $\mathbf{a} \cdot \mathbf{b}$ im allgemeinen,[4] indem wir (s. Figur 1) $\mathbf{a} = \mathbf{a}' + \mathbf{a}''$ mit $\mathbf{a}' \uparrow\uparrow \mathbf{b}$ und $\mathbf{a}'' \perp \mathbf{b}$ setzen. Dann folgt aus 2), (1) und (3)

$$\mathbf{a} \cdot \mathbf{b} = (\mathbf{a}' + \mathbf{a}'') \cdot \mathbf{b} = \mathbf{a}' \cdot \mathbf{b} + \mathbf{a}'' \cdot \mathbf{b} = k_1 |\mathbf{a}'| |\mathbf{b}| = k_1 |\mathbf{a}| |\mathbf{b}| \cos(\mathbf{a}, \mathbf{b}),$$

womit der Satz über das skalare Produkt bewiesen ist.

[4] Falls man beiderseitige Distributivität voraussetzt, folgt die Kommutativität

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

unmittelbar aus (1), da für beliebige Einheitsvektoren $\mathbf{e}_1, \mathbf{e}_2$

$$(\mathbf{e}_1 + \mathbf{e}_2) \perp (\mathbf{e}_1 - \mathbf{e}_2),$$

und deshalb

$$0 = (\mathbf{e}_1 + \mathbf{e}_2) \cdot (\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{e}_1 \cdot \mathbf{e}_1 - \mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_2 \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_1 - \mathbf{e}_1 \cdot \mathbf{e}_2$$

ist.

Andererseits muß wegen 1) für jedes $\mathbf{e}_1 \perp \mathbf{e}_2$ die Skalare k_2 in

$$\mathbf{e}_1 \times \mathbf{e}_2 = k_2 \mathbf{e}_3,$$

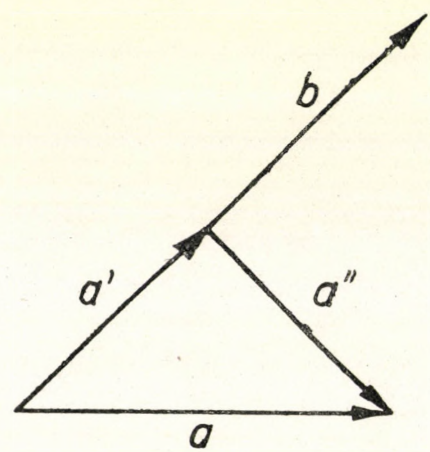
wo $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ ein Rechtssystem bilden, ebenfalls eine Konstante sein. Hieraus folgt wegen 3):

$$\mathbf{a} \times \mathbf{b} = k_2 |\mathbf{a}| |\mathbf{b}| \mathbf{e}, \tag{4}$$

falls $\mathbf{e} \perp \mathbf{a}$, $\mathbf{e} \perp \mathbf{b}$ und $\mathbf{a}, \mathbf{b}, \mathbf{e}$ ein Rechtssystem bilden. Endlich ist[5] für $\mathbf{a} \nparallel \mathbf{b}$ (Figur 1) wegen 2), (2) und (4)

$$\mathbf{a} \times \mathbf{b} = (\mathbf{a}' + \mathbf{a}'') \times \mathbf{b} = \mathbf{a}' \times \mathbf{b} + \mathbf{a}'' \times \mathbf{b} = \mathbf{a}'' \times \mathbf{b} = k_2 |\mathbf{a}''| |\mathbf{b}| \mathbf{e} =$$
$$= k_2 |\mathbf{a}| |\mathbf{b}| \sin(\mathbf{a}, \mathbf{b}) \mathbf{e},$$

wo $\mathbf{e} \perp \mathbf{a}$, $\mathbf{e} \perp \mathbf{b}$ und $\mathbf{a}, \mathbf{b}, \mathbf{e}$ ein Rechtssystem ist, womit auch die zweite Hälfte des Satzes 1 bewiesen ist.



Figur 1

b) Jetzt seien (Figur 2) $\mathbf{e}_{\varphi+\psi}$ bzw. $\mathbf{e}_{\varphi-\psi}$ Einheitsvektoren, die mit dem Einheitsvektor $\mathbf{e}$ die Winkel $\varphi+\psi$ bzw. $\varphi-\psi$ einschließen und mit $\mathbf{e}$ komplanar sind.

Es ist nach der Additionsregel zweier Vektoren

$$\mathbf{e}_{\varphi+\psi} + \mathbf{e}_{\varphi-\psi} = 2\mathbf{e}_{\varphi} \cos \psi.$$

Nun folgt aus 2)

$$(\mathbf{e}_{\varphi+\psi} + \mathbf{e}_{\varphi-\psi}) \cdot \mathbf{e} = \mathbf{e}_{\varphi+\psi} \cdot \mathbf{e} + \mathbf{e}_{\varphi-\psi} \cdot \mathbf{e},$$
$$(\mathbf{e}_{\varphi+\psi} + \mathbf{e}_{\varphi-\psi}) \times \mathbf{e} = \mathbf{e}_{\varphi+\psi} \times \mathbf{e} + \mathbf{e}_{\varphi-\psi} \times \mathbf{e}.$$

Führen wir laut 1) die Bezeichnung $\mathbf{e}_{\varphi} \cdot \mathbf{e} = f(\varphi)$, bzw. $\mathbf{e}_{\varphi} \times \mathbf{e} = f(\varphi) \mathbf{e}'$ ein, wo $\mathbf{e}' \perp \mathbf{e}_{\varphi}$, $\mathbf{e}' \perp \mathbf{e}$ ist und $\mathbf{e}_{\varphi}, \mathbf{e}, \mathbf{e}'$ ein Rechtssystem bilden, so erhalten

[5] Falls man beiderseitige Distributivität voraussetzt, so folgt das Alternieren

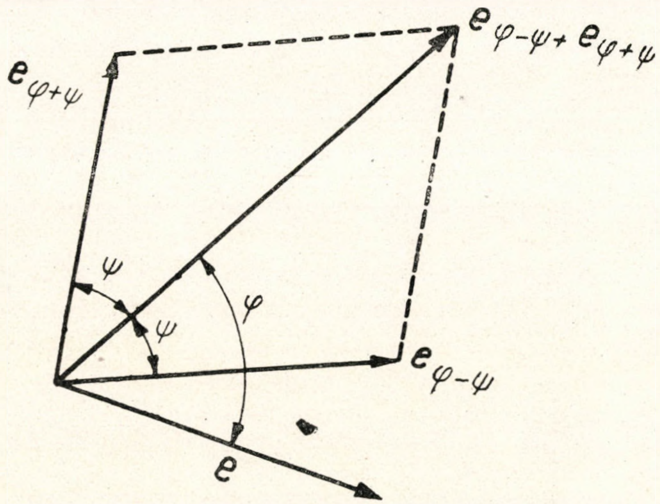
$$\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$$

unmittelbar aus (2):

$$0 = (\mathbf{a} + \mathbf{b}) \times (\mathbf{a} + \mathbf{b}) = \mathbf{a} \times \mathbf{a} + \mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{a} + \mathbf{b} \times \mathbf{b} = \mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{a}.$$

wir unter Berücksichtigung von 3) und vom Parallelismus der Vektoren $\mathbf{e}_{\varphi+\psi} \times \mathbf{e}$ und $\mathbf{e}_{\varphi-\psi} \times \mathbf{e}$ in beiden Fällen die Funktionalgleichung

$$2f(\varphi)\cos\psi = f(\varphi+\psi) + f(\varphi-\psi).$$



Figur 2

Wie man gleich sieht, sind $\cos\varphi$, $\sin\varphi$ und mit ihnen auch $c_1\cos\varphi + {}$ $+ c_2 \sin\varphi$ Lösungen dieser Gleichung. Wir wollen zeigen, daß dies die allgemeinste Lösung ist.

