ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO MATHEMATICA

TOMUS VI.

REDIGIT

Á. CSÁSZÁR

ADIUVANTIBUS

L. FUCHS, G. HAJÓS, F. KÁRTESZI, A. KÓSA, J. MOLNÁR, R. PÉTER, A. RÉNYI, J. SURÁNYI, P. SZÁSZ, P. TURÁN

1963

ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO BIOLOGICA
incepit anno MCMLVII.
SECTIO CHIMICA
incepit anno MCMLIX.
SECTIO GEOLOGICA
incepit anno MCMLVII.
SECTIO HISTORICA
incepit anno MCMLVII.
SECTIO IURIDICA
incepit anno MCMLIX.
SECTIO MATHEMATICA
incepit anno MCMLVIII.
SECTIO PHILOLOGICA
incepit anno MCMLVII.
SECTIO PHILOSOPHICA
incepit anno MCMLXII.
SECTIO PHYSICA
incepit anno MCMLXIII.

A kiadásért felelős:
az Eötvös Loránd Tudományegyetem rektora
Eredeti kiadásról készült változatlan utánnyomás
Minden jog fenntartva
Külföldi terjesztés:
KULTURA KÖNYV- ÉS HÍRLAP
KÜLKERESKEDELMI VÁLLALAT
BUDAPEST 62,
P. O. B. 149
This book is a reproduction of the original, published
in Budapest
All rights reserved
General Distributors:
KULTURA Hungarian Trading Company
for Books and Newspapers
BUDAPEST 62, P. O. B. 149,
Hungary
Printed in Hungary, 1969

68-2723-Felsőoktatási Jegyzetellátó Vállalat

REMARKS ON PROBABLE INFERENCE

By

J. ACZÉL

Mathematical Institute of the Kossuth Lajos University, Debrecen

(Received February 25, 1963)

1. R. T. Cox has given in 1946 ([9][1]) a system of conditions to be considered as natural for probable inferences of propositions i. e. conditional probabilities of events (he reproduced this argument in 1961 [10], and we will refer in what follows to this book). All functions figuring in his considerations were supposed to be differentiable.[2] There exists a great number of similar systems of conditions for unconditional probabilities (e. g. [1], [4]–[7], [14], [16]–[23]), which make no use of differentiability conditions. (Though not dealing directly with probabilities still the five papers [8], [12], [13], [24], [25] two of them by S. GOŁĄB (KRAKÓW) treating essentially similar problems under differentiability suppositions might be mentioned too.) In [2] (cf. also [3] 7.1.4) the author of the present paper showed how his considerations in [1] can be carried over in a more exact form to conditional probabilities.

In the present paper we wish to give a link with the investigations of R. T. Cox and show how his results can be obtained without differentiability conditions, resp. how they can be reduced to those of [1] and [2]. – The main interest lies in the rather intricate functional equation (10) of R. T. Cox ([10] (4.4)) in which a function of one variable is to be determined and which we reduce to a pair of functional equations for a function of two variables.

2. By denoting the probability of the event **A** under the condition **V** by

$$p(\mathbf{A}/\mathbf{V})$$

and supposing that $p(\mathbf{A}/\mathbf{V})$ takes every value between $0=p(\mathbf{0}/\mathbf{V})$ ($\mathbf{0}$ being the impossible event) and $e=p(\mathbf{V}/\mathbf{V})$, supposing further

$$p(\mathbf{A}\cdot\mathbf{B}/\mathbf{V}) = G[p(\mathbf{A}/\mathbf{B}\cdot\mathbf{V}),\ p(\mathbf{B}/\mathbf{V})] \tag{1}$$

for the probability of the conjoined events[3] (cf. [2] (13), [10] (3.1)), the associativity equation

$$(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$$

in Boolean algebra implies

$$G[x, G(y, z)]=G[G(x, y), z] \quad (x, y, z, G \in (0, e]). \tag{2}$$

[1] Numbers in brackets refer to the bibliography at the end of this paper.

[2] On p. 37 of [10] even the equation (7.1) $\eta(xy) = \eta(x)\,\eta(y)$ supposed only for entire x, y is solved by differentiation (!). Though there not – at the other places his differentiability suppositions prove to be essential in the considerations.

[3] In [2] (13) we allowed the function G to depend upon the condition **V**. Here we follow [10] and take the same G for all **V**-s.

In [9] and [10] G was supposed to be twice differentiable and the equation (2) was solved by differentiations. If we suppose G only *continuous and strictly monotonic* in x and y, and *such that for* $e=p(\mathbf{V}/\mathbf{V})$

$$G(e, x) = x, \tag{3}$$

holds, — suppositions suggested by (1) (cf. [2]) — then Theorem 1 of [2] implies already

$$G(x, y) = h^{-1}[h(x)+h(y)] \tag{4}$$

or with $g(x) = \exp h(x) > 0$:

$$G(x, y) = g^{-1}[g(x)g(y)] \tag{5}$$

i. e. by (1)

$$g[p(\mathbf{A}\cdot\mathbf{B}/\mathbf{V})] = g[p(\mathbf{A}/\mathbf{B}\cdot\mathbf{V})]\, g[p(\mathbf{B}/\mathbf{V})] \tag{6}$$

in accordance with the result of [10] p. 16. — Taking there $g(x)\equiv Cx$ and then $g(x)\equiv x$ is not quite motivated, e. g. one could take with the same right $h(x)\equiv x$ in (4). On the other hand *one can redefine the conditional probabilities* by

$$\bar{p}(\mathbf{A}/\mathbf{V}) = g[p(\mathbf{A}/\mathbf{V})]^a > 0 \tag{7}$$

($a\neq 0$ an *arbitrary* constant) so that — by (6) — *the product theorem*

$$\bar{p}(\mathbf{A}\cdot\mathbf{B}/\mathbf{V}) = \bar{p}(\mathbf{A}/\mathbf{B}\cdot\mathbf{V})\bar{p}(\mathbf{B}/\mathbf{V}) \tag{8}$$

holds — in accordance with [10] (3.7). — (3) and (5) imply $g(e)=1$ and so by (7)

$$\bar{p}(\mathbf{V}/\mathbf{V}) = 1.$$

We have excluded here $0=p(\mathbf{0}/\mathbf{V})$ because $G(0, y)$ is no more strictly monotonic. Nevertheless, by the formula $\mathbf{0}\cdot\mathbf{B}=\mathbf{0}$ of Boolean algebra, (8) remains valid with $p(\mathbf{0}/\mathbf{V})=0$.

3. R. T. Cox supposes further

$$\bar{p}(\tilde{}\mathbf{A}/\mathbf{V}) = f[\bar{p}(\mathbf{A}/\mathbf{V})] \tag{9}$$

for the probability of the contradictory event and derives by a rather intricate argument the functional equation

$$zf\left(\frac{f(y)}{z}\right) = yf\left(\frac{f(z)}{y}\right) \tag{10}$$

for this function besides the obvious

$$f[f(x)] = x \tag{11}$$

(following from $\tilde{}\tilde{}\mathbf{A} = \mathbf{A}$). In [9] and [10] again f was supposed to be twice differentiable and the functional equation (10) was solved by repeated diffe-

rentiations.[4] Here we will suppose f to be *continuous and strictly decreasing* in $(0,1]$, suppositions natural by (9). When examining (11), also $f(x)\in(0,1)$ for $x\in(0,1)$ will be supposed.

It is interesting that the functional equation (10) for a function of one variable can be solved by reducing it to a pair of functional equations for a function of two variables. For we define the function

$$F(x, y) = xf\left(\frac{y}{x}\right) \qquad \left[x, y, \frac{y}{x}\in(0,1]\right] \tag{12}$$

of two variables which by what precedes is *continuous, decreasing in* y and also *increasing in* x, because for $x_1<x_2$

$$\frac{y}{x_1}>\frac{y}{x_2},\ f\left(\frac{y}{x_1}\right)<f\left(\frac{y}{x_2}\right),\ x_1f\left(\frac{y}{x_1}\right)<x_2f\left(\frac{y}{x_2}\right)$$

[as by (7) and (9) $f>0$].

The function (12) is obviously homogeneous with exponent 1:

$$F(xz, yz) = F(x, y)z \tag{13}$$

and on the other hand (10) implies

$$\frac{1}{y}F[x, F(y, z)] = \frac{x}{y}f\left(\frac{F(y, z)}{x}\right) = \frac{x}{y}f\left[\frac{y}{x}f\left(\frac{z}{y}\right)\right] = \frac{z}{y}f\left[\frac{y}{z}f\left(\frac{x}{y}\right)\right] =$$
$$= \frac{z}{y}f\left(\frac{F(y,x)}{z}\right) = \frac{1}{y}F[z, F(y, x)]$$

i. e.

$$F[x, F(y, z)] = F[z, F(y, x)]. \tag{14}$$

In [1] (Theorem 2; cf. also [14]) we have solved the pair of functional equations

$$F(xz, yz) = F(x,y)z,$$
$$F[x, F(y, z)] = F[F(x, y),z],$$

and found

$$F(x, y) = (x^k+y^k)^{1/k}$$

as the general solution continous and strictly monotonic in both variables. The solution of (13)–(14) is not much more difficult either (cf. also [25]).

First we realize (cf. [15]) that (14) implies

$$F[F(x, y), F(u, v)] = F(v, F[u, F(x, y)]) = F(v, F[y, F(x, u)]) =$$
$$= F[F(x, u), F(y, v)].$$

[4] In [10] $f(x) = (bx^k+b^2)^{1/k}$ was found as general solution of (10).
A solution of (10) was neglected there: $f(x) = kx^b$ $(b^2+b = 1)$, which would follow from the case $c = 1$ of the differential equation $f' = b\left(\frac{f}{x}\right)^c$ of [10] (p. 21.).

The general solution continuous and strictly monotonic in both variables of this functional equation

$$F[F(x, y), F(u, v)] = F[F(x, u), F(y, v)]$$

of bisymmetry is (cf. e. g. [3] 6.4.3)

$$F(x, y) = h^{-1}(ah(x)+bh(y)+c) \tag{15}$$

where h is an arbitrary continuous and strictly monotonic function and $a \neq 0$, $b \neq 0$ are arbitrary constants. If we put in (15), we see immediately that (14) is satisfied if and only if $a=b^2$, i. e.

$$F(x, y) = h^{-1}(b^2h(x)+bh(y)+c) \qquad (b \neq 0) \tag{16}$$

is the general continuous strictly monotonic solution of (14) (cf. [15]; though in our consideration in all $F(x, y)$ -s $y \leqq x$ had to be supposed, nevertheless (16) remains true in this case too, as M. Hosszú has shown in a paper forthcoming in these *Annales*). F being decreasing in y, necessarily $b<0$. F is by all means increasing in x.

Thus (13) goes over into

$$h^{-1}(b^2h(xz)+bh(yz)+c) = h^{-1}(b^2h(x)+bh(y)+c)z$$

or if we denote

$$u = h(x), \quad v = h(y), \quad h_z(u) = h(h^{-1}(u)z), \tag{17}$$

then into

$$h_z(b^2u+bv+c) = b^2h_z(u)+bh_z(v)+c.$$

The general, say continuous solution of this equation (cf. [11] and [3] 2.2.6, 2.2.7) is

$$h_z(u) = A_zu+B_z, \tag{18}$$

the „constants" A_z, B_z depending also from z, f course. (There is a certain relation between A_z, B_z and b, c but we do not care for it at the moment.) Taking this dependence and (17) into consideration, (18) gives the functional equation

$$h(xz) = A(z)h(x)+B(z). \tag{19}$$

We can solve (19) easily (cf. [3] 3.1.3) by putting in first $x=1$

$$h(z) = A(z)h(1)+B(z)$$

and subtracting this from (19). Then for

$$g(x) = h(x)-h(1) \tag{20}$$

we get

$$g(xz) = A(z)g(x)+g(z). \tag{21}$$

By the symmetry of the left-hand-side, the right-hand-side is also symmetric in x, z:

$$A(z)g(x)+g(z) = A(x)g(z)+g(x). \tag{22}$$

We distinguish two cases:

1. $A(z)\equiv 1$. Then (21) goes over in to the CAUCHY-type equation

$$g(xz) = g(x)+g(z),$$

the general, say continuous solution of which in $(0,1]$ is (see e. g. [3] 2.1.2, 2.1.4)

$$g(x) = a \log x$$

and so by (20)

$$h(x) = a \log x+h(1).$$

(16) gives for this h the solution

$$F(x, y) = kx^{b^2}y^b \qquad (k = e^{[(b^2+b-1)h(1)+c]/a} \neq 0)$$

which satisfies (14) always but — this is a consequence of the above mentioned relation between $A(z)$, $B(z)$ and b, c — fulfills (13) only if

$$b^2+b=1.$$

Of the two solutions for b only

$$b = \frac{-1-\sqrt{5}}{2}$$

is negative. The obtained result

$$F(x, y) = kx^{b^2}y^b \qquad (b^2+b=1;\ k\neq 0) \tag{23}$$

settles the case 1. entirely and we can go over to the second case.

2. There exists a z_0 such that $A(z_0)\neq 1$. Then we put this z_0 into (22) and get

$$g(x) = a\,(A(x)-1) \qquad \left(a = \frac{g(z_0)}{A(z_0)-1}\right) \tag{24}$$

where $a\neq 0$ because of the strict monotony of h and g (cf. (20)). Resubstituting this into (21) we see that

$$A(xz) = A(x)A(z)$$

and this CAUCHY-type equation has in $(0,1]$

$$A(x) = x^k \qquad (k\neq 0)$$

as general, say continuous, non-constant solution (cf. again [3] 2.1.2, 2.1.4). (20) and (24) yield

$$h(x) = ax^k+h(1)-a$$

and by (16)

$$F(x, y)=(b^2x^k+by^k+C)^{1/k} \qquad \left(C = \frac{(b^2+b-1)\,[h(1)-a]+c}{a}\right)$$

which again always satisfies (14), of course, but (13) only if

$$C = 0.$$

(Again a consequence of the above mentioned relation between $A(z)$, $B(z)$ and b, c.) This gives the solution

$$F(x, y) = (b^2x^k + by^k)^{1/k} \qquad (b \neq 0, \quad k \neq 0) \tag{25}$$

and ends the process of solution. We have the

THEOREM 1. *The most general solutions, continuous and strictly monotonic in x and y, of the pair*

$$F(xz, yz) = F(x, y)z \tag{13}$$

$$F[x, F(y, z)] = F[z, F(y, x)] \tag{14}$$

of functional equations in $(0,1]$ *are*

$$F(x, y) = kx^{b^2}y^b \qquad (b^2 + b = 1, \quad k \neq 0) \tag{23}$$

and

$$F(x, y) = (b^2x^k + by^k)^{1/k} \quad (b \neq 0, \quad k \neq 0). \tag{25}$$

If we take (12) into consideration, f being decreasing, we get[4] the

THEOREM 2. *The most general continuous positive decreasing solutions of the functional equation*

$$zf\left(\frac{f(y)}{z}\right) = yf\left(\frac{f(z)}{y}\right) \tag{10}$$

in $(0,1]$ *are*

$$f(z) = kz^{(-1-\sqrt{5})/2} \qquad (k > 0) \tag{26}$$

and

$$f(z) = (b^2 + bz^k)^{1/k} \qquad (b < 0, \ k \neq 0). \tag{27}$$

Upto now we did not take (11) into consideration. One sees immediately that (26) does not satisfy this equation and (27) only if

$$b^2 + b^3 = 0, \quad b^2 = 1 \qquad \text{i. e. } b = -1.$$

So we arrived to the

COROLLARY 1. *The most general continuous decreasing solution of the pair*

$$zf\left(\frac{f(z)}{y}\right) = yf\left(\frac{f(y)}{z}\right) \tag{10}$$

$$f[f(x)] = x \tag{11}$$

of functional equations in $(0,1)$, *where* $0 < f(x) < 1$, *is*

$$f(x) = (1 - x^k)^{1/k} \qquad (k \neq 0). \tag{28}$$

By (9) this means

$$\bar{p}(\bar{\ }\mathbf{A}/\mathbf{V}) = [1-\bar{p}(\mathbf{A}/\mathbf{V})^k]^{1/k}.$$

In [10] (p. 22) again somewhat autocratically $k=1$ is taken. This can be motivated as follows: We have left in (7) the exponent a arbitrary. By choosing

$$\overset{\circ}{p}(\mathbf{A}/\mathbf{V}) = \bar{p}(\mathbf{A}/\mathbf{V})^k$$

we see that *one can redefine the conditional probabilities so that both*

$$\overset{\circ}{p}(\mathbf{A}\cdot\mathbf{B}/\mathbf{V}) = \overset{\circ}{p}(\mathbf{A}/\mathbf{B}\cdot\mathbf{V})\,\overset{\circ}{p}(\mathbf{B}/\mathbf{V}) \tag{29}$$

and

$$\overset{\circ}{p}(\bar{\ }\mathbf{A}/\mathbf{V}) = 1-\overset{\circ}{p}(\mathbf{A}/\mathbf{V}) \tag{30}$$

hold for conjoined resp. contradictory events, in accordance with [10] (3.7) and (4.5).[5] *By supposing f continuous in the points* 0 *and* 1 *too, we can extend our result also for* $\mathbf{A}=\mathbf{0}$ *and* $\mathbf{A}=\mathbf{V}$, excluded in Corollary 1.

4. Corollary 1 can be reformulated also to an assertion about the function F. In fact (12) and (11) imply

$$F[x, F(x,z)] = xf\left(\frac{F(x,z)}{x}\right) = xf\left[f\left(\frac{z}{x}\right)\right] = z.$$

On the other hand (14) gives

$$F[z, F(x,x)] = F[x, F(x,z)] = z.$$

Because of the supposed strict monotony of $F(z, y)$ in y, necessarily

$$F(x,x) = C \text{ (constant).} \tag{31}$$

(By (12) $C=F(x,x)=xf(1)$, so $f(1)=0$, $C=0$ must hold.)