Setzen wir in dieser Funktionalgleichung zuerst $\varphi = \frac{\pi}{2}$, $\psi = \frac{\pi}{2} - t$, dann $\varphi = \frac{\pi}{2} - t$, $\psi = \frac{\pi}{2}$, und endlich $\varphi = 0$, $\psi = t$, so erhalten wir der Reihe nach

$$f(\pi - t) + f(t) = 2f\left(\frac{\pi}{2}\right)\sin t,\ f(\pi - t) + f(-t) = 0,\ f(t) + f(-t) = 2f(0)\cos t.$$

Addieren wir die dritte Gleichung zur Differenz der ersten und zweiten, so wird:

$$f(t) = f(0)\cos t + f\left(\frac{\pi}{2}\right)\sin t.$$

Somit ist

$$f(t) = c_1 \cos t + c_2 \sin t$$

die allgemeine Lösung der Funktionalgleichung.

Für uns folgt in dem Falle der skalaren Operation aus (1):

$$c_2 = f\left(\frac{\pi}{2}\right) = 0,\ c_1 = k_1,\ \text{also}\ f(\varphi) = k_1\cos\varphi,$$

$$\mathbf{a}\cdot\mathbf{b} = k_1\,|\mathbf{a}|\cdot|\mathbf{b}|\cos(\mathbf{a},\mathbf{b});$$

in dem Falle der vektoriellen Operation aus (2):

$$c_1 = f(0) = 0, \quad c_2 = k_2, \quad \text{also} \quad f(\varphi) = k_2 \sin \varphi,$$

$$\mathbf{a} \times \mathbf{b} = k_2 |\mathbf{a}| |\mathbf{b}| \sin (\mathbf{a}, \mathbf{b}) \mathbf{e},$$

wo $\mathbf{e} \perp \mathbf{a}, \mathbf{e} \perp \mathbf{b}$ und $\mathbf{a}, \mathbf{b}, \mathbf{e}$ ein Rechtssystem bilden, w. z. b. w.[6]

III. Jetzt wenden wir uns der Multiplikation von Quaternionen zu. Das Produkt der Quaternionen $a + \mathbf{a}$ und $b + \mathbf{b}$ bezeichnen wir einfach mit $(a + \mathbf{a})(b + \mathbf{b})$ (während $\mathbf{a} \cdot \mathbf{b}$ bzw. $\mathbf{a} \times \mathbf{b}$ hier die *übliche* skalare bzw. vektorielle Vektorenmultiplikation bezeichnet). Es gilt der

SATZ 2. *Voraussetzungen:*

1) *Es gibt keine ausgezeichnete Richtung (im dreidimensionalen Vektorraum). Darunter verstehen wir, daß die Quaternionenmultiplikation drehungsautomorph ist, d. h. bei einer Drehung des Raumes bleibt der skalare bzw. vektorielle Teil des Quaternionenproduktes ungeändert bzw. ebenso gedreht.*

2) *Die Quaternionenmultiplikation ist beiderseitig distributiv.*

3) *Skalaren lassen sich aus der Quaternionenmultiplikation herausheben.*

4) *Die Quaternionenmultiplikation ist assoziativ:*

$$[(a + \mathbf{a})(b + \mathbf{b})](c + \mathbf{c}) = (a + \mathbf{a})[(b + \mathbf{b})(c + \mathbf{c})].$$

5) *Das Quaternionenprodukt zweier Einheitsvektoren ist eine Einheitsquaternion, d. h. eine Quaternion $a + \mathbf{a}$, für welche $a^2 + |\mathbf{a}|^2 = 1$ ist.*

Behauptungen:

A) *Falls die Voraussetzungen* 1), 2), 3) *erfüllt sind, so ist*

$$(a + \mathbf{a})(b + \mathbf{b}) = (ab + p\mathbf{a} \cdot \mathbf{b}) + (a\mathbf{b} + b\mathbf{a} + q\mathbf{a} \times \mathbf{b});$$

B) *falls die Voraussetzungen* 1), 2), 3), 4) *erfüllt sind, so ist*

$$(a + \mathbf{a})(b + \mathbf{b}) = (ab - p^2\mathbf{a} \cdot \mathbf{b}) + (a\mathbf{b} + b\mathbf{a} + q\mathbf{a} \times \mathbf{b});$$

C) *falls die Voraussetzungen* 1), 2), 3), 5) *erfüllt sind, so ist*

$$(a + \mathbf{a})(b + \mathbf{b}) = (ab \pm \mathbf{a} \cdot \mathbf{b}) + (a\mathbf{b} + b\mathbf{a} \pm \mathbf{a} \times \mathbf{b}),$$

wo die beiden Vorzeichens-Alternativen voneinander unabhängig sind;

[6] Die oben angegebene Lösungsmethode für die Funktionalgleichung $f(x+y) + f(x-y) = 2f(x) \cos y$ gibt eine einfachere Lösung eines in der Arbeit: ST. KACZMARZ, Sur l'équation fonctionnelle $f(x) + f(x+y) = \varphi(y) f\left(x + \frac{y}{2}\right)$, *Fund. Math.*, 6 (1924), S. 122—129 behandelten Teilproblems. Es sei hier bemerkt, daß auch für die komplexen Zahlen die Multiplikationsregel unter Voraussetzung der Distributivität und der Identifizierung der Horizontalvektoren mit Skalaren, aber natürlich ohne irgendwelche Invarianz-Voraussetzung, durch Lösung von Funktionalgleichungen des Typs

$$f\left(\frac{\varphi + \psi}{2}\right) = \frac{f(\varphi) + f(\psi)}{2}$$

ähnlich bewiesen werden kann. Für den bezüglichen *elementargeometrischen* Beweis s. T. SZELE, A komplex számok bevezetése vektoralgebrai alapon, *Matematikai Lapok*, 1 (1950), S. 349—362. L. FUCHS und T. SZELE, Introduction of complex numbers as vectors of the plane, *Amer. Math. Monthly*, 59 (1952), S. 628—631.

D) *falls endlich die Voraussetzungen* 1), 2), 3), 4), 5) *alle erfüllt sind, so ist*

$$(a+\mathbf{a})(b+\mathbf{b})=(ab-\mathbf{a}\cdot\mathbf{b})+(a\mathbf{b}+b\mathbf{a}\pm\mathbf{a}\times\mathbf{b})$$

und dies ist schon die übliche Regel der Quaternionenmultiplikation (da das Vorzeichen von $\mathbf{a}\times\mathbf{b}$ *nur die Orientierung des Systems entscheidet).*

Bemerkung: Es genügt die Assoziativität nur für Vektoren, d. h. für $a=b=c=0$ *vorauszusetzen, man könnte sich sogar auch auf ein speziales Einheitsvektorentrippel beschränken.*

Der *Beweis* beruht auf dem Satze 1. Wegen 2) wird

$$(a+\mathbf{a})(b+\mathbf{b})=ab+a\mathbf{b}+b\mathbf{a}+\mathbf{ab}$$

(wo ab, $a\mathbf{b}$, $b\mathbf{a}$ die gewöhnliche Multiplikation mit Skalaren, $\mathbf{ab}$ dagegen eine Quaternionenmultiplikation von Vektoren bedeutet).

$\mathbf{ab}$ ist eine Quaternion, deren skalarer Teil $s(\mathbf{ab})$ und vektorieller Teil $\mathbf{v}(\mathbf{ab})$, als Verknüpfungen betrachtet, auch einzeln distributiv sind. Es gilt nämlich

$$s[\mathbf{a}(\mathbf{b}+\mathbf{c})]=s(\mathbf{ab}+\mathbf{ac})=s(\mathbf{ab})+s(\mathbf{ac}),$$
$$\mathbf{v}[\mathbf{a}(\mathbf{b}+\mathbf{c})]=\mathbf{v}(\mathbf{ab}+\mathbf{ac})=\mathbf{v}(\mathbf{ab})+\mathbf{v}(\mathbf{ac}),$$

da bei der Addition von Quaternionen die skalaren und die vektoriellen Teile separat addiert werden.

Hieraus ergibt sich aber durch Anwendung des Satzes 1 unmittelbar:

$$s(\mathbf{ab})=p\,\mathbf{a}\cdot\mathbf{b},\quad \mathbf{v}(\mathbf{ab})=q\,\mathbf{a}\times\mathbf{b},$$

also

$$\mathbf{ab}=p\,\mathbf{a}\cdot\mathbf{b}+q\,\mathbf{a}\times\mathbf{b},$$
$$(a+\mathbf{a})(b+\mathbf{b})=(ab+p\,\mathbf{a}\cdot\mathbf{b})+(a\mathbf{b}+b\mathbf{a}+q\,\mathbf{a}\times\mathbf{b}),$$

was — wie man sich durch Einsetzen überzeugt — die Voraussetzungen 1), 2), 3) tatsächlich erfüllt, vomit A) bewiesen ist.

Es ist

$$(\mathbf{ab})\mathbf{c}=(p\,\mathbf{a}\cdot\mathbf{b}+q\,\mathbf{a}\times\mathbf{b})\mathbf{c}=p(\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}+q\,[p(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}+q\,(\mathbf{a}\times\mathbf{b})\times\mathbf{c}]=$$
$$=pq(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}+p(\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}+q^2(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$$

und ebenso

$$\mathbf{a}(\mathbf{bc})=pq\,\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})+p(\mathbf{b}\cdot\mathbf{c})\,\mathbf{a}+q^2\mathbf{a}\times(\mathbf{b}\times\mathbf{c}).$$

Es ist aber aus dem Vektorenkalkül bekannt, daß

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$$

und

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}-\mathbf{a}\times(\mathbf{b}\times\mathbf{c})=(\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}-(\mathbf{b}\cdot\mathbf{c})\,\mathbf{a}$$

ist. Die Quaternionenmultiplikation von Vektoren ist also dann und nur dann assoziativ (4)), falls

$$0=(\mathbf{ab})\mathbf{c}-\mathbf{a}(\mathbf{bc})=(p+q^2)\;[(\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}-(\mathbf{b}\cdot\mathbf{c})\,\mathbf{a}].$$

Da aber $(\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}-(\mathbf{b}\cdot\mathbf{c})\,\mathbf{a}\neq 0$ ist, trifft dies im allgemeinen dann und nur dann zu, wenn $p+q^2=0$ ist, was eben B) beweist, da

$$(a+\mathbf{a})(b+\mathbf{b})=(ab-q^2\mathbf{a}\cdot\mathbf{b})+(a\mathbf{b}+b\mathbf{a}+q\,\mathbf{a}\times\mathbf{b})$$

die Voraussetzungen 1), 2), 3), 4) auch wirklich erfüllt.

Gelten aber 1), 2), 3), 5) und wird 4) nicht vorausgesetzt, so muß das Quaternionenprodukt zweier Einheitsvektoren eine Einheitsquaternion sein, d. h. es muß

$$|\mathbf{e}_1\mathbf{e}_2|=|p\mathbf{e}_1\cdot\mathbf{e}_2+q\,\mathbf{e}_1\times\mathbf{e}_2|=|p\cos\varphi+q\,\mathbf{e}_3\sin\varphi|=1$$

bestehen, wo φ der Winkel der Vektoren $\mathbf{e}_1,\mathbf{e}_2$ ist und $\mathbf{e}_3$ orthogonal zur Ebene $(\mathbf{e}_1,\mathbf{e}_2)$ ist. Also muß

$$p^2\cos^2\varphi+q^2\sin^2\varphi=1$$

sein, was durch Einsetzen von $\varphi=0$, bzw. $\varphi=\frac{\pi}{2}$

$$p^2=1,\ q^2=1,$$
$$p=\pm 1,\ q=\pm 1,$$
$$(a+\mathbf{a})(b+\mathbf{b})=(ab\pm\mathbf{a}\cdot\mathbf{b})+(a\mathbf{b}+b\mathbf{a}\pm\mathbf{a}\times\mathbf{b})$$

ergibt, womit auch C) erledigt ist, denn diese Multiplikationsformel erfüllt tatsächlich die Voraussetzungen 1), 2), 3), 5).

Sind endlich alle Voraussetzungen erfüllt, so gilt

$$p=-q^2 \quad \textit{und} \quad p^2=q^2=1,$$

also

$$p=-q^2=-1,$$

d. h.

$$(a+\mathbf{a})(b+\mathbf{b})=(ab-\mathbf{a}\cdot\mathbf{b})+(a\mathbf{b}+b\mathbf{a}\pm\mathbf{a}\times\mathbf{b}),$$

womit der Beweis des Satzes 2 vollendet ist.

(Eingegangen am 7. September 1952.)

ОБ УМНОЖЕНИИ ВЕКТОРОВ И КВАТЕРНИОНОВ

Я. АЦЕЛ (Дебрецен)

(Р е з ю м е)

Автор доказывает, что из системы простых условий — среди которых наиболее существенное значение имеет закон дистрибутивности — следуют известные правила скалярного и векторного умножения векторов и умножения кватернионов. Из последнего выводятся теоремы, аналогичные относящейся сюда теоремы Ф р о б е н и у с а.

ON COMPOSED POISSON DISTRIBUTIONS, IV

(Remarks on the theory of differential processes)

By

ANDRÁS PRÉKOPA (Budapest)

(*Presented by* A. RÉNYI)

Introduction

In the present paper we shall use the following notations. Let ξ_t be a stochastic process. For the difference $\xi_{t_2}-\xi_{t_1}$ let us introduce the notation

$$\xi_J=\xi_{t_2}-\xi_{t_1}$$

where J means the finite time interval (t_1, t_2). Let further $W_\lambda(J)$ denote the following probability

$$W_\lambda(J)=Pr(\xi_J=\lambda)$$

and let $F(x, J)$ denote the probability distribution of the variable ξ_J. For the characteristic function of the distribution $F(x, J)$ we shall write $f(u, J)$:

$$f(u, J)=\int_{-\infty}^{\infty} e^{iux}dF(x, J).$$

The process on which our discussions are based will be considered in a finite, closed time interval I in which the process satisfies the following conditions. (By J we shall henceforth denote a subinterval of I.)

A) If $J_1, J_2, \ldots, J_n$ denote a subdivision of the interval I so that $I=J_1+J_2+\cdots+J_n$, the corresponding variables $\xi_{J_1}, \xi_{J_2}, \ldots, \xi_{J_n}$ are independent.

B) The variables ξ_J can only assume the values of a countable set of real numbers $\lambda_0=0, \lambda_1, \lambda_2, \ldots$ which set is independent of the special selection of J.

It follows from condition *A)* that this set must be closed under addition, since for every λ_k and λ_l it must contain a number λ_m such that $\lambda_k+\lambda_l=\lambda_m$, that is to say, it must form a semigroup with respect to addition.

C) $1-W_0(J)=1-Pr(\xi_J=0)$ is a continuous interval function, that is,

$$1-W_0(J)\to 0,$$

if J contracts to a fixed point.

It follows from condition *C)* that $1-W_0(J)$ is also uniformly continuous in I.[1] It follows likewise from condition *C)* that $1-f(u,J)$ is also uniformly continuous, notably, in a manner independent of u, for

$$|1-f(u,J)| \leqq 2(1-W_0(J)),$$

and, further, that the process ξ_t is weakly continuous, since

$$Pr(|\xi_J|>\varepsilon) \leqq 1-Pr(\xi_J=0) \to 0.$$

In consequence of condition *A)* and *C)*, ξ_I possesses an infinitely divisible distribution,[2] and therefore $\log f(u,I)$ can be represented in the canonical form:

$$\log f(u,I) = i\gamma(I)u - \frac{\sigma^2(I)}{2}u^2 +$$

$$+\int_{-\infty}^{0}\left(e^{iux}-1-\frac{iux}{1+x^2}\right)dM(x,I)+\int_{0}^{\infty}\left(e^{iux}-1-\frac{iux}{1+x^2}\right)dN(x,I) \tag{1}$$

where $\gamma(I)$ and $\sigma(I)$ are constants, $M(x,I)$ and $N(x,I)$ are non-decreasing functions in the intervals $(-\infty,0)$ and $(0,\infty)$, respectively; $M(-\infty)=$ $=N(\infty)=0$ and

$$\int_{-1}^{0}x^2dM(x,I)+\int_{0}^{1}x^2dN(x,I)<\infty.$$

In our case the general form (1) reduces to a simpler expression in which the concrete meaning of the functions $M(x,I)$ and $N(x,I)$ can be indicated more specifically.