Now (16) can satisfy (31) only if $b=-1$ and then (16) goes over with $\bar{h}(z)=h(z)-c$ into

$$F(x,y) = \bar{h}^{-1}(\bar{h}(x)-\bar{h}(y)) \tag{32}$$

similar to (4) (cf. [15]). Of course (23) cannot satisfy (31), and (25) satisfies (31) again only if $b=-1$, so we have – in full analogy with the result of [1] (Theorem 2) quoted above – (cf. also [25]) the

COROLLARY 2. *The most general solution, continuous and strictly monotonic in x and y of the system*

$$F(xz, yz) = F(x,y)z, \tag{13}$$

$$F[x, F(y,z)] = F[z, F(y,x)] \tag{14}$$

$$F(x,x) = C \tag{31}$$

[5] Also the idea in [10] of taking instead of $p(\mathbf{A}/\mathbf{V})$ an arbitrary function of it is somewhat vague, because – as we have just seen – after having reached (29) and (30) there remains nothing arbitrary in $p(\mathbf{A}/\mathbf{V})$ anymore.

of functional equations in (0,1) *is*

$$F(x, y) = (x^k - y^k)^{1/k} \qquad (k \neq 0).$$

(32) suggests the definition

$$\bar{p}(\mathbf{A}-\mathbf{B}/\mathbf{V}) = F[\bar{p}(\mathbf{A}/\mathbf{V}), \bar{p}(\mathbf{B}/\mathbf{V})] \qquad (\mathbf{B} \subseteq \mathbf{A})$$

and then the variant

$$F(x, xz) = xf(z)$$

of the seemingly less motivated definition (12) together with (8) and (9) reflects the obvious identity

$$\mathbf{A}-\mathbf{A}\cdot\mathbf{C} = \mathbf{A}\cdot\tilde{}\mathbf{C}.$$

On the other hand (cf. [10] pp. 24–25) by the known rules

$$\mathbf{A}+\mathbf{B} = \tilde{}(\tilde{}\mathbf{A}\cdot\tilde{}\mathbf{B}), \qquad \mathbf{A}\cdot\tilde{}\mathbf{B} = \tilde{}\mathbf{B}\cdot\mathbf{A}$$

of Boolean algebra and by (29), (30) we get for the *probability of disjoined events*

$$\mathring{p}(\mathbf{A}+\mathbf{B}/\mathbf{V}) = \mathring{p}[\tilde{}(\tilde{}\mathbf{A}\cdot\tilde{}\mathbf{B})/\mathbf{V}] = 1-\mathring{p}(\tilde{}\mathbf{A}\cdot\tilde{}\mathbf{B}/\mathbf{V}) = 1-\mathring{p}(\tilde{}\mathbf{A}/\tilde{}\mathbf{B}\cdot\mathbf{V})\mathring{p}(\tilde{}\mathbf{B}/\mathbf{V}) =$$

$$= 1-[1-\mathring{p}(\mathbf{A}/\tilde{}\mathbf{B}\cdot\mathbf{V})]\mathring{p}(\tilde{}\mathbf{B}/\mathbf{V}) = 1-\mathring{p}(\tilde{}\mathbf{B}/\mathbf{V})+\mathring{p}(\mathbf{A}/\tilde{}\mathbf{B}\cdot\mathbf{V})\mathring{p}(\tilde{}\mathbf{B}/\mathbf{V}) =$$

$$= \mathring{p}(\mathbf{B}/\mathbf{V})+\mathring{p}(\mathbf{A}\cdot\tilde{}\mathbf{B}/\mathbf{V}) = \mathring{p}(\mathbf{B}/\mathbf{V})+\mathring{p}(\tilde{}\mathbf{B}/\mathbf{A}\cdot\mathbf{V})\mathring{p}(\mathbf{A}/\mathbf{V}) = \mathring{p}(\mathbf{B}/\mathbf{V})+[1-$$

$$-\mathring{p}(\mathbf{B}/\mathbf{A}\cdot\mathbf{V})]\mathring{p}(\mathbf{A}/\mathbf{V}) = \mathring{p}(\mathbf{A}/\mathbf{V})+\mathring{p}(\mathbf{B}/\mathbf{V})-\mathring{p}(\mathbf{A}\cdot\mathbf{B}/\mathbf{V}).$$

In [2] we have begun with motivating this law in the case $\mathbf{A}\cdot\mathbf{B}=0$.

We remark yet that *the Theorems 1,2 and Corollaries 1,2 are valid instead of* (0,1] *also for other positive intervals and with some changes for negative varables too.*

Literature

[1] J. Aczél, A solution of some problems of K. Borsuk and L. Jánossy, *Acta Phys. Acad. Sci. Hung.*, **4** (1955), 351–362.

[2] J. Aczél, Über die Begründung der Additions-und Multiplikationsformeln von bedingten Wahrscheinlichkeiten, *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, **6** (1961), 110–122.

[3] J. Aczél, *Vorlesungen über Funktionalgleichungen und ihre Anwendungen*, (Lehrbücher und Monographien der exakt. Wiss., Math. Reihe XXV.) Basel–Stuttgart, 1961.

[4] G. A. Barnard, Statistical inference, *J. Roy. Statist. Soc. Ser. B.*, **11** (1949), 119–149.

[5] G. A. Barnard, The theory of information, *J. Roy. Statist. Soc. Ser. B.*, **13** (1951), 46–64.

[6] S. N. Bernstein, An attempt at an axiomatic exposition of the principles of the calculus of probabilities (Russ.), *Chark. Zap. Matem. O-va* (2), **15** (1917), 204–274.

[7] S. N. Bernstein, *Theory of the calculus of probabilities* (Russ.), 2d ed., Moscow 1934.

[8] A. Bohnenblust, Axiomatic characterization of L_p-spaces, *Duke Math. J.*, **6** (1940), 627–640.

[9] R. T. Cox, Probability, frequency and reasonable expectation, *Amer. J. Phys.*, **14** (1946), 1–13.

[10] R. T. Cox, *The algebra of probable inference*, Baltimore, 1961.
[11] Z. Daróczy, Notwendige und hinreichende Bedingungen für die Existenz von nichtkonstanten Lösungen linearer Funktionalgleichungen, *Acta Sci. Math. Szeged*, **22** (1961), 31–41.
[12] S. Gołąb, Przycznek do algebry działań w ciele rzeczywistych, *Rocznik Naukowo-Dydaktyczny W. S. P. w Krakowie, Mat.*, **1** (1954), 3–10.
[13] S. Gołąb, Zum distributiven Gesetz der reellen Zahlen, *Studia Math.*, **15** (1956), 353–358.
[14] I. J. Good, *Probability and the weighing of evidence*. Appendix III. London–New York, 1950.
[15] M. Hosszú, Some functional equations related with the associative law, *Publ. Math. Debrecen*, **3** (1954), 205–214.
[16] L. Jánossy, Remarks on the foundations of probability calculus, *Acta Phys. Acad. Sci. Hung.*, **4** (1955), 333–349.
[17] L. Jánossy, Überlegungen zu den Grundlagen der Wahrscheinlichkeitsrechnung. *Sitzb. Deutsch. Akad. Wiss. Berlin, Kl. Math. Phys. Tech.*, 1960, No. 2, 2–23.
[18] H. Richter, Zur Grundlegung der Wahrscheinlichkeitstheorie I, *Math. Ann.*, **125** (1952), 129–139.
[19] H. Richter, Zur Grundlegung der Wahrscheinlichkeitstheorie II, *Math. Ann.*, **125** (1952), 223–234.
[20] H. Richter, Zur Grundlegung der Wahrscheinlichkeitstheorie III. *Math. Ann.*, **125** (1952), 335–343.
[21] H. Richter, Zur Grundlegung der Wahrscheinlichkeitstheorie IV, *Math. Ann.*, **126** (1953), 362–374.
[22] H. Richter, Zur Grundlegung der Wahrscheinlichkeitstheorie V, *Math. Ann.*, **128** (1954), 305–339.
[23] H. Richter, *Wahrscheinlichkeitstheorie*. (Grundlehren der math. Wiss. LXXXV.) Berlin–Göttingen–Heidelberg, 1956.
[24] A. R. Schweitzer, Remarks on functional equations, *Bull. Amer. Math. Soc.*, **20** (1914), 72.
[25] A. R. Schweitzer, Remarks on functional equations. *Bull. Amer. Math. Soc.*, **21** (1915), 23–29.

ON GENERALIZED QUASI-STEP AND ALMOST-STEP PARABOLAS, RESPECTIVELY*

By

PAUL SZÁSZ

Department of Math. Analysis of the Eötvös Loránd University, Budapest

(Received March 7, 1963)

Assume $\alpha > -1$, $\beta > -1$. Let the zeros of the Jacobi polynomial

$$(1) \qquad \omega_n(x) = P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{n!\,2^n}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}[(1-x)^{\alpha+n}(1+x)^{\beta+n}]$$

$$(n = 0, 1, 2, \ldots)$$

(defined here by the well-known formula of Rodriguez) be denoted by

$$(2) \qquad x_{1n}, x_{2n}, \ldots, x_{nn}.$$

Then for the sum

$$(3) \qquad s_n = \sum_{\nu=1}^{n} \frac{1}{(1-x_{\nu n}^2)\omega_n'(x_{\nu n})^2}$$

which according to L. FEJÉR [1] has in the case $\alpha=\beta=0$ of the Legendre polynomials the value 1, it can be shown, that the formulas

$$(4) \qquad s_1 = \frac{1}{(\alpha+1)(\beta+1)}, \qquad s_n = \frac{1}{(\alpha+1)(\beta+1)}\prod_{\nu=2}^{n}\left(1-\frac{\alpha\beta}{(\alpha+\nu)(\beta+\nu)}\right)$$

are valid and

$$(5) \qquad s_n \to B(\alpha+1, \beta+1) = \frac{1}{2^{\alpha+\beta+1}}\int_{-1}^{1}(1-x)^{\alpha}(1+x)^{\beta}dx$$

if $n \to +\infty$.

The sum under (3) is with regard to (5) bounded, which in the case $\operatorname{sg}\alpha = \operatorname{sg}\beta$ follows from formula (4) itself. On the basis of this simple matter of fact which in the special case $\alpha=\beta$ of the ultraspherical polynomials was already shown by J. BALÁZS [2] (but in a completely different way), concerning the zeros under (2) of the Jacobi polynomial (1) the following can be proved.

THEOREM I. *In the case* $0 \leqq \alpha < 1$, $0 \leqq \beta < 1$ *the values*

$$(6) \qquad y_{1n}, y_{2n}', \ldots, y_{nn} \qquad (n = 1, 2, 3, \ldots)$$

having been chosen arbitrarily but with the restriction

$$(7) \qquad |y_{\nu n}'| \leqq \Delta$$

* Lecture delivered on the 18th of August, at the *International Congress of Mathematicians Stockholm* 1962.

where Δ is independent of n and ν, for each function $f(x)$, continuous in the closed interval $[-1, 1]$, for the polynomial $S_n(x)$ of degree $2n+1$ at most with the property

$$(8)\quad S_n(-1) = f(-1), \qquad S_n(1) = f(1), \qquad S_n(x_{\nu n}) = f(x_{\nu n}), \qquad S_n'(x_{\nu n}) = y'_{\nu n}$$
$$(\nu = 1, 2, \ldots, n; \quad n = 1, 2, 3, \ldots)$$

holds

$$(9)\qquad S_n(x) \to f(x) \qquad \text{if} \qquad n \to +\infty,$$

the convergence being uniform in each internal subinterval. In particular if $\alpha=\beta$ (i. e. for the ultraspherical polynomial $\omega_n(x)=P_n^{(\alpha)}(x)$) the convergence is uniform in the whole closed interval.

I have called in an earlier paper [3] the polynomial $S_n(x)$ of degree $2n+1$ at most defined by the $2n+2$ relations under (8) the *„generalized quasi-step parabola"* corresponding to these ones. This conception is a slight modification of that of the *„step parabola"* introduced by L. Fejér [4]. The above Theorem I was proved in my paper quoted above only for the case of the Legendre polynomial $P_n(x)$ and in that of the Tchebichef polynomial of the second kind $U_n(x)$. In the first case the theorem was a generalization of an approximation theorem of E. Egerváry and P. Turán [5]. L. Fejér [6] proved primarily a similar theorem concerning the *„generalized step parabola"* in the case when the fundamental points are the zeros of the Tchebichef polynomial of the first kind $T_n(x)$. His investigations were extended by G. Szegő [7] for Jacobi polynomials having solved for these the convergence problem of the generalized step parabolas completely.

Another type of step parabolas occurs in the following

Theorem II. *In the case $\alpha = \frac{1}{2}$, $\beta = -\frac{1}{2}$ the values under (6) having been chosen arbitrarily but with the restriction (7), for each function $f(x)$, continuous in the closed interval $[-1, 1]$, for the „generalized almost-step parabola" $S_n^*(x)$ of degree $2n$ at most belonging to the fundamental points under (2) and to this function $f(x)$, i. e. with the property*

$$(10)\qquad S_n^*(1) = f(1), \qquad S_n^*(x_{\nu n}) = f(x_{\nu n}), \qquad S_n^{*\prime}(x_{\nu n}) = y'_{\nu n}$$
$$(\nu = 1, 2, \ldots, n; \quad n = 1, 2, 3, \ldots),$$

holds

$$(11)\qquad S_n(x) \to f(x) \qquad \text{if} \qquad n \to +\infty$$

the convergence being uniform in the whole closed interval.

Disregarding a constant factor, the treated Jacobi polynomial $P_n^{\left(\frac{1}{2}, -\frac{1}{2}\right)}(x)$ is identical with

$$(12)\qquad \left[\frac{\sin(2n+1)\frac{\vartheta}{2}}{\sin\frac{\vartheta}{2}}\right]_{x=\cos\vartheta}$$

having the zeros

$$x_{\nu n} = \cos \frac{2\nu\pi}{2n+1} \qquad (\nu = 1, 2, \ldots, n). \tag{13}$$

The above Theorem II is an answer to a question put by P. TURÁN orally, concerning the convergence of this type of step parabolas I just named as „*generalized almost-step parabolas*".

Literature

[1] L. FEJÉR, Interpolációról (in Hungarian), *Mat. és Term. Értesitő*, **34** (1916), 209–229, spec. p. 221, form. (45), or by the same author, Über Interpolation, *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, Math.-phys. Kl.* (1916), 66–91, spec. p. 78, form. (39).

[2] J. BALÁZS, Megjegyzések a stabil interpolációról (in Hungarian), *Matematikai Lapok*, **11** (1960), 280–293, spec. p. 289, form. (4.7).

[3] P. SZÁSZ, On quasi-Hermite-Fejér interpolation, *Acta Math. Acad. Sci. Hung.*, **10** (1959), 413–439.

[4] L. FEJÉR, l. c. [1].

[5] E. EGERVÁRY and P. TURÁN, Notes on interpolation. V., *Acta Math. Acad. Sci. Hung.*, **9** (1958), 259–267, spec. pp. 264–265.

[6] L. FEJÉR, Über Weierstraßsche Approximation besonders durch Hermitesche Interpolation, *Mathematische Annalen*, **102** (1930), 707–725, spec. pp. 721–723.

[7] G. SZEGŐ, Über gewisse Interpolationspolynome, die zu den Jacobischen und Laguerreschen Abszissen gehören, *Mathematische Zeitschrift*, **35** (1932), 579–602, spec. pp. 583–592.

SUR DES CRITÈRES GÉNÉRAUX D'APPROXIMATION UNIFORME

Par

Á. CSÁSZÁR et J. CZIPSZER

Ie Chaire d' Analyse Mathématique, Université L. Eötvös, Budapest
et Institut de Mathématique de l'Académie des Sciences de Hongrie, Budapest

(Reçu le 13 avril 1963)

En généralisant le théorème d'approximation de Stone – Weierstrass, W. MAAK, G. NÖBELING et H. BAUER ont trouvé des résultats du type suivant ([3], [5]): On considère certaines classes de fonctions à valeurs réelles ou complexes (ou même vectorielles dans [5]), définies sur un ensemble commun E et on donne des critères nécessaires et suffisants pour qu'une fonction F, définie également sur E, puisse être approchée uniformément aussi près que l'on veut par des fonctions d'une classe du type considéré. Dans ce qui suit, nous préciserons partiellement l'un des théorèmes d'approximation de Nöbeling – Bauer dans ce sens que nous énoncerons les critères d'approximation pour un type de classe de fonctions qui est plus vaste (lorsqu'on se borne aux classes qui contiennent toutes les fonctions constantes) que celui envisagé en [5]. Puis nous montrerons par des contre-exemples que certaines tentatives de précisions plus poussées du théorème de NÖBELING – BAUER qu'on serait tenté de croire justifiées en vue des résultats antérieurs ne peuvent pas être menées à but.

§ 1.

Fixons un ensemble „fondamental" E. Dans la suite, „fonction" veut toujours dire „fonction réelle, définie sur E et *bornée*". $\mathfrak{S}$ étant une classe de fonctions, nous désignerons par $\overline{\mathfrak{S}}$ l'ensemble des fonctions qui sont limites uniformes de suites tirées de $\mathfrak{S}$, c'est-à-dire la fermeture de $\mathfrak{S}$ par rapport à la topologie de la convergence uniforme. On a donc évidemment $\overline{\overline{\mathfrak{S}}}=\overline{\mathfrak{S}}\supset\mathfrak{S}$.

Dans [5], on considère entre autres des classes non vides $\mathfrak{S}$ de fonctions qui remplissent les conditions suivantes:

(1) Si $f, g\in\mathfrak{S}$, on a $\alpha f+\beta g\in\mathfrak{S}$, α et β étant deux nombres réels quelconques;
(2) $f\in\mathfrak{S}$ entraîne $|f|\in\mathfrak{S}$;
(3) $f\in\mathfrak{S}$ implique $\min(1, f)\in\mathfrak{S}$.

Appelons une classe de cette sorte une *classe* (A). Si en outre

(4) $\mathfrak{S}$ contient toutes les fonctions constantes,

nous parlerons d'une *classe* (A^*). Rappelons que (1) et (2) entraînent, en vertu des formules

$$\min(f, g) = \frac{1}{2}(f+g-|g-f|),$$

$$\max(f, g) = \frac{1}{2}(f+g+|g-f|),$$

que $\min(f,g)\in\mathfrak{S}$, $\max(f,g)\in\mathfrak{S}$ chaque fois que $f,g\in\mathfrak{S}$. De cette façon, (3) s'ensuit de (1), (2) et (4), les classes (A^*) sont donc caractérisées par ces trois dernières conditions et aussi par (1), (4) et

(2′) $f, g\in\mathfrak{S}$ implique $\max(f,g)\in\mathfrak{S}$, $\min(f,g)\in\mathfrak{S}$,

eu égard à la formule

$$|f| = \max(f, -f).$$

De façon analogue, les classes (A) sont caractérisées par (1), (2′) et (3).