§ 1. The meaning of $M(x,I)$ and $N(x,I)$

THEOREM. *If conditions A), B), C) are fulfilled, then the general form* (1) *can be written in the following manner:*

$$\log f(u,I) = \sum_{k=1}^{\infty} C_{\lambda_k}(I)(e^{i\lambda_k u}-1), \tag{2}$$

where

$$C_{\lambda_k}(I) = \int_I W_{\lambda_k}(J), \qquad \lambda_k \neq 0 \tag{3}$$

[1] Let $\varphi(J)$ be an interval function continuous in the finite closed interval I; in this case $\varphi(J)$ is also uniformly continuous in I. For, in the contrary case it would be possible to find an ε_0 for which no δ exists, that is, if $\delta_n\to 0$, by selecting a suitable sequence J_n, we would obtain that $|\varphi(J_n)|>\varepsilon_0$ contrary to $|J_n|\leqq\delta_n\to 0$. Let d denote a point of condensation of the centres of the intervals J_n, then J_n will have a subsequence J_{k_n} whose centres converge to d. As the sequence of intervals J_{k_n} contracts to a single point, to d, it follows that $\varphi(J_{k_n})\to 0$, which is a contradiction.

[2] See [1], pp. 161—163.

and

$$\sum_{k=1}^{\infty} C_{\lambda_k}(I) = \int_I (1 - W_0(J)) < \infty. \tag{4}$$

In (3) and (4) the integrals in the interval I are taken in the sense of BURKILL.

PROOF. We first show that $\sigma(I) = 0$. In fact, the function $f(u, I)$ in formula (1) is the product of two characteristic functions:

$$f(u, I) = f_1(u, I) f_2(u, I)$$

where

$$f_1(u, I) = e^{-\frac{\sigma^2(I)}{2} u^2}$$

is the characteristic function of the normal distribution

$$F_1(x, I) = \frac{1}{\sqrt{2\pi}\sigma(I)} \int_{-\infty}^{x} e^{-\frac{y^2}{2\sigma^2(I)}} dy.$$

As

$$|F_1(x+h, I) - F_1(x, I)| \leqq \frac{|h|}{\sqrt{2\pi}\sigma(I)} \quad \text{and} \quad F(x, I) = \int_{-\infty}^{\infty} F_1(x-y, I)\, dF_2(y, I),$$

it follows that

$$|F(x+h, I) - F(x, I)| \leqq \frac{|h|}{\sqrt{2\pi}\sigma(I)}.$$

This is, however, a contradiction, because we have supposed that $F(x, I)$ is a step function. (For the proof, see [3], pp. 94—95.)

$f(u, I)$ is an almost periodic function of the variable u,

$$f(u, I) = \sum_{k=0}^{\infty} W_{\lambda_k}(I) e^{i\lambda_k u}.$$

Together with $f(u, I)$, $\log f(u, I)$ is also an almost periodic function. For the proof we have only to show that $|f(u, I)| \geqq \delta > 0$, because $\log z$ is continuous in the region $0 < \delta \leqq |z| \leqq 1$, and it is well known that any continuous function of an almost periodic function is itself almost periodic. In fact, if the decomposition $I = J_1 + J_2 + \cdots + J_n$ is carried out, where $|J_k|$ (the length of J_k) is sufficiently small to permit that in accordance with condition *C)* $|f(u, J_k)| \geqq \geqq \eta > 0$, it will follow that

$$|f(u, I)| = \prod_{k=1}^{n} |f(u, J_k)| \geqq \eta^n = \delta > 0.$$

Let us now multiply both sides of (1) by $\frac{1}{2T}$ and integrate in the interval $(-T, T)$; then by FUBINI's theorem, first integrating with respect to

u, we obtain

$$-\frac{1}{2T}\int_{-T}^{T}\log f(u, I)\,du=\int_{-\infty}^{0}\left(1-\frac{\sin Tx}{Tx}\right)dM(x, I)+\int_{0}^{\infty}\left(1-\frac{\sin Tx}{Tx}\right)dN(x, I).$$

If $T\to\infty$, then, as $\log f(u, I)$ is almost periodic, we shall obtain on the left side a finite limit and therefore, taking into account also that

$$1-\frac{\sin Tx}{Tx}\geqq 0,$$

we may apply FATOU's theorem to obtain

$$\int_{\infty}^{0}dM(x, I)+\int_{0}^{\infty}dN(x, I)<\infty. \tag{5}$$

Accordingly, $M(x, I)$ and $N(x, I)$ are of bounded variation and have a finite limit at the point $x=0$. In view of (5), formula (1) may be brought to the following form:

$$\log f(u, I)=i\gamma'(I)u+\int_{-\infty}^{0}(e^{iux}-1)\,dM(x, I)+\int_{0}^{\infty}(e^{iux}-1)\,dN(x, I), \tag{6}$$

where

$$\gamma'(I)=\gamma(I)-\int_{-\infty}^{0}\frac{x}{1+x^2}\,dM(x, I)-\int_{0}^{\infty}\frac{x}{1+x^2}\,dN(x, I).$$

Here we have $\gamma'(I)=0$, since the left side of (6) and the two last members of its right side are bounded functions of u.

$M(x, I)$ and $N(x, I)$ are step functions; this fact follows at once if we take into account that on the left side of (6) there stands an almost periodic function. For if we decompose the functions

$$M(x, I)=M_1(x, I)+M_2(x, I),\qquad N(x, I)=N_1(x, I)+N_2(x, I),$$

so that on the right sides the first member is a step function, while the second member is a continuous, non-decreasing function, then from (6), by a suitable rearrangement, we obtain

$$\varphi(u, I)=\int_{-\infty}^{0}e^{iux}dM(x, I)+\int_{0}^{\infty}e^{iux}dN(x, I),$$

if the almost periodic part is denoted by $\varphi(u, I)$. The right hand side can however be almost periodic only if $M(x, I)\equiv$ const. and $N(x, I)\equiv$ const. Indeed, this is readily seen if we form the expression[3] $M(\varphi(u, I)e^{-i\lambda u})$ and prove that this vanishes for all λ's, owing to the continuity of $M_2(x, I)$ and $N_2(x, I)$.

[3] $M(\varphi(u, I)\,e^{-i\lambda u})=\lim\limits_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\varphi(u, I)e^{-i\lambda u}du.$

We shall now show that $\log f(u, I)$ possesses the same exponents as $f(u, I)$, and further, that the other assertions of the theorem are also valid. For the proof we need a lemma.