Citons maintenant le théorème suivant de NÖBELING–BAUER ([5], p. 58, Kriterium 2):

(5) *Soit $\mathfrak{S}$ une classe (A). Pour qu'une fonction F appartienne à $\overline{\mathfrak{S}}$, il faut et il suffit qu'à chaque $\varepsilon>0$ on puisse faire correspondre un $\delta>0$ et un nombre fini de fonctions $f_1, \ldots, f_n\in\mathfrak{S}$ de façon que*

(6) $x, y\in E$, $|F(y)-F(x)|\geqq\varepsilon$ *entraînent*

$$\max_i [|f_i(y)-f_i(x)|] \geqq \delta$$

et

(7) $x\in E$, $|F(x)|\geqq\varepsilon$ *entraînent* $\max_i |f_i(x)|\geqq\delta$.

Pour des classes (A^*), (7) est évidemment superflu, donc on a le corollaire suivant ([5], p. 58, Korollar):

(8) *Soit $\mathfrak{S}$ une classe (A^*). Pour qu'une fonction F appartienne à $\overline{\mathfrak{S}}$, il faut et il suffit qu'à chaque $\varepsilon>0$ il corresponde un $\delta>0$ et des fonctions $f_1, \ldots, f_n\in\mathfrak{S}$ qui vérifient* (6).[1]

Nous allons étendre ce théorème aux classes non vides $\mathfrak{S}$ de fonctions, dites *classes* (B) et caractérisées par les propriétés suivantes:

(9) Si $f\in\mathfrak{S}$, on a $f+\alpha\in\mathfrak{S}$ pour tout nombre réel α;

(10) Si $f\in\mathfrak{S}$, on a $\alpha f\in\mathfrak{S}$ pour tout nombre $\alpha\geqq 0$;

(11) Si $f, g\in\mathfrak{S}$, on a $\min(f,g)\in\mathfrak{S}$ et $\max(f,g)\in\mathfrak{S}$.[2]

Si de plus $\mathfrak{S}$ contient avec f la fonction $-f$ (ou ce qui revient au même, si $\mathfrak{S}$ vérifie (10) pour tous les α réels), nous disons que $\mathfrak{S}$ est une classe (B) *symétrique*.

En vertu de (10) et (9), les classes (B) contiennent toutes les fonctions constantes. Évidemment toute classe (A^*) est en même temps une classe (B) symétrique, mais la réciproque n'est pas vraie.

Or, on peut énoncer les théorèmes d'approximation suivants pour les classes (B):

(12) *Soient $\mathfrak{S}$ une classe (B), F une fonction.*

a) *Pour que F appartienne à $\overline{\mathfrak{S}}$, il faut et il suffit qu'à chaque $\varepsilon>0$ il corresponde un $\delta>0$ et des fonctions $f_1, \ldots, f_n\in\mathfrak{S}$ tels que*

[1] Ce théorème reste valable lorsque $\mathfrak{S}$ est une *algèbre*, c'est-à-dire qu'il vérifie (1), il contient les constantes et, avec f et g, leur produit fg (voir [3], p. 190).

[2] De cette façon, les classes (B) coïncident avec les „familles ordonnatrices simples" de l'ouvrage [1], pp. 141 et 142. Un type de classes très proche (où la condition que les fonctions envisagées doivent être bornées est remplacée par une autre) a été considéré dans [2] („autarke Funktionensysteme", p. 236).

(13) $x, y \in E$, $F(y) - F(x) \geqq \varepsilon$ *entraînent*

$$\max_i [f_i(y) - f_i(x)] \geqq \delta.$$

b) *Si* $\mathfrak{S}$ *est symétrique,* (13) *peut être remplacé par la condition* (6).

La nécessité des conditions énoncées étant évidente (même avec $n=1$), leur suffisance sera la conséquence d'un critère (suffisant) dans lequel la classe $\mathfrak{S}$ envisagée est assujettie à des conditions encore moins restrictives. Dans ce but, appelons *classe* (C) une classe $\mathfrak{S}$ de fonctions, non vide et satisfaisant aux conditions suivantes:

(14) Les fonctions constantes appartiennent à $\mathfrak{S}$;

(15) $f \in \mathfrak{S}$ implique $f + \alpha \in \mathfrak{S}$ pour tout nombre réel α;

(16) Si $f, g \in \mathfrak{S}$, on a $\min(f, g) \in \mathfrak{S}$ et $\max(f, g) \in \mathfrak{S}$.

Une classe (B) constitue évidemment une classe (C).

Or, pour les classes (C), on peut affirmer:

(17) *Soit* $\mathfrak{S}$ *une classe* (C) *et* F *une fonction telle que,* $\varepsilon > 0$ *et* $\delta > 0$ *étant donnés, il existe des fonctions* $f_1, \ldots, f_n \in \mathfrak{S}$ *satisfaisant à* (13). *Alors* F *appartient à* $\overline{\mathfrak{S}}$.

Pour démontrer (17), soient $\alpha < \beta$ et $\alpha' < \beta'$ des nombres réels donnés et posons

$$U = F^{-1}((-\infty, \alpha]), \qquad V = F^{-1}([\beta, +\infty)).$$

Nous montrons qu'il existe une fonction $h \in \mathfrak{S}$ telle que

$$\alpha' \leqq h(x) \leqq \beta' \quad \text{pour } x \in E, \quad h(x) = \alpha' \quad \text{pour } x \in U, \tag{18}$$
$$h(x) = \beta' \qquad \text{pour} \qquad x \in V.$$

À cet effet, nous pouvons admettre que $U \neq 0 \neq V$, puisque au cas contraire $h(x) = \beta'$ ou $h(x) = \alpha'$ satisfait à nos conditions. Par hypothèse, il y a des fonctions $f_1, \ldots, f_n \in \mathfrak{S}$ qui satisfont à (13) avec $\varepsilon = \beta - \alpha$ et $\delta = \beta' - \alpha' + 2$. Par conséquent, pour tout $x \in U$, $y \in V$, il y a un i tel que $f_i(y) - f_i(x) \geqq \delta$.

Les fonctions f_i étant bornées, il y a des décompositions finies de U et de V

$$U = U_1 \cup \cdots \cup U_p, \qquad V = V_1 \cup \cdots \cup V_q$$

en parties non vides et telles que l'oscillation de chaque fonction f_i sur chacun des ensembles U_j et V_k soit 1 au plus. Donc, à chaque j et k $(1 \leqq j \leqq p,\ 1 \leqq k \leqq q)$ il correspond un $i(j, k)$ de sorte qu'on ait

$$f_{i(j,k)}(y) - f_{i(j,k)}(x) \geqq \delta - 2 = \beta' - \alpha'$$

pour $x \in U_j$ et $y \in V_k$. Posons

$$\alpha_{jk} = \sup_{x \in U_j} f_{i(j,k)}(x),$$

$$g_{jk}(x) = \min\{\max[f_{i(j,k)}(x), \alpha_{jk}] + \alpha' - \alpha_{jk}, \beta'\}.$$

On conclut aisément

$$\alpha' \leqq g_{jk}(x) \leqq \beta' \qquad \text{pour} \qquad x \in E,$$

$$g_{jk}(x) = \alpha' \qquad \text{pour} \qquad x \in U_j, \qquad g_{jk}(x) = \beta' \qquad \text{pour} \qquad x \in V_k.$$

La fonction

$$h(x) = \min_j \max_k g_{jk}(x)$$

vérifie par suite (18). De plus, les conditions (14) à (16) entraînent que $\mathfrak{S}$ contient avec les f_i les g_{jk} et aussi h, conformément à ce que nous avons affirmé.

Ceci posé, choisissons un entier $M>0$ tel que $|F(x)|<M$ pour $x\in E$, et fixons un entier $m>0$. Effectuons la construction précédente pour $\alpha = \frac{s-1}{m}$, $\beta = \frac{s}{m}$ et $\alpha' = -M$, $\beta' = \frac{s}{m}$, où s parcourt les entiers tels que $-Mm<s\leqq Mm$. Désignons par h_s la fonction qui résulte par cette construction et posons

$$r(x) = \max_s h_s(x).$$

En vertu de (16), on a $r\in\mathfrak{S}$. Soit $x_0\in E$. Pour un certain entier s_0, on a

$$\frac{s_0-1}{m} \leqq F(x_0) < \frac{s_0}{m} \qquad \text{et} \qquad -Mm < s_0 \leqq Mm.$$

Par conséquent (cf. (18))

$$h_s(x_0) = \frac{s}{m} \qquad \text{pour} \qquad s < s_0,$$

$$-M \leqq h_{s_0}(x_0) \leqq \frac{s_0}{m},$$

$$h_s(x_0) = -M \qquad \text{pour} \qquad s > s_0,$$

donc

$$\frac{s_0-1}{m} \leqq r(x_0) \leqq \frac{s_0}{m}$$

et finalement

$$|F(x_0)-r(x_0)| \leqq \frac{1}{m}.$$

Étant donné que $r\in\mathfrak{S}$, on conclut que $F\in\overline{\mathfrak{S}}$, c. q. f. d.[3]

Pour démontrer la suffisance de (12), supposons que $\mathfrak{S}$ est une classe (B) et que, pour une fonction F et $\varepsilon>0$, il existe un $\delta>0$ et des fonctions $f_1, \ldots, f_n\in\mathfrak{S}$ satisfaisant à (13). Un $\delta'>0$ étant donné, les fonctions $\frac{\delta'}{\delta} f_i$ $(1\leqq i\leqq n)$ satisfont à (13) avec δ' au lieu de δ et appartiennent à $\mathfrak{S}$ par suite de (10). Le critère (17) fournit donc $F\in\overline{\mathfrak{S}}$. De plus, si $\mathfrak{S}$ est une classe (B) symétrique et si, pour F et $\varepsilon>0$, le nombre $\delta>0$ et les fonctions $f_1, \ldots, f_n\in\mathfrak{S}$ peuvent être choisies de façon à satisfaire à (6), le même nombre $\delta>0$ et les $2n$ fonctions $f_1, \ldots,$

[3] Le critère (17) peut être déduit d'un critère de [5] (p. 70, Kriterium 4). Pourtant, nous avons préféré d'en donner une démonstration directe.

$f_n, -f_1, \ldots, -f_n$ satisferont évidemment à (13), ce qui achève la démonstration.

À titre d'application ultérieure de (17), nous démontrons la généralisation suivante de (8):

(19) *Soit* $\mathfrak{S}$ *une classe satisfaisant à* (2'), (4) *et*
(1') *Si* $f, g \in \mathfrak{S}$, *on a* $f+g \in \mathfrak{S}$.
Pour qu'une fonction F *appartienne à* $\overline{\mathfrak{S}}$, *il faut et il suffit qu'à chaque* $\varepsilon > 0$ *il corresponde un* $\delta > 0$ *et des fonctions* $f_1, \ldots, f_n \in \mathfrak{S}$ *qui vérifient* (13). *En remplaçant* (1') *par*
(1'') *Si* $f, g \in \mathfrak{S}$, *on a* $f-g \in \mathfrak{S}$,
(13) *peut être remplacé par* (6).

En effet, (2'), (4) et (1') impliquent évidemment que $\mathfrak{S}$ est une classe (C) et que $f \in \mathfrak{S}$ entraîne $mf \in \mathfrak{S}$ pour tout entier $m > 0$. Par suite, si $F, \varepsilon > 0, \delta > 0$ et $f_1, \ldots, f_n \in \mathfrak{S}$ satisfont à (13), et si l'on choisit à $\delta' > 0$ un entier $m > \frac{\delta'}{\delta}$, les fonctions $mf_i \in \mathfrak{S}$ $(1 \leqq i \leqq n)$ satisfont à (13) avec δ' au lieu de δ. De plus, (1'') implique que $f \in \mathfrak{S}$ entraîne $mf \in \mathfrak{S}$ pour tout entier m (de signe quelconque), de sorte que si F, $\varepsilon > 0$, $\delta > 0$ et $f_1, \ldots, f_n \in \mathfrak{S}$ satisfont à (6), les fonctions $mf_i \in \mathfrak{S}$ et $-mf_i \in \mathfrak{S}$ $(1 \leqq i \leqq n)$ vérifient (13) avec δ' au lieu de δ, l'entier m étant choisi comme plus haut.

§ 2.

Il est immédiat que si $\mathfrak{S}$ est une classe (A) ou (A^*) ou (B) ou une classe (B) symétrique, $\overline{\mathfrak{S}}$ est du même type. Mais il y a encore plus: Si $\mathfrak{S}$ est une classe (A), $\overline{\mathfrak{S}}$ contient avec f et g leur produit fg ([5], p. 59, Satz 1). Voici un résultat analogue pour les classes (B):

(20) *Si* $\mathfrak{S}$ *est une classe* (B), $\overline{\mathfrak{S}}$ *contient avec deux quelconques de ses éléments leur somme et, si* $\mathfrak{S}$ *est symétrique, leur produit.*

Première démonstration.[4] $\mathfrak{S}$ étant une classe (B), $\overline{\mathfrak{S}}$ l'est aussi. Si $f, g \in \overline{\mathfrak{S}}$, la classe $\overline{\mathfrak{S}}$ et la fonction $f+g$ vérifient les conditions du théorème (12)a). En effet, à $\varepsilon > 0$ on fait correspondre $\delta = \frac{\varepsilon}{2}$ et les fonctions f et g: si $f(y)+g(y) - [f(x)+g(x)] \geqq \varepsilon$, il faut avoir soit $f(y)-f(x) \geqq \frac{\varepsilon}{2}$, soit $g(y)-g(x) \geqq \frac{\varepsilon}{2}$, donc (13) est vérifié. Par conséquent $f+g \in \overline{\overline{\mathfrak{S}}} = \overline{\mathfrak{S}}$.

Si $\mathfrak{S}$ est symétrique, $\overline{\mathfrak{S}}$ l'est aussi. Montrons que la classe $\overline{\mathfrak{S}}$ et la fonction fg vérifient les conditions de (12)b). À $\varepsilon > 0$ on fait correspondre $\delta = \frac{\varepsilon}{2M}$ et les fonctions f et g, où M désigne une borne supérieure commune de $|f|$ et de $|g|$. Si

$$|f(y)\,g(y) - f(x)\,g(x)| \geqq \varepsilon,$$

[4] Cette démonstration est essentiellement la même que celle de Nöbeling–Bauer, donnée pour le théorème précité.

on a

$$\varepsilon \leqq |f(y)-f(x)|\cdot|g(y)|+|g(y)-g(x)|\cdot|f(x)| \leqq$$
$$\leqq M[|f(y)-f(x)|+|g(y)-g(x)|],$$

donc ou bien

$$|f(y)-f(x)| \geqq \frac{\varepsilon}{2M},$$

ou bien

$$|g(y)-g(x)| \geqq \frac{\varepsilon}{2M}.$$

C'est-à-dire, (6) est rempli. Par conséquent $fg \in \overline{\overline{\mathfrak{S}}} = \overline{\mathfrak{S}}$.

Deuxieme démonstration. Nous partons de la formule élémentaire

$$\xi+\eta = 2 \min_{-\infty<\gamma<\infty} \max (\xi+\gamma, \eta-\gamma)$$

ou ξ et η sont deux nombres réels quelconques. De plus, on voit sans peine que si I est un intervalle symétrique par rapport à l'origine et si I_n est l'ensemble (fini) des points qui divisent I en n parties égales (y compris les extrémités de I), alors

$$0 \leqq 2 \min_{\gamma \in I_n} (\max(\xi+\gamma, \eta-\gamma)-(\xi+\eta)) \leqq \frac{|I|}{n} \qquad \text{pour} \qquad \xi, \eta \in I,$$

où $|I|$ désigne la longueur de I. On tire de cette formule que si $f, g \in \overline{\mathfrak{S}}$ et si I contient toutes les valeurs de f et de g, alors la suite des fonctions

$$2 \min_{\gamma \in I_n} [\max (f+\gamma, g-\gamma)]$$

qui appartiennent également à $\overline{\mathfrak{S}}$, converge uniformément vers $f+g$. Par conséquent $f+g \in \overline{\overline{\mathfrak{S}}} = \overline{\mathfrak{S}}$.

Soit maintenant $\mathfrak{S}$ symétrique. On a évidemment pour tout nombre réel

$$\xi^2 = \max_{-\infty<\gamma<\infty} [\xi^2-(\xi-\gamma)^2] = \max_{-\infty<\gamma<\infty} (2\gamma\xi-\gamma^2),^5$$

plus précisément, si I et I_n ont la même signification que plus haut,

$$0 \leqq \xi^2 - \max_{\gamma \in I_n} (2\xi\gamma-\gamma^2) \leqq \frac{2|I|^2}{n} \qquad \text{pour} \qquad \xi \in I.$$

Par conséquent, si $f \in \overline{\mathfrak{S}}$ et $I \supset f(E)$, on a

$$\max_{\gamma \in I_n} [2\gamma f(x)-\gamma^2] \to f(x)^2$$

uniformément pour $x \in E$. Vu encore que

$$\max_{\gamma \in I_n} (2\gamma f-\gamma^2) \in \overline{\mathfrak{S}},$$

[5] Cf. p. ex. [6], p. 8. formule (7).

on conclut $f^2 \in \overline{\overline{\mathfrak{S}}} = \overline{\mathfrak{S}}$. Enfin si $f, g \in \overline{\mathfrak{S}}$, on tire de la formule

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$$

et des résultats précédents que $fg \in \overline{\mathfrak{S}}$.

REMARQUE. Grâce au théorème (20), le théorème (12)b) peut être ramené au théorème (8). En effet, si $\mathfrak{S}$ est une classe (B) symétrique, $\overline{\mathfrak{S}}$ est une classe (A^*) en vertu de (20) et si $\mathfrak{S}$ et F vérifient les hypothèses de (12)b), $\overline{\mathfrak{S}}$ et F vérifient à plus forte raison ces hypothèses, identiques à celles de (8), donc en vertu de (8) $F \in \overline{\overline{\mathfrak{S}}} = \overline{\mathfrak{S}}$. (Il ne s'agit pas de cercle vicieux, puisque la deuxième démonstration de (20) ne fait pas appel à (12)). Par contre, (12)a) est indépendant de (8).

§ 3.

1. En essayant de généraliser le théorème (5) de la même façon que nous avons généralisé le théorème (8), nous nous sommes posés la question suivante:

Soit $\mathfrak{S}$ une classe non vide de fonctions, jouissant des propriétés suivantes:

$$f, g \in \mathfrak{S} \text{ entraîne } f+g \in \mathfrak{S}, \ \min(f, g) \in \mathfrak{S}, \ \max(f, g) \in \mathfrak{S},$$

$$f \in \mathfrak{S} \text{ entraîne } \min(1, f) \in \mathfrak{S}, \ \max(-1, f) \in \mathfrak{S} \text{ et}$$

$$\alpha f \in \mathfrak{S} \text{ pour tout } \alpha \geqq 0.$$

Soit F une fonction qui vérifie le critère suivant:

À tout $\varepsilon > 0$ il existe un $\delta > 0$ et des fonctions $f_1, \ldots, f_n \in \mathfrak{S}$ telles que

$F(y) - F(x) \geqq \varepsilon$ entraîne $\max_i [f_i(y) - f_i(x)] \geqq \delta$ pour tout $x, y \in E$

et

$F(x) \geqq \varepsilon$ entraîne $\max_i [f_i(x)] \geqq \delta$,

$F(x) \leqq -\varepsilon$ entraîne $\min_i [f_i(x)] \leqq -\delta$ pour tout $x \in E$.

Ces hypothèses impliquent-elles $F \in \overline{\mathfrak{S}}$? L'exemple suivant montre que ce n'est pas le cas.

Soit $E = [0, 1]$ et soit $\mathfrak{S}$ l'ensemble des fonctions continues dont le graphique est composé d'un nombre fini de segments linéaires de la forme $y = ax + b$, où $a \geqq 0$, $b \geqq 0$. Posons

$$F(x) = \begin{cases} 0, & \text{si } 0 \leqq x < \frac{1}{2}, \\ x - \frac{1}{2}, & \text{si } \frac{1}{2} \leqq x \leqq 1. \end{cases}$$

Nous laissons au lecteur le soin de vérifier que $\mathfrak{S}$ et F satisfont aux conditions précédentes. De plus, $\mathfrak{S}$ contient les constantes $\alpha \geqq 0$. Reste à voir que $F \notin \overline{\mathfrak{S}}$.

Montrons d'abord que

(21) Si $f \in \mathfrak{S}$, $c > 0$ et pour un certain $x_0 \in E$ on a $f(x_0) < cx_0$, cette inégalité reste valable pour tout $x_0 < x \leqq 1$.

Supposé le contraire, soit x_1 le plus petit nombre pour lequel $x_0 < x_1 \leqq 1$, $f(x_1) = cx_1$ et soit $f(x) = ax + b$ dans un voisinage gauche de x_1. On a évidemment $a > c$ et $ax_1 + b = cx_1$, donc $b < 0$ contrairement à l'hypothèse.

Ceci établi, soit $\varepsilon > 0$ et supposons qu'il existe une fonction $f \in \mathfrak{S}$ telle que $|F(x) - f(x)| < \varepsilon$ pour $0 \leqq x \leqq 1$. On a donc

$$f\left(\frac{1}{2}\right) < \varepsilon = (2\varepsilon) \cdot \frac{1}{2}, \tag{22}$$

$$f(1) > \frac{1}{2} - \varepsilon. \tag{23}$$

Il suit de (21) et de (22)

$$f(x) < 2\varepsilon x \qquad \text{pour} \qquad \frac{1}{2} < x \leqq 1,$$

en particulier $f(1) < 2\varepsilon$, ce qui contredit à (23), si $\varepsilon \leqq \frac{1}{6}$. $F \in \overline{\mathfrak{S}}$ est donc impossible.

2. En remplaçant les classes (A^*) par les classes (B), nous avons considérablement affaibli l'exigence de linéarité. Or, si l'on va plus loin et, dans la définition des classes (B), on exige au lieu de (9) seulement que $\mathfrak{S}$ contienne toutes les fonctions constantes, le théorème (12) ne reste plus valable. Pour le voir, nous posons $E = [0, 1]$, et désignons par $\mathfrak{S}$ l'ensemble des fonctions non-négatives et des fonctions non-positives. Soit $F(x) = x - \frac{1}{2}$. Il est clair que $\mathfrak{S}$ vérifie (10) (même pour $\alpha < 0$) et (11), contient toutes les fonctions constantes, et F vérifie les conditions de (12)a) et celles de (12)b). Néanmoins $F \notin \overline{\mathfrak{S}}$, puisque de façon évidente $F \notin \mathfrak{S}$ et $\overline{\mathfrak{S}} = \mathfrak{S}$.

§ 4.

Les résultats précédents montrent que les classes (A^*) fermées par rapport à la topologie de la convergence uniforme coïncident avec les algèbres de fonctions fermées par rapport à la même topologie.[6] Or, celles-ci se caractérisent de la façon suivante:

(24) *$\mathfrak{S}$ étant une famille de fonctions* (réelles, bornées) *sur un ensemble E, pour que $\mathfrak{S}$ soit une algèbre fermée par rapport à la topologie de la convergence, uniforme, il faut et il suffit qu'il existe sur E une structure uniforme $\mathscr{U}$ telle que $\mathfrak{S}$ coïncide avec l'ensemble de toutes les fonctions* (réelles, bornées) *uniformément continues par rapport à $\mathscr{U}$.*[7]

[6] En effet, une classe (A^*) est une classe (B) symétrique et une classe (B) symétrique et fermée par rapport à la topologie de la convergence uniforme est une algèbre [v. (20)]; réciproquement, une algèbre fermée par rapport à la topologie en question est une classe (A^*), la fonction $|t|$ pouvant tre approchée uniformément dans un intervalle fini par des polynômes de t. Voir aussi [5], p. 59, Satz 1 et Satz 2.

[7] Voir [5], p. 64, Satz 6.

Pour les classes (B) (non-symétriques) fermées par rapport à la topologie de la convergence uniforme, on peut donner une caractérisation analogue. Pour cette fin, convenons d'appeler *structure quasi-uniforme* sur un ensemble E une famille $\mathscr{U}$ de relations binaires $\mathbf{U}$ sur E satisfaisant aux conditions:

(25) Toute relation $\mathbf{U}\in\mathscr{U}$ est réflexive;

(26) À deux relations $\mathbf{U}_1$, $\mathbf{U}_2\in\mathscr{U}$ il correspond une relation $\mathbf{U}\in\mathscr{U}$ telle que $x\mathbf{U}_i y$ implique $x\mathbf{U}y$ pour $i=1, 2$;

(27) À chaque relation $\mathbf{U}\in\mathscr{U}$ il correspond une relation $\mathbf{U}'\in\mathscr{U}$ telle que $x\mathbf{U}'y$ et $y\mathbf{U}'z$ entraînent $x\mathbf{U}z$.[8]

L'ensemble E, muni d'une structure quasi-uniforme, s'appelle *espace quasi-uniforme*. Par exemple, R désignant la droite numérique, les relations $\mathbf{U}_\varepsilon$ $(\varepsilon>0)$ définies par

$$x\mathbf{U}_\varepsilon y \text{ veut dire que } y-x<\varepsilon$$

forment une structure quasi-uniforme $\mathscr{V}$ sur R. Nous dirons qu'une fonction (réelle, bornée) f, définie dans un ensemble E, muni d'une structure quasi-uniforme $\mathscr{U}$, est *uniformément continue* par rapport à $\mathscr{U}$, si elle est uniformément continue en tant qu'application de l'espace quasi-uniforme $[E, \mathscr{U}]$ dans l'espace quasi-uniforme $[R, \mathscr{V}]$, c'est-à-dire si, pour $\varepsilon>0$ quelconque, il existe $\mathbf{U}\in\mathscr{U}$ tel que

$$x\mathbf{U}y \text{ implique } f(y)-f(x)<\varepsilon.$$

On voit aisément que si $\mathscr{U}$ est une structure quasi-uniforme sur E, la famille de toutes les fonctions uniformément continues par rapport à $\mathscr{U}$ est une classe (B) fermée par rapport à la topologie de la convergence uniforme. Réciproquement, soit $\mathfrak{S}$ une classe (B) sur l'ensemble E. Considérons les relations $\mathbf{U}(f_1, \ldots, f_n; \varepsilon)$ définies pour $n=1, 2, \ldots$, $f_i\in\mathfrak{S}$ $(1\leqq i\leqq n)$ et $\varepsilon>0$ de la manière suivante:

$$x\,\mathbf{U}(f_1, \ldots, f_n; \varepsilon)\, y \text{ veut dire que } f_i(y)-f_i(x)<\varepsilon$$
$$\text{pour } 1\leqq i\leqq n.$$

Il est aisé de vérifier que la famille de ces relations constitue une structure quasi-uniforme $\mathscr{U}$ sur E. De plus, les fonctions $f\in\mathfrak{S}$ sont uniformément continues par rapport à $\mathscr{U}$ puisque $x\mathbf{U}(f;\varepsilon)y$ implique $f(y)-f(x)<\varepsilon$. Enfin, si $\mathfrak{S}$ est fermée par rapport à la topologie de la convergence uniforme, toute fonction F uniformément continue par rapport à $\mathscr{U}$ appartient à $\mathfrak{S}$. En effet, $\varepsilon>0$ étant donné, il existe $f_1, \ldots, f_n\in\mathfrak{S}$ et $\delta>0$ tels que

$$x\mathbf{U}(f_1, \ldots, f_n; \delta)y \text{ implique } F(y)-F(x)<\varepsilon,$$

c'est-à-dire tels que

$$F(y)-F(x)\geqq\varepsilon \text{ entraîne } \max_i\,(f_i(y)-f_i(x))\geqq\delta.$$

Autrement dit, F satisfait à la condition (13) et par suite, en vertu de (12)a), $F\in\overline{\mathfrak{S}}=\mathfrak{S}$.

[8] Voir [1], p. 68. Cette notion est due à L. NACHBIN [4] et à D. TAMARI [7].

Par là, nous avons démontré le théorème suivant:

(28) *$\mathfrak{S}$ étant une famille de fonctions* (réelles, bornées) *sur un ensemble E, pour que $\mathfrak{S}$ soit une classe (B) fermée par rapport à la topologie de la convergence uniforme, il faut et il suffit qu'il existe sur E une structure quasi-uniforme $\mathcal{U}$ telle que $\mathfrak{S}$ coïncide avec l'ensemble de toutes les fonctions* (réelles, bornées) *uniformément continues par rapport à $\mathcal{U}$.*

Littérature

[1] Á. Császár, *Fondements de la topologie générale* (Budapest et Paris, 1960).
[2] H. Hahn, *Reelle Funktionen* (Leipzig, 1932).
[3] W. Maak, Eine Verallgemeinerung des Weierstraßschen Approximationssatzes, *Arch. Math.*, **6** (1955), 188–193.
[4] L. Nachbin, Sur les espaces uniformes ordonnés, *C. R. Paris*, **226** (1948), 774–775.
[5] G. Nöbeling–H. Bauer, Allgemeine Approximationskriterien mit Anwendungen, *Jber. Deutsch. Math.-Verein.*, **58** (1955), 54–72.
[6] F. Riesz, Sur la théorie ergodique des espaces abstraits, *Acta Sci. Math.*, **10** (1941), 1–20.
[7] D. Tamari, On a generalization of uniform structures and spaces, *Bull. Res. Council Israel*, **3** (1954), 417–428.

ON „SMALL" COEFFICIENTS OF THE POWER SERIES OF ENTIRE FUNCTIONS

By

ALFRÉD RÉNYI and CATHERINE RÉNYI

Department of Theory of Probability and Analysis of the Eötvös Loránd University Budapest

(Received May 6, 1963)

§ 1. Introduction

The second named author has proved (see [1]) some years ago the following

THEOREM A. *Let $f(z)$ be a transcendental entire function and let*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

and

$$f(z) = \sum_{n=0}^{\infty} b_n (z-1)^n$$

be the power series of $f(z)$ around the points[1] *$z=0$ and $z=1$ respectively. Let $Z_0(n)$ resp. $Z_1(n)$ denote the number of zeros in the sequence $a_0, a_1, \ldots, a_{n-1}$ resp. $b_0, b_1, \ldots, b_{n-1}$. Then*

$$\liminf_{n\to\infty} \frac{Z_0(n)+Z_1(n)}{n} \leq 1.$$

Theorem A can be formulated somewhat vaguely in the following way: if there are „many" zeros among the coefficients of the power series of a transcendental entire function about the point $z=0$, then there can not be „too many" zeros among the coefficients of its power series about the point $z=1$.

In the present paper we shall prove a generalization of Theorem A, whose statement can be expressed somewhat vaguely as follows: if „many" of the coefficients of the power series of an entire transcendental function about the point $z=0$ are „small", then there can not be „too many" „small" coefficients in the power series of the same function about the point $z=1$. Of course one has to give a precise meaning to the word „small" in this context.

Roughly speaking we shall call a coefficient of the power series of an entire function $f(z)$ „small" if it is smaller in absolute value than the corresponding coefficient of the power series of an other entire function $g(z)$ which is of smaller order of magnitude than $f(z)$. To make this definition quite definite, we have to decide how to compare the orders of magnitude of the entire functions $f(z)$ and $g(z)$. This could be done for example by comparing the rate of increase for $r\to\infty$ of the maximum modulus $M_f(r) = \operatorname*{Max}_{|z|=r} |f(z)|$ of $f(z)$ with that of $g(z)$, i. e. with $M_g(r) = \operatorname*{Max}_{|z|=r} |g(z)|$.

[1] The theorem remains valid if we consider the power series of $f(z)$ around two arbitrary different points α and β.

For our purposes it is however more convenient to compare the orders of magnitude of $f(z)$ and $g(z)$ by comparing the maximal terms of their power series. We shall prove the following

THEOREM 1. *Let* $f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n (z-1)^n$ *be a transcendental entire function and* $g(z) = \sum_{n=0}^{\infty} c_n z^n$ *be an arbitrary entire function. Put*

$$\mu_f(r) = \operatorname*{Max}_n |a_n| r^n$$

and

$$\mu_g(r) = \operatorname*{Max}_n |c_n| r^n.$$

Let us suppose that there exists a number δ *such that* $0<\delta<1$ *and*

$$\liminf_{r\to\infty} \frac{\mu_g(r^{1/\delta})}{\mu_f(r)} < \infty. \tag{1.1}$$

Let $S_0(n)$ *denote the number of indices* $k < n$ *for which* $|a_k| \leq |c_k|$ *and* $S_1(n)$ *the number of indices* $k < n$ *for which* $|b_k| \leq |c_k|$. *Then we have*[2]

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1+2\delta. \tag{1.2}$$

Let us make some remarks.

REMARK 1. If we choose $g(z) \equiv 0$, then (1.1) holds for any $\delta>0$, further in this case $S_0(n) = Z_0(n)$ and $S_1(n) = Z_1(n)$, where $Z_0(n)$ resp. $Z_1(n)$ denote the number of zeros in the sequence $a_0, a_1, \ldots, a_{n-1}$ resp. $b_0, b_1, \ldots, b_{n-1}$. Thus in this special case Theorem 1 reduces to Theorem A, i. e. Theorem 1 is a generalization of Theorem A.

REMARK 2. If $f(z)$ is an entire function of order ϱ_f and $g(z)$ an entire function of order ϱ_g, then according to a well known theorem (see e. g. [2], Chapter IV. Problem 51)

$$\limsup_{r\to\infty} \frac{\log\log \mu_f(r)}{\log r} = \varrho_f$$

and similarly

$$\limsup_{r\to\infty} \frac{\log\log \mu_g(r)}{\log r} = \varrho_g. \tag{1.3}$$

Thus if $\varrho_g<\varrho_f$ and we choose a sequence r_k $(k = 1, 2, \ldots)$ such that $\lim_{k\to\infty} r_k = \infty$ and $\lim_{k\to\infty} \frac{\log\log \mu_f(r_k)}{\log r_k} = \varrho_f$ then we have for any δ with $\delta>\frac{\varrho_g}{\varrho_f}$

$$\lim_{k\to\infty} \frac{\mu_g(r_k^{1/\delta})}{\mu_f(r_k)} = 0.$$

[2] Clearly only the case $\delta<\frac{1}{2}$ is interesting as otherwise (1.2) is trivial.

It follows that the conditions of Theorem 1 hold with any δ for which $\delta > \frac{\varrho_g}{\varrho_f}$ and thus we obtain

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1+2\frac{\varrho_g}{\varrho_f}. \tag{1.4}$$

In case $\varrho_g = 0$ and $\varrho_f > 0$ further also in case $\varrho_f = \infty$ and $\varrho_g < \infty$ it follows

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1.$$

We shall show later (Theorem 4) that (1.4) can be improved, namely the factor 2 on the right hand side of (1.4) is unnecessary.

Of course, Theorem 1 can also be applied to certain pairs of entire functions which are either both of order 0 or both of order ∞.

REMARK 3. Clearly instead of considering the power series of $f(z)$ about $z=0$ and $z=1$ we could consider its power series about $z=\alpha$ and $z=\beta$ where α and β are arbitrary complex numbers and $\alpha \neq \beta$. As a matter of fact in this case we have to apply Theorem 1 to the function $f[(\beta-\alpha)z+\alpha]$ instead of $f(z)$ and $g[(\beta-\alpha)z]$ instead of $g(z)$.

§ 2 contains estimations of certain interpolatory polynomials. In **§ 3** we give the proof of Theorem 1, further of the related Theorems 2,3 and 4. The method of proof of Theorem 1 is the same as that of Theorem A given in [1] but besides the tools used there, the results of **§ 2** are also needed. **§ 4** contains the discussion of an example, which shows that our results are not far from being best possible.