LEMMA. If a stochastic process satisfies conditions *A)*, *B)* and *C)*, then

$$\log f(u, I)=\int_I (f(u,J)-1), \tag{7}$$

and the sequence in the definition of the Burkill integral will uniformly converge with respect to u to $\log f(u, I)$, or, more precisely,

$$\left|\log f(u, I)-\sum_{k=1}^{n}(f(u,J_k)-1)\right| \leqq K \max(1-W_0(J_k)) \to 0 \tag{8}$$

$$\text{if } \max|J_k| \to 0.$$

PROOF. We know that $|f(u,J)|^2$, together with $f(u, J)$, is an almost periodic characteristic function, and that, owing to $|f(u, J)|^2 \geqq \delta^2 > 0$, the function $\log|f(u, J)|^2$ is likewise almost periodic. Let us denote in the distribution defined by $|f(u, J)|^2$ the probability of the value 0 by $W_0^*(J)$, then

$$W_0^*(J)=\sum_{k=0}^{\infty} W_{\lambda_k}^2(J)=M(|f(u,J)|^2). \tag{9}$$

From (9) we have

$$W_0^*(J)-1 \leqq W_0^2(J)+1-W_0(J)-1=W_0(J)(W_0(J)-1)$$

and therefore, if $|J|$ is so small that $W_0(J)>\frac{1}{2}$, then

$$1-W_0(J)<2(1-W_0^*(J)). \tag{10}$$

We shall now show that $1-W_0(J)$ is of bounded variation.

$$\begin{aligned}\sum_{k=1}^{n}(1-W_0(J_k)) &\leqq 2\sum_{k=1}^{n}(1-W_0^*(J_k))=2M\left(\sum_{k=1}^{n}(1-|f(u,J_k)|^2)\right) \leqq \\ &\leqq -2M\left(\sum_{k=1}^{n}\log|f(u,J_k)|^2\right)=-2M(\log|f(u,I)|^2)=K,\end{aligned} \tag{11}$$

where use has been made of the inequality $1-x \leqq -\log x$ $(0<x\leqq 1)$. Thus (8) can be established in the following manner:

$$\begin{aligned}\left|\log f(u, I)-\sum_{k=1}^{n}(f(u,J_k)-1)\right| &\leqq \sum_{k=1}^{n}|\log f(u,J_k)-(f(u,J_k)-1)| \leqq \\ &\leqq \sum_{k=1}^{n}\sum_{l=2}^{\infty}\frac{1}{l}|f(u,J_k)-1|^l \leqq \frac{1}{2}\sum_{k=1}^{n}\frac{|f(u,J_k)-1|}{1-|f(u,J_k)-1|} \leqq \\ &\leqq \sum_{k=1}^{n}|f(u,J_k)-1|^2 \leqq \sum_{k=1}^{n}(1-W_0(J_k))^2 \leqq K\max(1-W_0(J_k)) \to 0.\end{aligned}$$

As the sequence figuring in the expression (8) converges uniformly to $\log f(u, I)$, it follows that

$$M\left(\sum_{k=1}^{n}(f(u,J_k)-1)e^{-i\lambda u}\right) \to M(\log f(u,I)e^{-i\lambda u}),$$

that is, if the Fourier coefficients of $\log f(u, I)$ are denoted by $C_{\lambda_k}(I)$, then

$$(12)\qquad \begin{cases} C_{\lambda_k}(I) = \int\limits_I W_{\lambda_k}(J), & \lambda_k \neq 0 \\ C_0(I) = \int\limits_I (W_0(J) - 1), \end{cases}$$

and, in view of (6), we have

$$(12a)\qquad -C_0(I) = \sum_{k=1}^{\infty} C_{\lambda_k}(I) < \infty.$$

Consequently, the theorem is proved.

It may happen that in the formula (2) $C_{\lambda_k}(I)$ belonging to a certain λ_k is equal to 0. A sufficient condition for this is that in every point of the interval we have

$$\frac{W_{\lambda_k}(J)}{|J|} \to 0$$

during J contracts to the point in question. For, in this case, we have

$$(13)\qquad \int\limits_I W_{\lambda_k}(J) = 0.$$

The proof is very simple. As the interval function $\frac{W_{\lambda_k}(J)}{|J|}$ is continuous in the closed interval I, it is therefore also uniformly continuous. In other words, if $\varepsilon > 0$ is arbitrary and $\max |J_l| < \delta$, it follows that $\frac{W_{\lambda_k}(J_l)}{|J_l|} < \varepsilon$ and therefore $\sum\limits_{l=1}^{n} W_{\lambda_k}(J_l) < \varepsilon |I|$, whence (13) follows.

§ 2. A direct proof

For the theorem we have just proved, we can also give a direct proof,[4] if we make use of a theorem known in the theory of almost periodic functions.

Thus in this case we do not make use of formula (1), but we require the lemma. For by the lemma, as is seen from (12), the Fourier coefficients of the function $\log f(u, I) - C_0(I)$ are $C_{\lambda_k}(I) \geqq 0$ $(\lambda_k \neq 0)$. It is known that if all Fourier coefficients of an almost periodic function are non-negative, then their sum is convergent:[5]

$$(14)\qquad \sum_{k=1}^{\infty} C_{\lambda_k}(I) < \infty.$$

[4] A direct proof of formula (2) is discussed in the first section of [3] for the case when the totality of the numbers λ_k is identical with the set of non-negative integers.

[5] See [1], p. 62.

It follows from (14) that the Fourier series of $f(u,I)-C_0(I)$ converges uniformly, and therefore

$$\log f(u,I)-C_0(I)=\sum_{k=1}^{\infty} C_{\lambda_k}(I)e^{i\lambda_k u}. \tag{15}$$

If 0 is substituted for u, we obtain

$$-C_0(I)=\sum_{k=1}^{\infty} C_{\lambda_k}(I), \tag{16}$$

therefore

$$\log f(u,I)=\sum_{k=1}^{\infty} C_{\lambda_k}(I)\,(e^{i\lambda_k u}-1)$$

which exactly proves (2), whilst (12), (14) and (16) are proving (3) and (4).

The proof of (14) can simply be carried out in case the totality of the numbers λ_k is identical with the totality of integers. In fact, in this case $\log f(u,I)$ is a periodic function with period 2π. As $f(u,I)=\overline{f(-u,I)}$, the real part of $\log f(u,I)$ is an even function, whilst its imaginary part is an odd function. Consequently,

$$\frac{1}{2\pi}\int_0^{2\pi}(\log|f(u,I)|-C_0(I))\cos ku\,du=\begin{cases}\dfrac{C_k(I)+C_{-k}(I)}{2} & \text{if } k\neq 0,\\ 0 & \text{if } k=0.\end{cases}$$

The integrand is a continuous even function and its Fourier coefficients are non-negative, whence [6]

$$\sum_{k\neq 0}(C_k(I)+C_{-k}(I))<\infty.$$

§ 3. The explicit form of $W_k(I)$

Let us consider the case in which the values λ_k are integers. Then (2) can be written in the following manner:

$$\log f(u,I)=\sum_{k\neq 0} C_k(I)\,(e^{i\lambda_k u}-1). \tag{17}$$

Let us now express the probabilities $W_k(I)$ with the aid of the interval functions $C_k(I)$. It follows from (17) that the variable ξ_I which will now be denoted by $\xi(I)$, can be represented as follows:

$$\xi(I)=\sum_{k=-\infty}^{\infty} k\,\xi_k\ I), \quad \text{or} \quad \xi(I)=\sum_{k=1}^{\infty} k\,(\xi_k(I)-\xi_{-k}(I)), \tag{18}$$

[6] A special case of PALEY's theorem which can easily be proved is the following: if the Fourier coefficients of an even and continuous function are non-negative, then their sum will converge. As a matter of fact,

$$\frac{1}{2}S_n(0)\leq\frac{n+1}{2n+1}S_n(0)\leq\frac{1}{2n+1}\sum_{k=n}^{2n}S_k(0)\leq\sigma_{2n}(0)$$

where σ_n denotes the n-th Fejér mean.

where the variables $\xi_k(I)$ are independent and have a Poisson distribution with a mean value $C_k(I)$. Let us now consider the second expression of (18). Here the differences of variables of Poisson distribution are standing. In general, if ξ and η are two independent variables of Poisson distribution with mean values λ and μ, the difference

$$\zeta=\xi-\eta$$

can assume arbitrary integral values and

$$Pr(\zeta=n)=\lambda^n e^{-(\lambda+\mu)}\sum_{k=0}^{\infty}\frac{(\lambda\mu)^k}{k!(k+n)!} \qquad (n=0,1,2,\ldots)$$

and

$$Pr(\zeta=-n)=\mu^n e^{-(\lambda+\mu)}\sum_{k=0}^{\infty}\frac{(\lambda\mu)^k}{k!(k+n)!} \qquad (n=0,1,2,\ldots).$$