§ 2. Estimates of certain interpolatory polynomials

Let $Q_n(z)$ $(n=1, 2, \ldots)$ be the unique polynomial of degree $2n-1$ which satisfies the following conditions:

$$\frac{Q_n^{(j)}(0)}{j!} = p_j \qquad \text{for} \qquad j = 0, 1, \ldots, n-1 \tag{2.1}$$

and

$$\frac{Q_n^{(j)}(1)}{j!} = q_j \qquad \text{for} \qquad j = 0, 1, \ldots, n-1 \tag{2.2}$$

where p_j and q_j $(j=0, 1, \ldots, n-1)$ are arbitrary complex numbers. An explicit formula for $Q_n(z)$ has been given by P. JOHANSEN [3]; this can be written as follows:

$$\begin{aligned} Q_n(z) = (1-z)^n \sum_{k=0}^{n-1} z^k \left[\sum_{s=0}^{k} p_{k-s} \binom{n+s-1}{s}\right] + \\ + z^n \sum_{k=0}^{n-1} (z-1)^k \left[\sum_{s=0}^{k} (-1)^s q_{k-s} \binom{n+s-1}{s}\right]. \end{aligned} \tag{2.3}$$

Formula (2.3) can be transfromed as follows:

$$
(2.4)\qquad \begin{aligned} Q_n(z) = \sum_{m=0}^{2n-1} z^m \Bigg\{ \sum_{k=0}^{n-1} \binom{n}{m-k} (-1)^{m-k} \Bigg[\sum_{s=0}^{k} p_{k-s} \binom{n+s-1}{s} \Bigg] \Bigg\} + \\ + \sum_{t=0}^{n-1} z^{n+t} \Bigg\{ \sum_{k=t}^{n-1} \binom{k}{t} (-1)^{k-t} \Bigg[\sum_{s=0}^{k} (-1)^s q_{k-s} \binom{n+s-1}{s} \Bigg] \Bigg\}. \end{aligned}
$$

Thus we obtain

$$
\frac{Q_n^{(n)}(z)}{n!} = \sum_{t=0}^{n-1} z^t \binom{n+t}{n} Q_{n,t}
$$

where

$$
\begin{aligned} Q_{n,t} = (-1)^n \sum_{k=t}^{n-1} \binom{n}{n+t-k} (-1)^{k-t} \Bigg[\sum_{s=0}^{k} p_{k-s} \binom{n+s-1}{s} \Bigg] + \\ + \sum_{k=t}^{n-1} \binom{k}{t} (-1)^{k-t} \Bigg[\sum_{s=0}^{k} (-1)^s q_{k-s} \binom{n+s-1}{s} \Bigg]. \end{aligned}
$$

Now evidently

$$
\begin{aligned} |Q_{n,t}| \leq \Big(\max_{0 \leq u \leq n-1} |p_u| \Big) \sum_{k=t}^{n-1} \binom{n}{n+t-k} \Bigg[\sum_{s=0}^{k} \binom{n+s-1}{s} \Bigg] + \\ + \Big(\max_{0 \leq v \leq n-1} |q_v| \Big) \sum_{k=t}^{n-1} \binom{k}{t} \Bigg[\sum_{s=0}^{k} \binom{n+s-1}{s} \Bigg]. \end{aligned}
$$

As however

$$
\sum_{s=0}^{k} \binom{n+s-1}{s} = \binom{n+k}{k} < 2^{n+k}
$$

we obtain

$$
\sum_{k=t}^{n-1} \binom{n}{n+t-k} \Bigg[\sum_{s=0}^{k} \binom{n+s-1}{s} \Bigg] < 2^{n+t} \cdot 3^n < 12^n
$$

further

$$
\sum_{k=t}^{n-1} \binom{k}{t} \Bigg[\sum_{s=0}^{k} \binom{n+s-1}{s} \Bigg] < 8^n.
$$

Thus we obtain

$$
|Q_{n,t}| < 12^n \Big(\max_{0 \leq u \leq n-1} |p_u| + \max_{0 \leq v \leq n-1} |q_v| \Big) \qquad \text{for} \qquad t = 0, 1, \ldots, n-1.
$$

Thus we have proved the following

LEMMA 1. *Let $Q_n(z)$ $(n=1, 2, \ldots)$ be the unique polynomial of degree $2n-1$ determined by the conditions* (2.1) *and* (2.2). *Then we have*

$$
\max_{|z|=\varrho} \frac{|Q_n^{(n)}(z)|}{n!} < 12^n \Big(\max_{0 \leq u \leq n-1} |p_u| + \max_{0 \leq v \leq n-1} |q_v| \Big) \sum_{t=0}^{n-1} \binom{n+t}{n} \varrho^t.
$$

§ 3. Proof of Theorems 1–4

We can suppose without the restriction of generality, that $f(z)$ is real for real z, i. e. that all the coefficients a_n (and thus all the b_n) are real. As a matter of fact suppose that Theorem 1 is already proved for such functions. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is an arbitrary entire function, put $f_1(z) = \sum_{n=0}^{\infty} a_n' z^n$ resp. $f_2(z) = \sum_{n=0}^{\infty} a_n z^n$ where a_n' and a_n'' denotes the real and imaginary part of a_n respectively $(n=0, 1, \ldots)$. Clearly

$$\mu_f(r) \leq \mu_{f_1}(r) + \mu_{f_2}(r)$$

thus if $f(z)$ satisfies (1.1), at least one of the functions $f_1(z)$ and $f_2(z)$ satisfies it too. Suppose e. g. that $f_1(z)$ satisfies (1.1) i. e.

$$\liminf_{r\to\infty} \frac{\mu_g(r^{1/\delta})}{\mu_{f_1}(r)} < \infty.$$

Put $f(z) = \sum_{n=0}^{\infty} b_n (z-1)^n$, further $b_n = b_n' + ib_n''$; now if $|a_n| \leq |c_n|$ then $|a_n'| \leq |c_n|$, and similarly if $|b_n| \leq |c_n|$ then $|b_n'| \leq |c_n|$. Thus if $S_0'(n)$ resp. $S_1'(n)$ denotes the number of indices $k < n$ for which $|a_n'| \leq |c_n|$ resp. $|b_n'| \leq |c_n|$ then $S_0(n) \leq S_0'(n)$ and $S_1(n) \leq S_1'(n)$. As by supposition the statement of Theorem 1 is valid for $f_1(z)$, it follows that it is valid for $f(z)$ too. Therefore in what follows we shall suppose that $f(z)$ is real on the real axis, i. e. that the coefficients a_n and b_n are all real numbers.

To prove Theorem 1 we shall need the following Lemma A, which has been used in proving Theorem A in [1].

LEMMA A. *If $h(x)$ is a real function which has n continuous derivatives in the closed interval $[0, 1]$ and if $N(n)$ denotes the number of zeros of $h^{(n)}(x)$ in the interval $[0, 1]$ further $Z_0(n, h)$ and $Z_1(n, h)$ denotes the number of zeros in the sequence $h(0), h'(0), \ldots, h^{(n-1)}(0)$ and in the sequence $h(1), h'(1), \ldots, h^{(n-1)}(1)$ respectively, then we have*

$$Z_0(n, h) + Z_1(n, h) - n \leq N(n).$$

We shall apply Lemma A to the function

$$h_n(z) = f(z) - Q_n(z) \qquad (n = 1, 2, \ldots) \tag{3.1}$$

where $Q_n(z)$ is the unique polynomial of order $2n-1$ defined by the following conditions:

$$\frac{Q_n^{(j)}(0)}{j!} = \begin{cases} 0 & \text{for } 0 \leq j < \delta' n \\ a_j & \text{if } \delta' n \leq j \leq n-1 \text{ and } |a_j| \leq |c_j| \\ 0 & \text{if } \delta' n \leq j \leq n-1 \text{ and } |a_j| > |c_j| \end{cases} \tag{3.2a}$$

$$\frac{Q_n^{(j)}(1)}{j!} = \begin{cases} 0 & \text{for } 0 \leq j < \delta' n \\ b_j & \text{if } \delta' n \leq j \leq n-1 \text{ and } |b_j| \leq |c_j| \\ 0 & \text{if } \delta' n \leq j \leq n-1 \text{ and } |b_j| > |c_j|. \end{cases} \tag{3.2b}$$

Here δ' is an arbitrary number such that $\delta < \delta' < 1$.

Clearly if $h_n(z)$ is defined by (3.1) and $Q_n(z)$ by (3.2a–b) we have

$$S_0(n) \leq Z_0(n, h_n) + \delta' n \tag{3.3}$$

and

$$S_1(n) \leq Z_1(n, h_n) + \delta' n. \tag{3.4}$$

Thus it follows by Lemma A that denoting by $N_n(n)$ the number of zeros of $h_n^{(n)}(z)$ in the interval [0, 1], we have

$$S_0(n) + S_1(n) \leq N_n(n) + n(1 + 2\delta'). \tag{3.5}$$

Now we shall need *Jensen*'s theorem which we shall use in the following form (see [4], p. 127, formula (13)):

LEMMA B. *Let $H(z)$ be an analytic function which is regular in a circle $|z| \leq R$ $(R > 1)$ and $H(o) \neq 0$. Let $\mathscr{N}_0$ denote the number of zeros of $H(z)$ in the (closed) unit circle, and choose a number ϱ with $1 < \varrho < R$. Then*

$$\mathscr{N}_0 \leq \frac{\operatorname*{Max}_{|z|=\varrho} \log \left| \dfrac{H(z)}{H(0)} \right|}{\log \varrho}. \tag{3.6}$$

It follows from (3.5) and (3.6) that if $h_n^{(n)}(o) \neq 0$, then for any $\varrho > 1$ we have

$$\frac{S_0(n) + S_1(n)}{n} \leq \frac{\operatorname*{Max}_{|z|=\varrho} \log \left| \dfrac{h_n^{(n)}(z)}{h_n^{(n)}(0)} \right|}{n \log \varrho} + (1 + 2\delta'). \tag{3.7}$$

Now let us choose a sequence r_s of positive numbers, such that $\lim_{s \to \infty} r_s = \infty$ and

$$\lim_{s \to \infty} \frac{\mu_g\left(r_s^{\frac{1}{\delta}}\right)}{\mu_f(r_s)} = G < \infty.$$

Such a sequence r_s exists by supposition. Let n_s denote the index of the maximal term of the series $\sum_{n=0}^{\infty} a_n z^n$ for $|z| = r_s$, i. e. suppose

$$|a_{n_s}| r_s^{n_s} \geq |a_m| r_s^m \qquad \text{for} \qquad m = 0, 1, \ldots \tag{3.8}$$

It follows by Lemma 1 that

$$\left| \frac{Q_{n_s}^{(n_s)}(0)}{f^{(n_s)}(0)} \right| \leq \frac{2 \operatorname*{Max}_{t \geq \delta' n_s} |c_t|}{|a_{n_s}|} 12^{n_s} \leq C_1 (12 \lambda_{n_s})^{n_s} \tag{3.9}$$

where $C_1 > 0$ is a constant and

$$\lambda_{n_s} = \frac{1}{r_s^{\frac{\delta' - \delta}{\delta}}}.$$

As clearly

$$\lim_{s\to\infty} \lambda_{n_s} = 0 \tag{3.10}$$

if follows that

$$\lim_{s\to\infty} \left| \frac{Q_{n_s}^{(n_s)}(0)}{f^{(n_s)}(0)} \right| = 0 \tag{3.11}$$

and thus also $h_{n_s}^{(n_s)}(0) \neq 0$ for sufficiently large s.

By Lemma 1 we have also for $|z| = \varrho > 1$

$$\left| \frac{Q_{n_s}^{(n_s)}(z)}{f^{(n_s)}(0)} \right| \leq \frac{2\left(\operatorname*{Max}_{t \geq \delta' n} |c_t|\right) 12^{n_s} \sum_{j=0}^{n_s-1} \binom{n_s+j}{j} \varrho^j}{|a_{n_s}|}.$$

Similarly as in (3.9) we obtain

$$\operatorname*{Max}_{|z|=\varrho} \left| \frac{Q_{n_s}^{(n_s)}(z)}{f^{(n_s)}(0)} \right| \leq C_2 (12\lambda_{n_s})^{n_s} \sum_{j=0}^{n_s-1} \binom{n_s+j}{j} \varrho^j.$$

where $C_2 > 0$ is a constant.

Now if s is sufficiently large, we have $12\lambda_{n_s}\varrho < 1$ and therefore

$$\operatorname*{Max}_{|z|=\varrho} \left| \frac{Q_{n_s}^{(n_s)}(z)}{f^{(n_s)}(0)} \right| \leq \frac{C_2}{(1 - 12\lambda_{n_s}\varrho)^{n_s+1}}.$$

It follows by (3.10) that

$$\lim_{s\to\infty} \frac{\operatorname*{Max}_{|z|=\varrho} \log \left| \frac{Q_{n_s}^{(n_s)}(z)}{f^{(n_s)}(0)} \right|}{n_s} = 0. \tag{3.12}$$

Finally we have for $|z| = \varrho$ by (3.8)

$$\left| \frac{f^{(n_s)}(z)}{f^{(n_s)}(0)} \right| \leq \sum_{l=0}^{\infty} \left| \frac{a_{n_s+l}}{a_{n_s}} \right| \binom{n_s+l}{l} \varrho^l \leq \frac{1}{\left(1 - \frac{\varrho}{r_s}\right)^{n_s+1}}$$

and thus

$$\lim_{s\to\infty} \frac{\operatorname*{Max}_{|z|=\varrho} \log \left| \frac{f^{(n_s)}(z)}{f^{(n_s)}(0)} \right|}{n_s} = 0. \tag{3.13}$$

As we have evidently

$$\operatorname*{Max}_{|z|=\varrho} \left| \frac{h_{n_s}^{(n_s)}(z)}{h_{n_s}^{(n_s)}(0)} \right| \leq \frac{\operatorname*{Max}_{|z|=\varrho} \left| \frac{f^{(n_s)}(z)}{f^{(n_s)}(0)} \right| + \operatorname*{Max}_{|z|=\varrho} \left| \frac{Q_{n_s}^{(n_s)}(z)}{f^{(n_s)}(0)} \right|}{1 - \left| \frac{Q_{n_s}^{(n_s)}(0)}{f^{(n_s)}(0)} \right|} \tag{3.14}$$

comparing (3.11), (3.12), (3.13) and (3.14) it follows that

$$\lim_{s\to\infty} \frac{\operatorname*{Max}_{|z|=\varrho} \log \left| \dfrac{h_{n_s}^{(n_s)}(z)}{h_{n_s}^{(n_s)}(0)} \right|}{n_s} = 0. \tag{3.15}$$

Combining (3.7) with (3.15) we obtain

$$\limsup_{s\to\infty} \frac{S_0(n_s)+S_1(n_s)}{n_s} \leq 1+2\delta' \tag{3.16}$$

and as (3.16) holds for any $\delta'>\delta$, we obtain finally

$$\liminf_{s\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1+2\delta.$$

Thus Theorem 1 is proved.

New we prove the following

THEOREM 2. *Let* $f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n (z-1)^n$ *be a transcendental entire function and* $g(z) = \sum_{n=0}^{\infty} c_n z^n$ *an arbitrary entire function for which the suppositions of Theorem 1 hold. Let* $Z_0(n)$ *denote the number of zeros in the sequence* $a_0, a_1, \ldots, a_{n-1}$ *and* $S_1(n)$ *the number of indices* $k<n$ *for which* $|b_k| \leq |c_k|$. *Then*

$$\liminf_{n\to\infty} \frac{Z_0(n)+S_1(n)}{n} \leq 1+\delta.$$

The proof of Theorem 2 is almost the same as that of Theorem 1; as however we do not use now (3.3), only (3.4), we get δ instead of 2δ.

The following result contains each of Theorem 1 and Theorem 2 as special cases.

THEOREM 3. *Let* $f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n (z-1)^n$ *be a transcendental entire function, further* $g_1(z) = \sum_{n=0}^{\infty} c_n z^n$ *and* $g_2(z) = \sum_{n=0}^{\infty} d_n z^n$ *be entire functions, such that putting*

$$\mu_{g_1}(r) = \operatorname*{Max}_n |c_n| r^n$$

and

$$\mu_{g_2}(r) = \operatorname*{Max}_n |d_n| r^n$$

we have for some δ_1 *and* $\delta_2 (\delta_1>0, \delta_2>0, \delta_1+\delta_2<1)$

$$\liminf_{r\to\infty} \frac{\mu_{g_1}(r^{1/\delta_1})+\mu_{g_2}(r^{1/\delta_2})}{\mu_f(r)} < \infty.$$

Let $S_0(n)$ denote the number of indices $k < n$ for which $|a_k| \leq |c_k|$ and $S_1(n)$ the number of indices $k < n$ for which $|b_k| \leq |c_k|$. Then

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1+\delta_1+\delta_2.$$

The proof of Theorem 3 is essentially the same as that of Theorem 1; the only differences consist in that now we define the polynomial $Q_n(z)$ of order $2n-1$ $(n=1, 2, \ldots)$ as follows:

$$\frac{Q_n^{(j)}(0)}{j!} = \begin{cases} 0 & \text{for } 0 \leq j < \delta_1' n \\ a_j & \text{if } \delta_1' n \leq j \leq n-1 \text{ and } |a_j| \leq |c_j| \\ 0 & \text{if } \delta_1' n \leq j \leq n-1 \text{ and } |a_j| > |c_j| \end{cases}$$

$$\frac{Q_n^{(j)}(1)}{j!} = \begin{cases} 0 & \text{for } 0 \leq j < \delta_2' n \\ b_j & \text{if } \delta_2' \leq j \leq n-1 \text{ and } |b_j| \leq |d_j| \\ 0 & \text{if } \delta_2' n \leq j \leq n-1 \text{ and } |b_j| > |d_j|. \end{cases}$$

Here δ_1 is an arbitrary number for which $\delta_1 < \delta_1' < 1$ and δ_2' an arbitrary number for which $\delta_2 < \delta_2' < 1$.

Evidently if we choose $g_1(z) = g_2(z) = g(z)$, then Theorem 3 reduces to Theorem 1, while if $g_1(z) \equiv 0, g_2(z) = g(z)$, then Theorem 3 reduces to Theorem 2.

Now we shall prove the following Theorem 4, which gives a sharper inequality than (1.4).

THEOREM 4. *If $f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n (z-1)^n$ is an entire function of order $\varrho_f (0 < \varrho_f \leq \infty)$ and $g(z) = \sum_{n=0}^{\infty} c_n z^n$ an entire function of order ϱ_g where $0 \leq \varrho_g < \varrho_f$, and if $S_0(n)$ denotes the number of indices $k < n$ for which $|a_k| \leq |c_k|$ and $S_1(n)$ the number of indices $k < n$ for which $|b_k| \leq |c_k|$ then*

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1 + \frac{\varrho_g}{\varrho_f}.$$

PROOF OF THEOREM 4. Let us put

$$\varphi(z) = \sum_{|a_n| \leq |c_n|} a_n z^n$$

and

$$f^*(z) = f(z) - \varphi(z) = \sum_{|a_n| > |c_n|} a_n z^n = \sum_{n=0}^{\infty} a_n^* z^n = \sum_{n=0}^{\infty} b_n^* (z-1)^n.$$

Then clearly $f^*(z)$ has the same order ϱ_f as $f(z)$ further $\varphi(z)$ is at most of order ϱ_g. Let $Z_0^*(n)$ denote the number of zeros in the sequence $a_0^*, a_1^*, \ldots, a_{n-1}^*$, then clearly

$$Z_0^*(n) = S_0(n) \tag{3.17}$$

We can without restricting the generality suppose that $c_n \geq 0$ as evidently $\sum_{n=0}^{\infty} |c_n| z^n$ is an entire function of the same order as $g(z)$ [This follows e. g. by (1.3)]. Put

$$g^*(z) = 2 \sum_{n=0}^{\infty} z^n \left[\sum_{m=n}^{\infty} \binom{m}{n} c_m \right] = \sum_{n=0}^{\infty} c_n^* z^n.$$

It is easy to see that $g^*(z)$ is also an entire function of order ϱ_g. As a matter of fact, one has

$$g^*(z) = 2g(z+1).$$

Now as

$$f^*(z) = \sum_{n=0}^{\infty} \left[b_n - \sum_{\substack{m=n \\ |a_m| \leq |c_m|}}^{\infty} \binom{m}{n} a_m \right] (z-1)^n$$

we obtain

$$|b_k^*| \leq |b_k| + \sum_{m=k}^{\infty} \binom{m}{k} c_k \qquad (k = 0, 1, \ldots).$$

Thus if $|b_k| \leq |c_k|$ we have $|b_k^*| \leq |c_k^*|$. It follows that if $S_1^*(n)$ denotes the number of indices $k < n$ for which $|b_k^*| \leq |c_k^*|$, then

$$S_1(n) \leq S_1^*(n). \tag{3.18}$$

Now we can apply Theorem 2 to the function $f^*(z)$ of order ϱ_f and the function $g^*(z)$ of order ϱ_g with any $\delta > \frac{\varrho_g}{\varrho_f}$ and thus we obtain, taking (3.17) and (3.18) into account

$$\liminf_{n\to\infty} \frac{S_0(n) + S_1(n)}{n} \leq \liminf_{n\to\infty} \frac{Z_0^*(n) + S_1^*(n)}{n} \leq 1 + \frac{\varrho_g}{\varrho_f}$$

Thus Theorem 4 is proved.