These probabilities can be expressed with the aid of the Bessel function of n-th order

$$J_n(x)=\left(\frac{x}{2}\right)^n\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{x}{2}\right)^{2k}}{k!(k+n)!}$$

in the following manner:

$$\left.\begin{aligned}Pr(\zeta=n)&=\left(\frac{1}{i}\sqrt{\frac{\lambda}{\mu}}\right)^n e^{-(\lambda+\mu)}J_n(2i\sqrt{\lambda\mu})\\ Pr(\zeta=-n)&=\left(\frac{1}{i}\sqrt{\frac{\mu}{\lambda}}\right)^n e^{-(\lambda+\mu)}J_n(2i\sqrt{\lambda\mu}),\end{aligned}\right\} \qquad (n=0,1,2,\ldots)$$

or. with a uniform method of writing:

$$Pr(\zeta=n)=\left(\frac{1}{i}\right)^{|n|}\left(\frac{\lambda}{\mu}\right)^{\frac{n}{2}} e^{-(\lambda+\mu)}J_{|n|}(2i\sqrt{\lambda\mu}). \tag{19}$$

By making use of (19) and on the basis of (18) we obtain

$$\begin{gathered}W_k(I)=e^{-\lambda(I)}\sum_{\Sigma n r_n=k}\prod\left(\frac{1}{i}\right)^{|r_n|}\left(\frac{C_n(I)}{C_{-n}(I)}\right)^{\frac{r_n}{2}}J_{|r_n|}(2i\sqrt{C_n(I)C_{-n}(I)}),\\ k=0,\pm 1,\pm 2,\ldots \quad (n=1,2,\ldots;\ r_n \text{ is an integer})\end{gathered} \tag{20}$$

where the product is extended over all finite systems of values r_n for which $\Sigma n r_n=k$, whilst the summation relates to all such systems, and

$$\lambda(I)=\sum_{n\neq 0}C_n(I).$$

Finally, I express my sincere thanks to ALFRED RÉNYI for his valuable remarks.

INSTITUTION FOR APPLIED MATHEMATICS
OF THE HUNGARIAN ACADEMY OF SCIENCES

(Received 25 October 1952)

Literature

[1] J. Favard, *Leçons sur les fonctions presques-périodiques* (Paris, 1939).
[2] P. Lévy, *Processus stochastiques et mouvement Brownien* (Paris, 1948).
[3] A. Rényi, On composed Poisson distributions, II, *Acta Math. Acad. Sci. Hung.*, **2** (1951), pp. 83–98.

ОБОБЩЕННЫЕ РАСПРЕДЕЛЕНИЯ ТИПА ПУАССОНА, IV

(Замечания к теории дифференциальных процессов)

А. ПРЕКОПА (Будапешт)

(Резюме)

Пусть ξ_t стохастический процесс расмотренный на конечном отрезке времени I. Обозначим разность $\xi_{t_2}-\xi_{t_1}$ с ξ_J, где J — отрезок (t_1, t_2) $(J\subset I)$, а $W_\lambda(J)$ — вероятность

$$W_\lambda(J) = Pr(\xi_J = \lambda).$$

Если имеем следующие условия: *A)* ξ_J есть процесс с независимыми приращениями; *B)* ξ_J принимают только значения фиксированного счетного множества $\lambda_0=0, \lambda_1, \lambda_2, \ldots$, с положительной вероятностю; *C)* $1-W_0(J)\to 0$, когда J стягивается в фиксированную точку; то справедлива следующая

Теорема. Логарифм характеристической функции величины ξ_J может быть представлен в виде:

$$\log f(u, I) = \sum_{k=1}^{\infty} C_{\lambda_k}(I)\,(e^{i\lambda_k u}-1),$$

где

$$C_{\lambda_k}(I) = \int_I W_{\lambda_k}(J) \qquad (\lambda_k \neq 0)$$

и

$$\sum_{k=1}^{\infty} C_{\lambda_k}(I) = \int_I \bigl(1-W_0(J)\bigr) < \infty.$$

Интегралы принимаются в смысле Бэркилла.

В § 1 теорема доказывается при помощи формулы Леви [1] и одной леммы, а в § 2 доказательство той же теоремы опирается на теорию почти периодических функций без использования формулы Леви. В § 3 дается явная форма функции отрезок $W_{\lambda_k}(J)$.

NEUE HERLEITUNG DER HYPERBOLISCHEN TRIGONOMETRIE DURCH VERWENDUNG DER GRENZKUGEL

Von

PAUL SZÁSZ (Budapest)

(*Vorgelegt von* G. Hajós)

Es sei $p(a)$ der Grenzkreisbogen von der Höhe a und bezeichne $\Pi(a)$ wie üblich den Parallelwinkel, der dem Abstand a angehört (Fig. 1). Schon J. Bolyai[1] hat bewiesen, daß der Grenzkreisbogen

$$p(a)\ \mathrm{tg}\, \Pi(a) = k \tag{1}$$

von a unabhängig, d. h. eine Weltkonstante ist. Die Höhe dieses Grenzkreisbogens ist also der Abstand x vom Parallelwinkel $\Pi(x) = 45^0$ (Fig. 2).

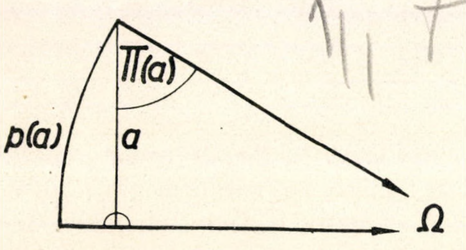


Fig. 1

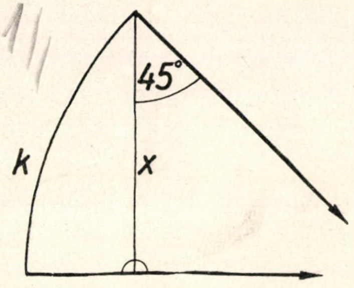


Fig. 2

V. F. Kagan[2] hat für den Satz unter (1) mit alleiniger Hilfe der Grenzkugel einen einfacheren Beweis gegeben und auf Grund dieses Satzes die klassische Formel

$$\mathrm{tg}\, \frac{\Pi(a)}{2} = e^{-\frac{a}{k}} \tag{2}$$

in bemerkenswerter Weise von neuem hergeleitet, und dadurch die hyperbolische Trigonometrie gewonnen.

[1] J. Bolyai, *Appendix. Scientiam spatii absolute veram exhibens, etc.* (Marosvásárhely, 1832), besonders § 30. Deutsch bei P. Stäckel, *Wolfgang und Johann Bolyai. Geometrische Untersuchungen,* Bd. II (Leipzig und Berlin, 1913), S. 198.

[2] V. F. Kagan, *N. I. Lobacsevszkij, Geometriai vizsgálatok a párhuzamosok elméletének köréből* (ungarisch), übersetzt von György Bizám, zum Drucke befördert von Ferenc Kárteszi, (Budapest, 1951), besonders S. 130—141.

In vorliegender Note wird für die Gleichungen (1) und (2) wieder eine andere Herleitung angegeben. Der von KAGAN verwendete Satz, laut welchem

$$\lim_{a=0} \frac{p(a)}{a} = 1$$

ausfällt, ist eine unmittelbare Folge unserer Darstellung.[3] Zunächst wird durch Verwendung einer anderen Konfiguration von KAGAN[4] der Satz (1) einfacher bewiesen, aus der mitgewonnenen hyperbolischen Winkeltrigonometrie des rechtwinkligen Dreiecks die Gleichung

$$\operatorname{tg} \frac{\Pi(a)}{2} = e^{-\frac{a}{l}}$$

hergeleitet und endlich gezeigt, daß hier l gleich der Bogenlänge von k ist.