§ 4. An example

We shall discuss now an example, which shows that Theorem 4 is in a certain sense best possible.

Let us consider the function

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{a^n}}{(a^n)!}$$

where $a > 1$ is an integer. Clearly $f(z)$ is an entire transcendental function of order 1. Let us choose for $g(z)$ the function

$$g(z) = \sum_{n=0}^{\infty} \frac{z^n}{(n!)^{1/\vartheta}}$$

where $\frac{1}{a} < \vartheta < 1$. It is easy to see that $g(z)$ is an entire function of order ϑ. We consider now the power series of $f(z)$ around the point $z = 1$; we have

$$f(z) = \sum_{n=0}^{\infty} \frac{[1+(z-1)]^{a^n}}{(a^n)!} = \sum_{m=0}^{\infty} b_m(z-1)^m$$

where

$$b_m = \frac{1}{m!} \sum_{a^n \geq m} \frac{1}{(a^n - m)!}.$$

Let us compare now b_m with

$$c_m = \frac{1}{(m!)^{1/\vartheta}}$$

for a value of m which lies in the interval

$$a^{n-1} \leq m < \frac{a^n \vartheta}{1+\varepsilon} \qquad (n = 1, 2, \ldots)$$

where ε is an arbitrary positive number.

We have clearly for any m satisfying (4.1)

$$b_m \leq \frac{2}{m!\left[\left(\frac{1+\varepsilon}{\vartheta} - 1\right) m\right]!}.$$

It follows by *Stirling*'s formula that if $n \geq n_0$ where n_0 is a suitably chosen large number and if m satisfies (4.1) we have

$$b_m \leq c_m.$$

It follows that

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \geq 1 + \frac{\frac{\vartheta}{1+\varepsilon} - \frac{1}{a}}{1 - \frac{1}{a}}. \tag{4.2}$$

As for ε we can choose an arbitrary small positive number and a can be chosen arbitrarily large, it follows that the lower bound in (4.2) can be arbitrarily near to $1+\vartheta$. As however by Theorem 4

$$\liminf_{n\to\infty} \frac{S_0(n)+S_1(n)}{n} \leq 1+\vartheta \tag{4.3}$$

it follows that Tehorem 4 (and thus also Theorem 2) is best possible in the sense that $1+\vartheta$ can not be replaced in general by any smaller number in (4.3).

Literature

[1] C. Rényi, On a conjecture of G. Pólya, *Acta Math. Acad. Sci. Hung.*, **7** (1956), 145–150.
[2] G. Pólya und G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, Vol. 2., Berlin, Springer 1925.
[3] P. Johansen, Über osculierende Interpolation, *Skandinavisk Aktuarietidskrift*, 1931.
[4] P. Erdős and A. Rényi, On the number of zeros of successive derivatives of analytic functions, *Acta Math. Acad. Sci. Hung.*, **7** (1956), 125–144.

ON EXTREMAL VALUES OF MAPPINGS. I

By

I. JUHÁSZ

Department of Math. Analysis of the Eötvös Loránd University, Budapest

(Received May 22, 1963)

§ 1.

This paper was indicated by a well-known notice of F. HAUSDORFF: the set of the values attained by an arbitrary real-valued function of one variable on its local extremal value places is enumerable ([1], p. 363). In the following we shall generalize this notice, examine its converse cases and we shall deal with other questions of similar type.

For this purpose, first of all we introduce some notations and definitions. Mapping means always mapping of a topological space into a partial ordered set. A point x of this space will be a local extremal value place if it has such a neighbourhood — say U — that for any $y \in U$ we have either $f(y) \geq f(x)$ or $f(y) \leq f(x)$. We call x local minimum or local maximum respectively. We note that the definition implicitely contains that these values of the function are comparable.

We shall say that a topological space has the characteristic $\sigma(k)$ — or shortly that it is a $\sigma(k)$-space — if the set of the values of a mapping attained on its local extremal places has a power $\leq k$ when mapping it to any partial ordered set. (Here k is an infinite cardinal number.)

We denote the topological weight of the topological space R by $Bg(R)$; it is the least infinite cardinal number w for which R has a basis with a power $\leq w$.

We call the least infinite cardinal number s for which R has an everywhere dense subset with a power $\leq s$ the degree of separability of R and we shall denote it by $Sg(R)$. By the local weight of a point x of R we mean — and denote by $Rg(x)$ — the rank of the filter of neighbourhoods of x. Thus that is the least infinite cardinal number r for which x has a fundamental system of neighbourhoods with a power $\leq r$. Finally we shall denote the proximity weight of a proximity space R ([3], p. 175) by $Pg(R)$; that is the least infinite cardinal nomber p for which R has a proximity base ([3], p. 175) of a power $\leq p$.

§ 2.

After these we can state the following generalization of the notice mentioned above:

THEOREM 1. *If R is a topological space and $Bg(R)=k$, then R is a $\sigma(k)$ space.*

PROOF. It is sufficient to examine the set of local minima. Let F be the set of the values of a mapping attained on its local minimum places. Assign to an arbitrary value $f \in F$ such an open set of R, suppose G, on which f is a minimal value. Clearly to two different values f_1 and f_2 different sets G_1 and G_2 will correspond. As we can take these sets G from an arbitrary fixed base of

R, therefore there exists a one-to-one mapping of F into a base with a power $\leq k$, which furnishes the statement.

With this theorem we have given a sufficient condition for a topological space being a $\sigma(k)$-space. In the following – in order to examine the converse of Theorem I – we shall give a necessary condition concerning the same question.

THEOREM II. *If R is a $\sigma(k)$ space, then $Sg(R) \leq k$.*

PROOF. The proof of this theorem is based on the following

LEMMA. If R is a topological space and there is a transfinite sequence of points in R with the power k such that every point of tihs sequence has a neighbourhood which does not contain any previous points of the sequence, then there exists a mapping of R (actually into a well-ordered set) such that the power of the values of this mapping attained on local extremal value places is at least k.

PROOF OF THE LEMMA. Let us order the elements of the transfinite sequence H according to the increasing of their indices. Let us add a new element ξ to the ordered set which is larger than any of the others, denoting the new ordered set by $\{H, \xi\}$. Let us map R on $\{H, \xi\}$ in such a way that the points of H should correspond to themselves and the points of $R \setminus H$ should correspond to ξ. Obviously, the elements of H – as the elements of the well-ordered set $\{H, \xi\}$ – form a set with the power k of values of the mapping attained on local minimum places. By this, the lemma is proved.

The lemma being established, we prove the theorem itself: Our statement is justified – according to the lemma – if we show that from the supposition $Sg(R) > k$ it follows that there is a sequence in R having the property mentioned in the lemma and whith a power larger than k. In fact, assume $Sg(R) > k$. We shall define a well-ordered sequence of points of R by transfinite induction. Put $Sg(R) = \aleph_\alpha$ and let ω_α be the initial number of $\aleph_\alpha$. Let P_0 be an arbitrary point of the sapce R. If P_ν is defined for all $\nu < \mu < \omega_\alpha$ then there exists a $P_\mu \in R$ which has a neighbourhood containing no points P_ν with $\nu < \mu$ because the power of the set of the points already defined is less then $Sg(R)$. As this procedure can be carried out for each $\mu < \omega_\alpha$, therefore we have defined a transfinite sequence S having the property mentioned above and for which $\overline{\overline{S}} = Sg(R) > k$. This completes the proof of Theorem II.

Comparing Theorem II with Theorem I we get immediately:

COROLLARY 1. A pseudometric space R is a $\sigma(k)$ space if and only if $Bg(R) \leq k$.

This follows obviously from the well-known fact, that for every pseudometric space $Bg(R) = Sg(R)$ ([2], p. 77.)

Let us note that we can come to this result also in another way. In fact, let R be a pseudometric space for which $Bg(R) > k$. According to R. H. BING's metrization theorem, R has an $\aleph_0$-discrete base B ([2], p. 191). Then B is union of an enumerable family of discrete systems of sets. Each of these systems is composed of mutually non intersecting sets. But as $\overline{\overline{B}} > k$, there exists at least one among these systems of sets with a power larger than k. However, a space containing more than k mutually non intersecting open sets, obviously cannot be a $\sigma(k)$ space.

It is also clear from this proof that the following converse of the notice of Hausdorff is true: if R is a pseudometric space for which $Bg(R) > \aleph_0$, then there exists a *real function* on R such that the set of the values taken by this function on the local extremal value places is not enumerable. Even, as in the above proof defined system of open sets is discret, it is easy to see, we can define an uniform continuous real function on R of the above mentioned property.

We shall prove another converse of Theorem I – similar to Corollary 1 – concerning completely regular spaces. This follows also from Theorem II though not so directly than Corollary 1.

COROLLARY 2. If R is a completely regular space with the property $\sigma(k)$, then $Bg(R) \leq 2^k$.

PROOF. As it is known ([3], p. 176, 3.), we can introduce into R a proximity structure for which $Pg(R) = Bg(R)$. According to Theorem II, $Sg(R) \leq k$. Let us take an everywhere dense subset S in R with $\overline{\overline{S}} \leq Sg(R)$. As it is known ([3], p. 176, III.), we have for the proximity weight of S $Pg(S) = Pg(R) = Bg(R)$. But it is clear that $Pg(S) \leq 2^{\overline{\overline{S}}} \leq 2^{Sg(R)} \leq 2^k$, by which we proved our statement.

We conjecture that the statement of Corollary 1 is also true for completely regular spaces.

Now we shall prove a proposition which is fairly surprising after the precedings , – that in general, from knowing that a space is a $\sigma(k)$ space, we can not conclude on its topological weight, not even in the case of T_1-spaces. To speak more strictly, we prove the following.

THEOREM III. *Let $m > k$ be two arbitrary infinite cardinal numbers. Then there exist a topological space R which is a T_1-space, a $\sigma(k)$ space, and yet $Bg(R) \geq m$, indeed $Rg(p) \geqq m$ for any $p \in R$.*

PROOF. Cheose an arbitrary set R with $\overline{\overline{R}} = m$. We define a topology in R if we take as open sets, beside the empty set, those subsets of R the complement of which has a power $< k$. This is clearly a T_1-topology. First of all we prove that $Rg(p) \geq m$ for any $p \in R$. Suppose $p \in R$ and $Rg(p) < m$. This would mean that the filter of neighbourhoods of p has a base B for which $\overline{\overline{B}} \leq Rg(p)$. As R is a T_1-space, the intersection of the sets of B contains only the point p, consequently the union of their complements covers $R \setminus \{p\}$. But as the power of any of these complements is less than k, this would imply $\overline{\overline{R \setminus \{p\}}} \leq k \cdot \overline{\overline{B}} < m$, which contradicts our original supposition $\overline{\overline{R}} = m$.

Now we prove that R is a $\sigma(k)$ space. Let us consider an arbitrary mapping of R, and assume that the power of the set of the function values, attained for example on the local minimum places, is larger than k. Select from R a set H of local minimum places for which $\overline{\overline{H}} > k$ and to different points of H correspond different function values. For the sake of simplicity, let us identify the points of H with the corresponding function values. H has the property that for any $p \in H$ p has a neighbourhood in which p is a minimal value, thus the dower of the set of those points of H which are not larger than p is less than k.

Let T be a subset of H with $\overline{\overline{T}} \leq k$. Let us consider those $x \in H$ for which there exists a $y(x) \in T$ not less than x. According to our preceding results, to any $y \in T$ there exist less than k $x \in H$ whith $y \nless x$, so that the power of the set of all such x is at most k. This means that we can select from H one more point which is larger than any element of T.

If now we define a sequence of points of H by transfinite induction in which an arbitrary point is larger than any preceding one, then – if $k = \aleph_\alpha$ – this procedure may not stop before $\omega_{\alpha+1}$. But then the power of the set of points preceding the ω_α-th point p_{ω_α}, is k, which is, however, a contradiction, because all these points are less than $p_{\omega_\alpha} \in H$, by which our statement is proved.

After these, the question naturally arises whether the following statement, sharper than Theorem III, is true: given $m > k \geq \aleph_0$ does there exist a T_1-space R which is a $\sigma(k)$ space and for which $Bg(R) = m$. From the foregoing proof of Theorem III, it is clear, that if we suppose the generalized continuum hypothesis, and if m is either a regular cordinal number, or if $m = \aleph_\alpha$ and $k < \aleph_{cf(\alpha)}$, then $Bg(R) = m$ is true; namely, then the power of the set of all open sets – i. e. all subsets of R, possesing complements with a power less than k – is m.

Finally, let us note the following fact, not belonging to the subject but rather interesting: From the proof of Corollary 2 of Theorem II one can see immediately that if R is a proximity space, then $Pg(R) \leq 2^{Sg(R)}$. Thus if we admit the generalized continuum hypothesis and $Pg(R)$ is a cardinal number with an index of second kind, then $Pg(R) = Bg(R) = Sg(R)$. Similarly, if R is a completely regular space, then $Bg(R) \leq 2^{Sg(R)}$ and admitting the generalized continuum hypothesis, we get $Bg(R) = Sg(R)$, as soon as $Bg(R)$ is a cardinal number with an index of second kind.

*

In composing my paper I had the continual and most efficient help of Professor ÁKOS CSÁSZÁR, on this account I express him my sincerest thanks. I am also thankful to my friends I. DANCS, CS. KISS and G. KORVIN for the English translation of this paper.

Literature

[1] F. HAUSDORFF, *Grundzüge der Mengenlehre* (Leipzig, 1914).
[2] H.-J. KOWALSKY, *Topologische Räume* (Basel-Stuttgart, 1961).
[3] Н. С. РАММ – А. С. ШВАРЦ, Геометрия близости, униформная геометрия и топология, *Мат. Сборник*, **33 (75)** (1953), 157–180.

ON CERTAIN SETS OF INTEGERS

By

I. KÁTAI

Department of Algebra of the Eötvös Loránd University, Budapest

(Received May 20, 1963)

Let $n \geqq 2$ integer, and $c_1, c_2, \ldots, c_{n+1}$ integers, different from zero. A set of positive integers $u_1, u_2, \ldots$ will be called an A set, if the holding of the equality

$$c_1 u_{i_1} + \ldots + c_n u_{i_n} = c_{n+1} u_{i_{n+1}}$$

implies $i_1 = i_2 = \ldots = i_n = i_{n+1}$.

Let $A(x)$ denote the greatest number of integers which can be selected from $1, 2, \ldots, [x]$ to form an A-set. Let be $a(x) = \dfrac{A(x)}{x}$.

THEOREM I. *If* $c_1 + \ldots + c_n \neq c_{n+1}$, *then* $\varliminf_{x \to \infty} a(x) > 0$.

THEOREM II. *If* $c_1 + \ldots + c_n = c_{n+1}$, *then* $\lim_{x \to \infty} a(x) = 0$.

PROOF OF THEOREM I is very simple. If $c_1 + \ldots + c_n \neq c_{n+1}$, then exist m positive integer, that $c_1 + \ldots + c_n \not\equiv c_{n+1} \pmod{m}$. Let $u_i = im + 1$, $i = 1, 2, \ldots$ This is an A-set, and $a(x) \geqq \dfrac{\sum\limits_{im+1 \leq x} 1}{x}$. Thus $\varliminf_{x \to \infty} a(x) \geqq \dfrac{1}{m}$, and hence follows the result.

The Theorem II in the case $n = 2$, $c_1 = c_2 = 1$ was conjectured by P. TURÁN and P. ERDŐS. This conjecture was solved by K. F. ROTH [1]. The method of proof of Theorem II is a simple generalisation of method used in [1].

If $au_1 + b$, $au_2 + b$, ... form an A-set then so do $u_1, u_2, \ldots$ and conversely. Further, if x, y are positive integers, then we have the following inequality

$$A(x+y) \leq A(x) + A(y)$$

$$(1) \qquad A(xy) \leq xA(y), \qquad A(x) \leq A\left(\left\{\left[\frac{x}{y}\right] + 1\right\} \cdot y\right) \leq \frac{x+y}{y} A(y)$$

$$(2) \qquad a(xy) \leq a(y), \qquad a(x) \leq \left(1 + \frac{y}{x}\right) \cdot a(y)$$

Let $u_1, u_2, \ldots, u_u$ be any A-set selected from $1, 2, \ldots, M$. Let $e(\Theta) = e^{2\pi i \Theta}$, and

$$S = \sum_{k=1}^{U} e(\alpha u_k),$$

where α is a real number. For each α there exist h, q such that

$$\alpha = \frac{h}{q} + \beta, \qquad (h, q) = 1, \qquad q \leq M^{1/2}, \qquad q|\beta| \leq M^{-\frac{1}{2}} \tag{3}$$

Let $m < M$, and put $S' = \frac{a(m)}{q}\left[\sum_{T=1}^{q} e\left(\frac{rh}{q}\right)\right] \cdot \sum_{n=1}^{M} e(\beta n)$ (obviously $S' = 0$, if $q > 1$).

K. F. Roth proved, that

$$|S - S'| < Ma(m) - U + O(mM^{1/2}) \tag{4}$$

in the case $n = 2$, $c_1 = c_2 = 1$. The proof in the general case is a similar one.