§ 1. Die hyperbolische Winkeltrigonometrie des rechtwinkligen Dreiecks

Es sei ABC ein rechtwinkliges Dreieck ($\sphericalangle C = 90^0$) mit den Bestimmungsstücken

$$\overline{BC} = a, \ \overline{CA} = b, \ \overline{AB} = c, \ \sphericalangle A = \lambda, \ \sphericalangle B = \mu.$$

Da auf der Grenzkugel die euklidische ebene Geometrie gilt, wenn die Grenzkreise Geraden genannt werden, so folgt aus der klassischen Konfiguration von J. BOLYAI[5]

$$\sin \lambda = \frac{p(a)}{p(c)}. \tag{3}$$

Wir stellen nun die Konfiguration von V. F. KAGAN[6] her. Auf die Ebene des Dreiecks errichten wir in A eine Senkrechte mit dem einen unendlich fernen Punkte Ω und legen durch C die Grenzkugel mit dem Mittelpunkt Ω, von dem die Achse $A\Omega$ in D, $B\Omega$ in E geschnitten werde (Fig. 3). Infolge der Konstruktion steht die Ebene $AC\Omega$ auf ABC senkrecht, die Gerade BC steht deshalb auf $AC\Omega$ ebenfalls senkrecht und es ist somit $\sphericalangle BC\Omega = 90^0$, also $\sphericalangle CB\Omega = \Pi(a)$. Wird noch von C auf $B\Omega$ das Lot $\overline{CF} = d$ gefällt, so ist daher im rechtwinkligen Dreieck BCF (mit der Hypotenuse a) im Sinne des Satzes (3)

$$p(d) = p(a) \sin \Pi(a). \tag{4}$$

Da weiter im Grenzkugeldreieck CDE offenbar $\sphericalangle D = \lambda$ ausfällt, ferner $\overset{\frown}{CD} = p(b)$ und $\overset{\frown}{CE} = p(d)$ ist, so besteht nach der euklidischen Geometrie

[3] Schon F. SCHUR hat darauf hingewiesen, daß der Beweis dieses Satzes auch bei J. BOLYAI [1] fehlt, vgl. P. STÄCKEL, loc. cit. [1], Bd. II, S. 272.

[4] V. F. KAGAN, loc. cit. [2], S. 132.

[5] J. BOLYAI, loc. cit. [1], § 25.

[6] Siehe [4].

der Grenzkugel mit Rücksicht auf (4) die Formel

(5) $$\operatorname{tg} \lambda = \frac{p(a)}{p(b)} \sin \Pi(a).$$

Aus (3) und (5) ergibt sich durch Elimination von λ

(6) $$p(c)^2 = p(a)^2 + \frac{p(b)^2}{\sin^2 \Pi(a)}.$$

Ebenso gilt

$$p(c)^2 = p(b)^2 + \frac{p(a)^2}{\sin^2 \Pi(b)},$$

und aus diesen beiden Gleichungen folgt

$$p(a) \operatorname{tg} \Pi(a) = p(b) \operatorname{tg} \Pi(b).$$

Damit ist der Satz (1) bewiesen, da doch a und b voneinander unabhängig sind.

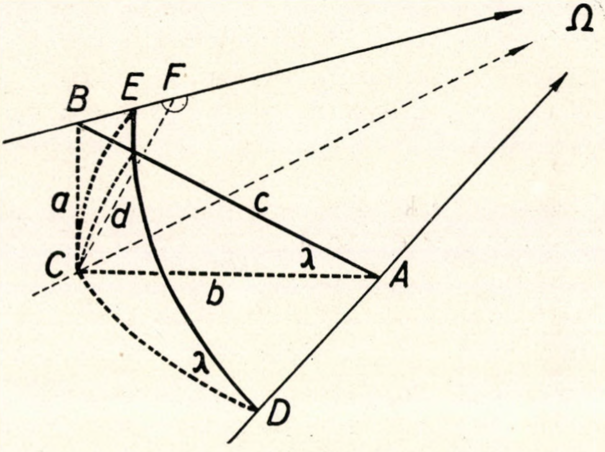


Fig. 3

Auf Grund dieses Satzes nimmt (3) bzw. (5) die Gestalt

(I) $$\sin \lambda = \frac{\operatorname{ctg} \Pi(a)}{\operatorname{ctg} \Pi(c)}$$

bzw.

(II) $$\operatorname{tg} \lambda = \cos \Pi(a) \operatorname{tg} \Pi(b)$$

an, und (6) kann sogleich auf die Form

(III) $$\sin \Pi(c) = \sin \Pi(a) \sin \Pi(b)$$

gebracht werden.

Mit Rücksicht auf (III) ergibt sich aus (I) und (II)

(IV) $$\cos \lambda = \frac{\cos \Pi(b)}{\cos \Pi(c)},$$

(II) und die ähnliche Gleichung für den Winkel μ ergeben

(V) $$\operatorname{tg} \lambda \operatorname{tg} \mu = \sin \Pi(c),$$

aus (IV) und der Gleichung (I) ähnlichen Formel folgt endlich

$$\text{(VI)} \qquad \frac{\sin \mu}{\cos \lambda} = \sin \Pi(a).$$

(I)—(VI) sind die Lobatschefskijschen Grundformeln[7] der hyperbolischen Winkeltrigonometrie des rechtwinkligen Dreiecks.

§ 2. Bestimmung des Parallelwinkels

Im Besitze der Winkeltrigonometrie des rechtwinkligen Dreiecks kann die klassische Gleichung zwischen den Parallelwinkeln $\alpha_1, \alpha_2, \alpha$, die den Abständen $a_1, a_2, a = a_1 + a_2$ entsprechen, in folgender Weise gewonnen werden. (Den mit lateinischen Buchstaben bezeichneten Abständen angehörige Parallelwinkel seien immer mit den entsprechenden griechischen Buchstaben bezeichnet.)

Es seien auf eine Gerade in derselben Richtung die Abstände $\overline{AB} = a_1$, $\overline{BC} = a_2$ abgetragen, wobei also $\overline{AC} = a$ ist, und in B sei auf diese Gerade eine Senkrechte errichtet mit dem einen unendlich fernen Punkte Ω (Fig. 4). Wegen der Bedeutung von α_1 und α_2 ist dann $\sphericalangle \Omega AB = \alpha_1$, $\sphericalangle \Omega CB = \alpha_2$. Von A werde auf $C\Omega$ das Lot $\overline{AD} = t$ gefällt, dem also der Parallelwinkel $\tau = \sphericalangle \Omega AD$ entspricht, und es sei $\sphericalangle CAD = \lambda$. Infolge $\tau = \alpha_1 - \lambda$ ist

$$(7) \qquad \sin \tau = \sin \alpha_1 \cos \lambda - \cos \alpha_1 \sin \lambda.$$

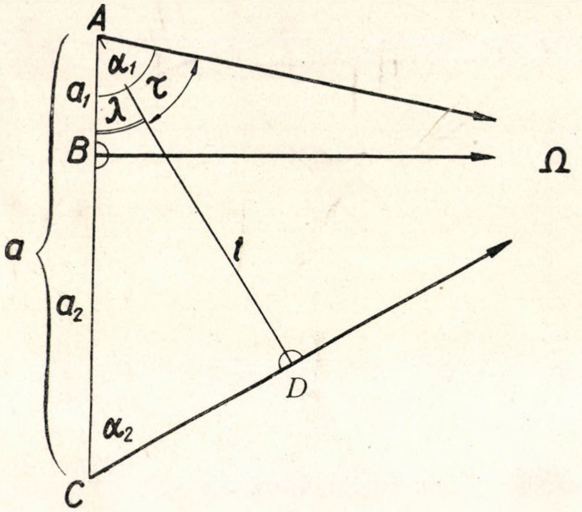


Fig. 4

Es besteht aber im rechtwinkligen Dreieck ACD nach (IV)

$$\cos \lambda = \frac{\cos \tau}{\cos \alpha}$$

[7] Vgl. F. ENGEL, *Nikolaj Iwanowitsch Lobatschefskij etc.* (Leipzig, 1898), S. 20, Formel (14).