Let $m \equiv 0 \pmod n$, $nN = m^4$, and let now $u_1, u_2, \ldots, u_u$ be a maximal A-set selected from $1, 2, \ldots, nN$, so that $U = A(nN)$.

We denote all the integers from $u_1, \ldots, u_u$ which are $\equiv 0 \pmod n$ by $nv_1, nv_2, \ldots, nv_v$. Obviously we have

$$U = A(nN) = a(nN) \cdot nN$$

and

$$A(nN) - (n-1)A(N) \leq V \leq A(N).$$

We define the functions $f(\alpha)$, $F(\alpha)$ by

$$f(\alpha) = \sum_{k=1}^{V} e(\alpha v_k), \qquad F(\alpha) = a(m) \cdot \sum_{\nu=1}^{N} e(\alpha\nu). \tag{5}$$

We now show that for any α

$$f(\alpha) - F(\alpha) = O(N\{a(m) - a(nN)\} + N^{3/4}. \tag{6}$$

If $q = 1$ in (3), this follows at once from (4). Let $\|\alpha\|$ denote as usually the distance of α from the nearest integer.

We have obviously

$$\sum_{\nu=1}^{N} e(\alpha\nu) = O(\|\alpha\|^{-1}). \tag{7}$$

In the case $q > 1$ is $\|\alpha\| > M^{-\frac{1}{2}}$ and thus we get $A(\alpha) = O\left(N^{\frac{3}{4}}\right)$ which implies (6).

The Parseval-formula gives immediately

$$V = \int_0^1 f(c_1\alpha) \ldots f(c_n\alpha) f(-c_{n+1}\alpha)\, d\alpha. \tag{8}$$

Using the triangle inequality $n-1$ times we get

$$f(c_1\alpha) \ldots f(c_n\alpha) f(-c_{n+1}\alpha) - F(c_1\alpha) \ldots F(c_n\alpha) F(-c_{n+1}^{\alpha})(-c_{n+1}\alpha) =$$

$$= O\{a^n(m) N^{n+1}[a(m) - a(nN) + N^{-\frac{1}{4}}]\}. \tag{9}$$

Finally, by (6) and (7), if $0 < \eta < \alpha < 1-\eta$, then we have

$$f(\alpha) = O\left(\frac{a(m)}{\eta} + N\{a(m) - a(nN) + N^{-\frac{1}{4}}\}\right). \tag{10}$$

Let $0 < \eta < \dfrac{1}{2 \max\limits_{\nu=1,\dots,n+1} |c_\nu|}$, and put $n+1 = 2\left[\dfrac{n}{2}\right] + l$, where $l=1$ or 2.

Further, let denote $b_1, b_2, \dots$ positive constants.

Using the inequality of the arithmetic and geometric means we have

$$\left|\int_{\eta}^{1-\eta} f(c_1\alpha)\dots f(c_n\alpha) f(-c_{n+1}\alpha)\, d\alpha\right| \leq b_1 \operatorname*{supr}_{\|\alpha\| > b_2\eta} |f(\alpha)|^l \int_{\eta}^{1-\eta} |f(c_1\alpha)\dots f(c_n\alpha)|\, d\alpha \leq$$

$$\leq b_3\left\{\left(\frac{a(m)}{\eta}\right)^l + N^l\left[a(m) - a(nN) + N^{-\frac{1}{4}l}\right]\right\} \cdot [Na(m)]^{\left[\frac{n}{2}\right]}. \tag{11}$$

On other hand we get from (9)

$$\int_{-\eta}^{\eta} f(c_1\alpha)\dots f(-c_{n+1})\, d\alpha =$$

$$= \int_{-\eta}^{\eta} F(c_1\alpha)\dots F(-c_{n+1}\alpha)\, d\alpha + O\left([a(m)N]^n \cdot N \cdot \left\{a(m) - a(nN) + N^{-\frac{1}{4}}\right\}\eta\right). \tag{12}$$

Further using (5) and (7), we have

$$\left|\int_{-\eta}^{\eta} F(c_1\alpha)\dots F(-c_{n+1}\alpha)\, d\alpha - \int_{-\frac{1}{2}}^{\frac{1}{2}} F(c_1\alpha)\dots F(-c_{n+1}\alpha)\, d\alpha\right| \leqq b_4 \cdot a^{n+1}(m)\eta^{-n} \tag{13}$$

and

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(c_1\alpha)\dots F(-c_{n+1}\alpha)\, d\alpha = a^{n+1}(m) \sum_{\substack{c_1X_1+\dots+c_nX_n = c_{n+1}X_{n+1} \\ N \geq X_i > 0,\ \text{integer}}} 1 \geqq$$

$$\geqq b_5 a^{n+1}(m) N^n. \tag{14}$$

Finally summing the results (11), (12), (13), (14) we have

$$b_5 \cdot a^{n+1}(m) \cdot N^n < b_6\left\{b_4 \cdot a^{n+1}(m) \cdot \eta^{-n} + [a(m)N]^n N\eta\left[a(m) - a(nN) + N^{-\frac{1}{4}}\right] + \right.$$

$$\left. + \left(\frac{a(m)}{\eta}\right)^l [Na(m)]^{\left[\frac{n}{2}\right]} + N^{l+\left[\frac{n}{2}\right]} \cdot a^{\left[\frac{n}{2}\right]}(m) \cdot \left[\left(a(m) - a(nN)\right)^l + N^{-\frac{1}{4}l}\right]\right\}.$$

Introducing the notations $\delta = (N\eta)^{-1}$, $\delta = d\cdot a(m)$, since $\frac{1}{x} \leq a(x) \leq 1$ is valid, we have

$$a^l(m)\cdot\delta^l\left(\frac{a(m)}{N}\right)^{\left[\frac{n}{2}\right]}\cdot N \leq d\cdot a^{n+1}(m)$$

and if $d<\frac{b_6}{2b_5}$, then the relation

$$a^{n+1}(m) < b_7\left[a^{n-1}(m)+N^{1-\left[\frac{n}{2}\right]}\cdot a^{\left[\frac{n}{2}\right]}(m)\right]\cdot\left[a(m)-a(nN)+N^{-\frac{1}{4}}\right]$$

holds, whence follows

$$a^{\left[\frac{n}{2}\right]+l}(m) < b_8\left\{a(m)-a(nN)+N^{-\frac{1}{4}}\right\}. \tag{15}$$

Write $m=n^{4^x}$, $b(x) = a(m)$, so we get

$$b^{\left[\frac{n}{2}\right]+l}(x) < b_8\left\{b(x)-b(x+1)+\frac{1}{n^{4^x}}\right\}.$$

Now $b(x)$ is a decraising function, by (2). Let P be a positive integer, so we have

$$Pb^{\left[\frac{n}{2}\right]+l}(2P) \leq \sum_{X=P}^{2P-1}[b(x)]^{\left[\frac{n}{2}\right]+l} \leqq b_9\cdot\left\{b(P)-b(2P)+\frac{1}{P}\right\} \leq b_{10}$$

whence

$$b(x) \leq \frac{b_{11}}{x^{\left[\frac{n}{2}\right]+l}}.$$

If $n^{4^x} < y \leq n^{4^{x+1}}$, then

$$a(y) \leqq \left(1+\frac{n^{4^x}}{y}\right)a(n^{4^x}) \leq \frac{2b_{11}}{x^{\frac{1}{\left[\frac{n}{2}\right]+l}}} \quad \text{is valid.}$$

Hence follows that

$$a(y) = O\left\{\frac{1}{(\log\log y)^{\frac{1}{\left[\frac{n}{2}\right]+l}}}\right\}. \tag{16}$$

Q. E. D.

Consider an integer $n \geqq 2$ and the integers $c_1, c_2, \ldots, c_{n+1}$. Let $m > 0$ be an integer for which the conditions $c_i \not\equiv 0 \pmod{m}$, $i = 1, 2, \ldots, n+1$ hold.

A set $a_1, a_2, \ldots, a_A$ of incongruent residues to the modulus m is called a D-set, if the holding of the congruence

$$c_1 a_{i_1} + \cdots + c_n a_{i_n} \equiv c_{n+1} a_{i_{n+1}} \pmod m \tag{17}$$

implies $i_1 = i_2 = \ldots = i_n = i_{n+1}$.
Let $D(m)$ denote the greatest number of residues mod m to form a D-set, and $d(m) = \dfrac{D(m)}{m}$.

THEOREM III. *If*

$$c_1 + \cdots + c_n \neq c_{n+1} \tag{18}$$

$$\text{then } \varliminf_{m\to\infty} d(m) > 0.$$

THEOREM IV. *If* $c_1 + \ldots + c_n = c_{n+1}$, *then* $\lim\limits_{m\to\infty} d(m) = 0$ where

$$c_1 \not\equiv 0 \pmod m, \ldots, c_{n+1} \not\equiv 0 \pmod m.$$

PROOF OF THEOREM III. In the case (18) there exist a $m_0 > 0$ integer that $c_1 + \ldots + c_n \not\equiv c_{n+1} \pmod{m_0}$ holds. Let m be an integer satisfying the condition $m > 3m_0(|c_1| + \ldots + |c_{n+1}|)$. Consider the set of $a_i = im_0 + 1$, $i = 1, 2, \ldots, K$;

$$K \overset{\text{def}}{=} \left[\frac{m}{3m_0(|c_1| + \cdots + |c_{n+1}|)}\right]. \tag{19}$$

Since the inequality $|c_1 a_{i_1} + \ldots + c_n a_{i_n} - c_{n+1} a_{i_{n+1}}| < m$ is valid, so from the congruence

$$c_1 a_{i_1} + \cdots + c_n a_{i_{n+1}} \equiv c_{n+1} a_{i_{n+1}} \pmod m$$

follows the equality

$$c_1 a_{i_1} + \cdots + c_n a_{i_{n+1}} = c_{n+1} a_{i_{n+1}}.$$

On the other hand the relation $c_1 a_{i_1} + \ldots + c_n a_{i_n} - c_{n+1} a_{i_{n+1}} \equiv c_1 + \ldots + c_n - c_{n+1} \not\equiv$ $\not\equiv O \pmod{m_0}$ holds, and so a_i is a D-set. Obviously $d(m) \geqq \dfrac{1}{4m_0(|c_1| + \ldots |c_{n+1}|)}$ whence follows the result.

PROOF OF THEOREM IV. Let x be an integer, $1 \leq x \leq m$. Let $C(x, m)$ denote the greatest number of residues mod m) that can be selected from the set $1, 2, \ldots, x$ to form an D-set.
Clearly the equation $C(m, m) = D(m)$ is valid. Let x, y denote integers, then we have

$$C(x+y, m) \leqq C(x, m) + C(y, m) \quad \text{whence}$$

$$C(xy, m) \leqq xC(y, m). \tag{20}$$

Let $a_1, \ldots, a_A$ be a maximal D-set selected from

$$1, 2, \ldots, \quad \left[\frac{m}{2(|c_1| + \cdots + |c_{n+1}|)}\right] \overset{\text{def}}{(=} L) \bmod m.$$

Evidently this is an A-set. Applying (17) we get

$$C(L, m) \leq A(L) = O\left(\frac{m}{(\log\log m)^{\frac{1}{\left[\frac{n}{2}\right]+l}}}\right).$$

Substitue $x = L$ and $y = 3(|c_1| + \ldots + |c_{n+1}|)$ so the equality

$$D(m) = O\left(\frac{m}{(\log\log m)^{\frac{1}{\left[\frac{n}{2}\right]+l}}}\right) \quad \text{holds.}$$

Q. E. D.

Literature

[1] K. F. Roth, On certain sets of integers, *Journal London Math. Soc.*, **28** (1953), 104–109.

ÜBER DIE VERTEILUNG VON PUNKTEN AUF DER KUGEL[1]

Von

J. STROHMAJER

Lehrstuhl der Geometrie, Eötvös Loránd Universität, Budapest

(Eingegangen am 30. April 1963)

Der holländische Biologe TAMMES [10] hat folgendes Problem aufgeworfen[2]: *Wie muß man auf einer Kugel $n \geqq 2$ Punkte verteilen, damit der minimale Abstand unter ihnen maximal sei?*

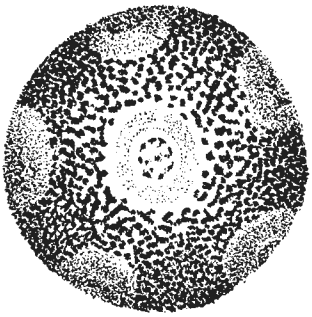


Fig. 1.

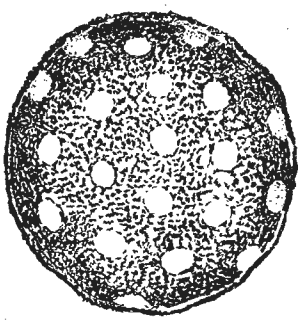


Fig. 2.

Die Lösung des Problemes ist für die Fälle $n \leqq 12$ und $n = 24$ bekannt,[3] in den Fällen $n = 13, 14, 15, 16, 17, 20, 25, 32, 33, 42$, kennt man nur gewisse günstige Konfigurationen.[4]

[1] Diesem Problemkreis hat J. MOLNÁR meine Aufmerksamkeit zugelenkt.

[2] TAMMES hat dieses Problem aufgeworfen, als er die Verteilung von gewissen Öffnungen auf der Oberfläche des kugelförmigen Pollens untersuchte, und beobachtete das diese Öffnungen meistens so verteilt sind, daß sie möglichst weit voneinander seien. Figur 1 und 2 (ERDTMAN [3]) stellen die Pollen Gypsophila fastigata und Suaeda maritima dar. Die in eckige Klammern gesetzte Nummern weisen auf unser Litteraturverzeichnis hin.

[3] Für $n = 2$ ist die Lösung trivial; für $n = 3, 4, 6, 12$ ist die Lösung eine Folge eines Satzes von FEJES TÓTH [4]; für $n = 5$ hat die Lösung selbst TAMMES schon gekannt; die Fälle $n = 7, 8, 9$ haben SCHÜTTE und VAN DER WAERDEN [9] gelöst, die Lösung für $n = 10, 11$ stammt von DANZER [1], [2]; die Vermutung von VAN DER WAERDEN für den Fall $n = 24$ hat ROBINSON [7] bewiesen.

[4] Für $n = 13, 14, 15, 16$ hat SCHÜTTE [8], für $n = 19, 20, 32, 42$, VAN DER WAERDEN [11] und für $n = 17, 25, 33$ JUCOVIČ [5] gute Konfigurationen angegeben. Während der Korrektur habe ich erfahren, daß M. GOLDBERGER für $n = 33$ eine noch günstigere Konfiguration gefunden hat [Siehe M. GOLDBERGER, Packing of 33 equal Circles on a Sphere, *El.Math.*, *B.* **18** (1963), 99–100]. Für $n = 122$ s. VAN DER WAERDEN [11], [12].

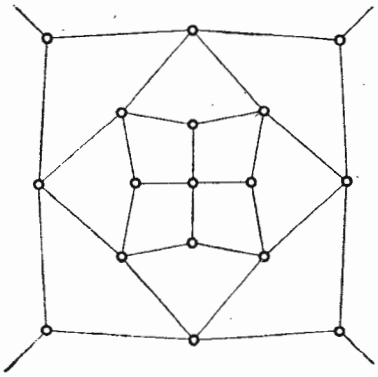


Fig. 3.

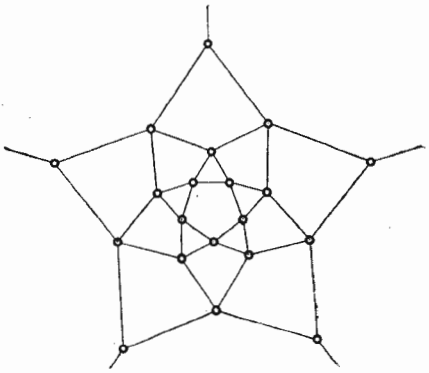


Fig. 4.

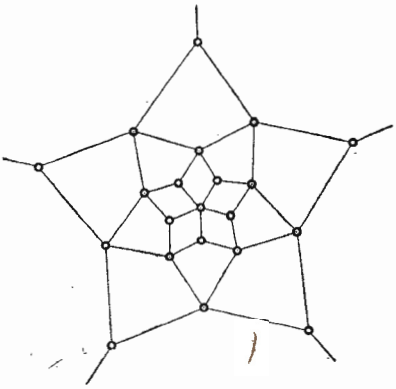


Fig. 5.

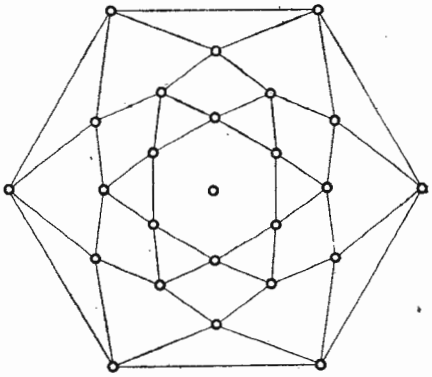
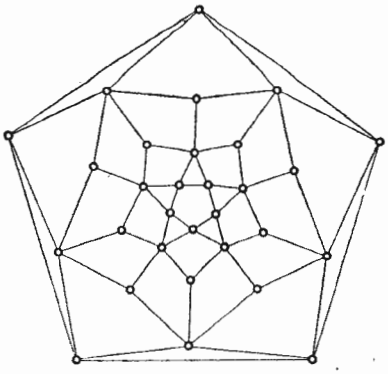
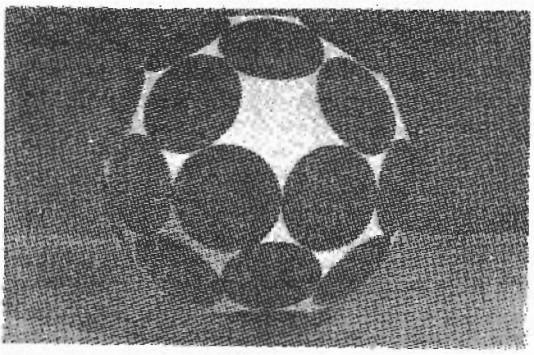


Fig. 6.



a)



b)

Fig. 7.

Die Figuren[5] 3, 4, 5, 6, 7, 8, 9 geben für die Fälle $n=18, 21, 22, 26, 30, 31, 52$ weitere günstige Konfigurationen an. Die Konfigurationen $n=18, 26, 52$ stammen von D. KÓLYA, die für $n=21, 22$ von JUDIT GYŐRFFY, endlich die für $n=31$ von K. BÖRÖCZKY.