und im Sinne von (VI)

$$\sin\lambda = \cos\alpha_2 \sin\tau,$$

(7) nimmt also nach Division durch $\sin\tau$ die Gestalt

$$1 = \frac{\sin\alpha_1}{\cos\alpha}\operatorname{ctg}\tau - \cos\alpha_1\cos\alpha_2$$

an. Im Sinne von (I) ist ferner

$$\operatorname{ctg}\tau = \operatorname{ctg}\alpha \sin\alpha_2,$$

somit lautet die gewonnene Gleichung

$$1 = \sin\alpha_1 \sin\alpha_2 \frac{1}{\sin\alpha} - \cos\alpha_1\cos\alpha_2,$$

woher durch Einführung der Halbwinkel sich

$$\operatorname{ctg}\frac{\alpha}{2} + \operatorname{tg}\frac{\alpha}{2} = \operatorname{ctg}\frac{\alpha_1}{2}\operatorname{ctg}\frac{\alpha_2}{2} + \operatorname{tg}\frac{\alpha_1}{2}\operatorname{tg}\frac{\alpha_2}{2}$$

ergibt. Die Wurzeln dieser Gleichung zweiten Grades für $\operatorname{ctg}\frac{\alpha}{2}$ sind offenbar $\operatorname{tg}\frac{\alpha_1}{2}\operatorname{tg}\frac{\alpha_2}{2}$ und $\operatorname{ctg}\frac{\alpha_1}{2}\operatorname{ctg}\frac{\alpha_2}{2}$. Die erste kommt aber nicht in Betracht, da doch $\operatorname{ctg}\frac{\alpha}{2} > 1$, $\operatorname{tg}\frac{\alpha_1}{2}\operatorname{tg}\frac{\alpha_2}{2} < 1$ ausfällt. Also wird

$$\operatorname{ctg}\frac{\alpha}{2} = \operatorname{ctg}\frac{\alpha_1}{2}\operatorname{ctg}\frac{\alpha_2}{2}$$

die gewünschte Gleichung.[8] Mit wachsendem a durchläuft α beständig abnehmend das Intervall $0^0 < \alpha < 90^0$, ist also eine stetige Funktion von a, und $\operatorname{ctg}\frac{\alpha}{2}$ ist deshalb auch stetig. Aus dieser Funktionalgleichung folgt daher[9]

$$\operatorname{ctg}\frac{\alpha}{2} = e^{\frac{a}{l}} \tag{8}$$

wobei l eine gewisse Strecke bedeutet.

Zur Bestimmung dieser Strecke l nehmen wir die Rektifikation des Grenzkreisbogens p vor.

(8) zieht

$$\operatorname{ctg}\alpha = \operatorname{sh}\frac{a}{l}$$

nach sich, also ist im Sinne von (1)

$$\frac{p(a)}{k} = \operatorname{sh}\frac{a}{l},$$

[8] Vgl. F. Engel, loc. cit.[7], S. 20. Formel (11).

[9] A. L. Cauchy, *Analyse algébrique* (Paris, 1821), Ch. V § I, problème II, Oeuvres IIe série, t. III, S. 100—102.

und so strebt

$$\frac{\frac{p(a)}{k}}{\frac{a}{l}} \to 1 \tag{9}$$

für $p(a) \to 0$, da in diesem Falle a fortiori $a \to 0$. Der Grenzkreisbogen p werde nun in n Teile $p_1, p_2, \ldots, p_n$ geteilt und $s_1, s_2, \ldots, s_n$ seien die entsprechenden Sehnen. Es ist bekanntlich

$$\min \frac{\frac{p_i}{k}}{\frac{s_i}{l}} \leqq \frac{\frac{p_1}{k} + \frac{p_2}{k} + \cdots + \frac{p_n}{k}}{\frac{s_1}{l} + \frac{s_2}{l} + \cdots + \frac{s_n}{l}} \leqq \max \frac{\frac{p_i}{k}}{\frac{s_i}{l}}.$$

Strebt der größte Bogen gegen 0, so streben hier auf Grund von (9) sowohl die untere, als auch die obere Schranke gegen 1, also

$$\frac{s_1 + s_2 + \cdots + s_n}{l} \to \frac{p_1 + p_2 + \cdots + p_n}{k} = \frac{p}{k}$$

d. h.

$$s_1 + s_2 + \cdots + s_n \to \frac{p}{k}\, l.$$

Der Grenzkreisbogen p ist demnach rektifizierbar und hat die Länge $\frac{p}{k} l$. Für den Fall $p = k$ ergibt sich insbesondere

$$l = \text{der Bogenlänge von } k. \tag{10}$$

Wird die Länge von k ebenfalls mit k bezeichnet, so nimmt (8) mit Rücksicht auf (10) die Gestalt

$$\operatorname{ctg} \frac{\alpha}{2} = e^{\frac{a}{k}} \tag{8*}$$

an. Damit ist auch der Beweis des Satzes (2) fertig.

Da infolge (8*)

$$\sin \alpha = \frac{1}{\operatorname{ch} \frac{a}{k}}, \quad \cos \alpha = \operatorname{th} \frac{a}{k}, \quad \operatorname{tg} \alpha = \frac{1}{\operatorname{sh} \frac{a}{k}}$$

ist, gehen die Formeln (I)—(VI) des § 1 in die bekannten Grundformeln der hyperbolischen Trigonometrie über.

(Eingegangen am 12. Januar 1953)

НОВОЕ ОБОСНОВАНИЕ ГИПЕРБОЛИЧЕСКОЙ ТРИГОНОМЕТРИИ С ПОМОЩЬЮ ОРИСФЕРЫ

П. САС (Будапешт)

(Резюме)

Применяя одну конфигурацию В. Ф. Кагана, автор снова показывает что продукт некоторой дуги $p(a)$ высоты a орицикла на тангенс угла параллельности $\Pi(a)$

$$p(a)\operatorname{tg}\Pi(a)=k$$

не зависит от a, то есть является постоянной дугой орицикла. На основе полученной таким образом уголной тригонометрии прямоуголного треуголника выводится уравнение

$$\operatorname{tg}\frac{\Pi(a)}{2}=e^{-\frac{a}{l}}.$$

Наконец ректифицированием дуги орицикла доказывается, что l равно здесь длине дуги k, так что если обозначать эту длину также через k, то получается

$$\operatorname{tg}\frac{\Pi(a)}{2}=e^{-\frac{a}{k}}.$$

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 du 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, französicher, 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.

24 forint

INDEX

Rédei, L., Vollidealringe im weiteren Sinn, I.
Л. Рэдэи, Полные идеальные кольца в общем смысле, I 243
Sz.-Nagy, Gy., Wertverteilung bei Polynomen mit lauter reellen Nullstellen und Koeffizienten.
Д. С.-Надь, Распределение значений полиномов с действительными коэффициентами и действительными корнями 269
Rényi, A., and *Turán, P.*, On the zeros of polynomials.
А. Реньи и П. Туран, О корнях многочленов 275
Sz.-Nagy, B., A moment problem for self-adjoint operators.
Б. С.-Надь, Проблема моментов для самосопряженных операторов 285
Ganea, T., Remark on R-equivalent spaces.
Т. Ганеа, Замечание об R-эквивалентных пространствах 295
Freud, G., Restglied eines Tauberschen Satzes, II.
Г. Фрайд, Об остаточном члене некоторой теоремы типа Таубера, II . . . 299
Aczél, J., Bemerkungen über die Multiplikation von Vektoren und Quaternionen.
Я. Ацел, Об умножении векторов и кватернионов 309
Prékopa, A., On composed Poisson distributions, IV.
А. Прекопа, Обобщенные распределения типа Пуассона, IV 317
Szász, P., Neue Herleitung der hyperbolischen Trigonometrie durch Verwendung der Grenzkugel.
П. Сас, Новое обоснование гиперболической тригонометрии с помощью орисферы 327

Technikai szerkesztő: Fuchs László

A kiadásért felelős: Mestyán János. Műszaki felelős: Tóth Ferenc

A kézirat beérkezett: 1953. II. 6. — Példányszám: 750. — Terjedelem: 8 és fél (A/5) ív, 6 ábra.

Csongrádmegyei Nyomdaipari Vállalat, Szeged 532 Felelős vezető: Vincze György