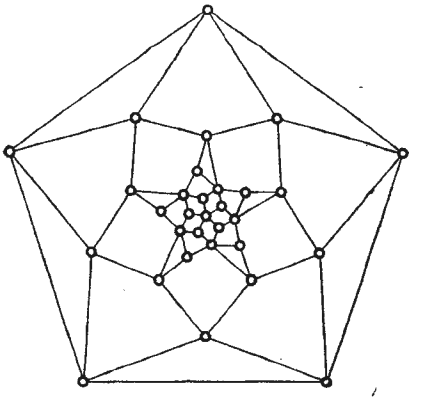


Fig. 8.

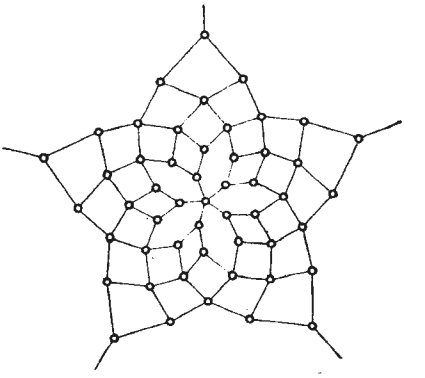


Fig. 9.

In den Folgenden machen wir einige Bemerkungen über die Struktur und Günstigkeit solcher Konfigurationen. Dazu geben wir zuerst unserem Problem eine neue Fassung.

Schlägt man um die Punkte des zum gegebenen n gehörenden extremalen Punktsystems nicht ineinander greifende Kreise maximalen Halbmessers, so erhält man die dichteste Kreislagerung von n Kreisen auf der Kugel.[6]

Unser obiges Problem kann nun folgendermaßen gefaßt werden:

Es ist auf der Kugel die dichteste Lagerung von n kongruenten, nicht ineinander greifenden Kreisen zu bestimmen.

In den erwähnten Konstruktionen für Werte $n \geqq 13$ kommen überall, mit Ausnahme des Falles $n=20$, Kreiskränze vor.[7] Ein Kreiskranz besteht aus $2k$ zyklisch angeordneten Kreisen, unter denen ein jeder die zwei Nachbarkreise berührt, und deren Mittelpunkte abwechselnd auf dem einen und dem anderen von zwei parallelen Kugelkreisen liegen. Auf jedem dieser Kugelkreise bestimmen also die Mittelpunkte der zum Kranz gehörigen Kreise je ein reguläres n-Eck (Fig. 10).

Um die Leistung unserer Kreiskonstruktionen schätzen zu können, vergleichen wir ihre Dichte D_n mit der durch ROBINSON [7] angegebenen oberen

[5] Die Figuren – mit Ausnahme von Fig. 7b – stellen das Schema der stereographischen Projektionen der Punkte dar. Es sind jene Punktpare verbunden deren sphärischer Abstand minimal ist.

[6] Unter der Dichte nicht ineinander greifender Kreise verstehen wir das Verhältnis ihrer Inhaltsumme zur gesamten Kugeloberfläche.

[7] Siehe J. MOLNÁR [6].

Schranke Δ_n.[8] Die Strukturen der Konfigurationen und die numerischen Resultate enthält folgende Tabelle.[9]

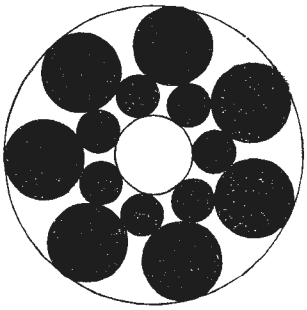


Fig. 10.

n	Struktur der Konfiguration	r_n	D_n	Δ_n
18	1, 4-4-4-4, 1	24° 46′	0,828	0,8659
21	1, 5-5-5-5	22° 28′ 30″	0,798	0,8625
22	1, 5-5-5-5, 1	22° 12′	0,815	0,8620
26	1, 6-6-6-6, 1	20° 22′	0,813	0,8620
30	5-5-5, 5-5-5	19° 4′ 2″	0,822	0,8635
31	1, 5-5, 5, 5-5-5	18° 50′	0,830	0,8639
52	1, 5, 10-10-10-10, 5, 1	14° 23′	0,813	0,8741

Figur 11 stellt die Werte D_n und die Funktion Δ_n dar.[10]

[8] Die Robinsonsche Schranke bezieht sich auf die Werte $n \geqq 12$.

[9] r_n bedeutet den zu D_n gehörenden Radiuswert. Ein Kreiskranz von $2k$ Kreisen wird in der Tabelle durch $k-k$ bezeichnet.

[10] Für $n = 19, 23, 27, 28, 29$ stehen in unserer Figur jenen Konvfigunrationen entommene Werte, die aus den für die nächstgrößeren Werte bekannten Konfigurationen durch Weglassen eines Kreises bzw. einiger Kreise entstehen.

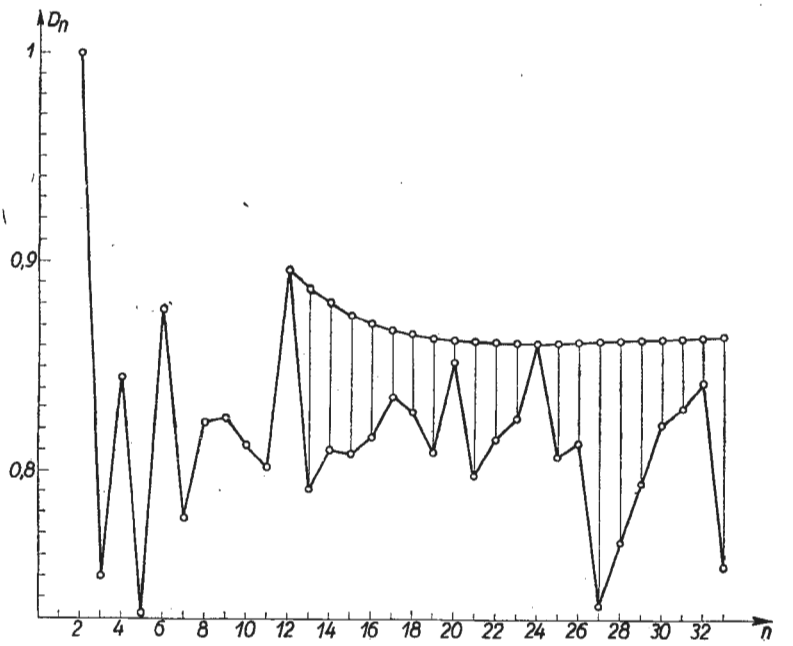


Fig. 11.

Literaturverzeichnis

[1] L. DANZER, Verteilung von elf Punkten auf der Kugel. In Oberwolfach am 25. Okt. 1958. gehaltener Vortrag.

[2] L. DANZER, Über optimale Lagerungen von 7 bis 11 kongruenten Kreisen auf der 2-Sphäre, Oberwolfach, Kolloquium der diskreten Geometrie, 23–29 Juli. 1962.

[3] G. ERDTMAN, B. BERGLUND, and J. PRAGLOWSKI, An introduction to a scandinavian pollen flora, *Grana Palynologica II* (1961), 2–92, Pl. 1–Pl 74.

[4] L. FEJES TÓTH, Über die Abschätzung des kürzesten Abstand zweier Punkte eines auf einer Kugelfläche liegenden Punktsystems, *Jber. d. Deutschen Math.–Ver.*, **53** (1943), 66–68.

[5] F. JUCOVIČ, Umiestenie 17, 25 a 33 bodov na guli, *Mat.–Fiz. Časopis*, **9** (1959), 173–176.

[6] J. MOLNÁR, Körök elhelyezése gömbön, *Mat. Lapok*, **4** (1953), 113–123.

[7] R. M. ROBINSON, Arrangement of 24 points on a sphere, *Math. Ann.*, **144** (1961), 17–48.

[8] K. SCHÜTTE, Überdeckung der Kugel mit höchstens acht Kreisen, *Math. Ann.*, **129** (1955), 181–186.

[9] K. SCHÜTTE, und B. L. VAN DEN WAERDEN, Auf welcher Kugel haben 5, 6, 7, 8, oder 9 Punkte mit Mindestabstand 1 Platz? *Math. Ann.*, **123** (1951), 96–124.

[10] P. M. L. TAMMES, On the origin of number and arrangement of the places of exit on the surface of pollengrains, *Rec. Trav. bot. Néerl.*, **27** (1930), 1–84.

[11] B. L. VAN DER WAERDEN, Punkte auf der Kugel, *Math. Ann.*, **125** (1952), 213–222.

[12] B. L. VAN DER WAERDEN, Berichtigung zu der Arbeit B. L. van der Waerden in Zürich: Punkte auf der Kugel. Drei Zusätze, Math. Ann., 125 (1952), 213–222., *Math. Ann.*, **152** (1963), 94.

TRANSPOSITION DE STRUCTURES SYNTOPOGÈNES

Par

ÁKOS CSÁSZÁR

Ie Chaire d'Analyse Mathématique, Université Eötvös Loránd, Budapest

Dédié à M. B. Sz.-Nagy à l'occasion de son 50e anniversaire

(Reçu le 21 mai 1963)

Il existe, dans la topologie générale, plusieurs méthodes qui servent à transposer d'un espace à un autre une topologie, une structure uniforme ou, plus généralement, une structure syntopogène. Telles sont par exemple la construction de l'image réciproque (d'une topologie, d'une structure uniforme, d'une structure syntopogène), la construction de la topologie quotient, la construction de l'image d'une structure syntopogène, la construction de la complétion d'une structure uniforme, la compactification d'une topologie, etc.

Le but de cette note est d'étudier une méthode générale de transposition de structures syntopogènes, méthode qui contient comme cas particuliers les constructions énumérées et encore d'autres. Pour la terminologie et la notation, nous renvoyons le lecteur à l'édition allemande [1] de notre traité sur les fondements de la topologie générale.

Je remercie M. J. Czipszer* pour des remarques de grande valeur.

1. *Relations prétopogènes.* Considérons un ensemble E et une relation binaire $\prec$, définie sur l'ensemble 2^E des parties de E et moins fine que la relation d'inclusion (c'est-à-dire telle que $A \prec B$ implique $A \subset B$). Une relation de cette sorte sera appelée *relation prétopogène* sur l'ensemble E.

Le plus simple exemple d'une relation prétopogène est celui d'une relation $\prec$ qui ne subsiste qu'entre un seul couple d'ensembles $A \subset B$. On sait que, $A \subset B \subset E$ étant donnés, l'*ordre topogène élémentaire* $<_{A,B}$ est défini comme suit:

(1.1) $X <_{A,B} Y$ *si et seulement si* $X=0$ *ou* $Y=E$ *ou* $X \subset A \subset B \subset Y$.[1]

De plus, $<_{A,B}$ est un ordre topogène biparfait. On observe aisément que $<_{A,B}$ est le moins fin des ordres semi-topogènes qui sont plus fins que la relation $\prec$ ne subsistant qu'entre A et B.

De façon plus générale, on peut affirmer:

(1.2) *Pour une relation prétopogène quelconque* $\prec$ *sur* E, *il existe un ordre semi-topogène* $<$, *désigné par* $\prec^e$, *qui est le moins fin de tous les ordres semi-topogènes plus fins que la relation* $\prec$, *à savoir*

$$< \; = \; \prec^e = \bigcup \{<_{A,B} : A \prec B\} \tag{1.3}$$

lorsqu'il existe au moins un couple A, B *tel que* $A \prec B$, *et* $< \; = \; \prec^e = <_{0,E}$ *si la relation* $\prec$ *est vide* (c'est-à-dire si elle ne subsiste pour aucun couple A, B).

Si $<$ est un ordre semi-topogène sur E et $\mathfrak{D}$ désigne une famille de parties

* Défunt tragiquement le 15 juin 1963.

[1] [1], p. 56, formule (5.50).

de E, on obtient une relation prétopogène en posant $A \prec B$ si et seulement si $A < B$ et $A \in \mathfrak{D}$, $B \in \mathfrak{D}$; autrement dit, la restriction sur $\mathfrak{D} \times \mathfrak{D}$ de la relation $<$ constitue une relation prétopogène $\prec = < | \mathfrak{D} \times \mathfrak{D}$. Pour cette relation prétopogène, on a évidemment

$$(1.4) \qquad (< | \mathfrak{D} \times \mathfrak{D})^e \subset \prec;$$

le signe d'inclusion peut être remplacé ici par le signe d'égalité si

(1.5) *$A < B$ implique l'existence de C, D tels que*

$$A \subset C < D \subset B, \qquad C \in \mathfrak{D}, \qquad D \in \mathfrak{D},$$

et réciproquement l'égalité $(< | \mathfrak{D} \times \mathfrak{D})^e = <$ implique (1.5) pour $0 \neq A < B \neq E$. Si la condition (1.5) est remplie, nous disons que la famille $\mathfrak{D}$ est *dense* pour l'ordre semi-topogène $<$. Les formules $0 < 0$ et $E < E$ entraînent que, si $\mathfrak{D}$ est dense pour $<$, on a $0 \in \mathfrak{D}$ et $E \in \mathfrak{D}$.

2. *Transposition d'ordres semi-topogènes.* Considérons deux ensembles E et E' et une application h d'une famille $\mathfrak{D} \subset 2^E$ dans $2^{E'}$, *monotone* en ce sens que

$$(2.1) \qquad X \subset Y, \ X \in \mathfrak{D}, \ Y \in \mathfrak{D} \quad \textit{impliquent} \quad h(X) \subset h(Y).$$

Si $\prec$ est une relation prétopogène sur E, on définit sur E' une relation prétopogène $\prec'$ en posant

(2.2) *$A' \prec' B'$ si et seulement si il existe $A, B \in \mathfrak{D}$ tels que*

$$A' = h(A), \qquad B' = h(B), \qquad A \prec B.$$

Ceci établi, soit donné un ordre semi-topogène $<$ sur E, désignons par $\prec$ la restriction $< | \mathfrak{D} \times \mathfrak{D}$, puis par $\prec'$ la relation prétopogène définie à partir de $\prec = < | \mathfrak{D} \times \mathfrak{D}$ au moyen de (2.2), et posons enfin $<^h = \prec'^e$, c'est-à-dire

$$(2.3) \qquad <^h = \bigcup \{ <_{h(A), h(B)} : A \in \mathfrak{D}, \ B \in \mathfrak{D}, \ A < B \},$$

ou $<^h = <_{0,E'}$ si la famille $\{\ldots\}$ est vide.

L'ordre semi-topogène $<^h$ sur E' est appelé le *transposé* de $<$ par l'application h.

Dans ce qui suit, nous supposerons toujours que h est une application monotone [c'est-à-dire satisfaisant à (2.1)] de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$. Nous admettrons parfois que h est *bimonotone*, c'est-à-dire qu'elle est monotone et qu'en même temps

$$(2.4) \qquad X, \ Y \in \mathfrak{D}, \ h(X) \subset h(Y) \quad \textit{impliquent} \quad X \subset Y.$$

Une application bimonotone est évidemment biunivoque.

(2.5) *$<_1$ et $<_2$ étant deux ordres semi-topogènes sur E, $<_1 \subset <_2$ entraîne $<_1^h \subset <_2^h$; si h est bimonotone et si $\mathfrak{D}$ est dense pour $<_1$, alors $<_1^h \subset <_2^h$ entraîne $<_1 \subset <_2$.*

Démonstration. La première assertion étant évidente, supposons que h est bimonotone, $\mathfrak{D}$ est dense pour $<_1$ et qu'on a $<_1^h \subset <_2^h$. Posons $A <_1 B$; il existe $C, D \in \mathfrak{D}$ tels que $A \subset C <_1 D \subset B$, donc $h(C) <_1^h h(D)$, par conséquent $h(C) <_2^h h(D)$. Ainsi, soit il existe C_1, $D_1 \in \mathfrak{D}$ tels que $h(C) \subset h(C_1) \subset h(D_1) \subset h(D)$

et $C_1 <_2 D_1$, soit $h(C)=0$, soit $h(D)=E'$. Dans le premier cas, il vient $A \subset C \subset \subset C_1 <_2 D_1 \subset D \subset B$ et $A <_2 B$. Si $h(C)=0$, on a $h(0)=0$ parce que, $\mathfrak{D}$ étant dense pour $<_1$, $0 \in \mathfrak{D}$ et h est monotone, d'où résulte $C=0$ puisque h est bimonotone, et finalement $A=0$, $A <_2 B$. On raisonne similairement si $h(D)=E'$.

(2.6) $\{<_\gamma : \gamma \in \Gamma \neq 0\}$ *étant une famille d'ordres semi-topogènes sur E, on a*

$$(\bigcup_{\gamma \in \Gamma} <_\gamma)^h = \bigcup_{\gamma \in \Gamma} <_\gamma^h.$$

En effet, si

$$A'(\bigcup <_\gamma)^h B',$$

on a pour $A' \neq 0$, $B' \neq E'$

$$A' \subset h(A) \subset h(B) \subset B', \qquad A(\bigcup <_\gamma)B, \qquad A,\ B \in \mathfrak{D},$$

donc $A <_\gamma B$ pour au moins un $\gamma \in \Gamma$. Il en résulte $A' <_\gamma^h B'$ et

$$A'(\bigcup <_\gamma^h)B'.$$

Cette formule subsiste évidemment pour $A'=0$ ou $B'=E'$. Par conséquent,

$$(\bigcup <_\gamma)^h \subset \bigcup <_\gamma^h,$$

et (2.5) implique

$$\bigcup <_\gamma^h \subset (\bigcup <_\gamma)^h.$$

En ce qui concerne la proposition suivante, remarquons que si $<_1$ et $<_2$ sont deux ordres semi-topogènes sur E, le produit $<_1 <_2$ de ces deux relations est également un ordre semi-topogène sur E.[2]

(2.7) *Soient* $<_1$ *et* $<_2$ *deux ordres semi-topogènes sur E. Si* $\mathfrak{D}$ *est dense pour* $<_1$ *ou* $<_2$, *alors*

(2.8) $$(<_1 <_2)^h \subset <_1^h <_2^h.$$

Si h est bimonotone, alors

(2.9) $$<_1^h <_2^h \subset (<_1 <_2)^h.$$

DÉMONSTRATION. Admettons que $\mathfrak{D}$ est dense pour $<_1$ et posons $A'(<_1 <_2)^h B'$. On a alors pour $A' \neq 0$, $B' \neq E'$

$$A' \subset h(A) \subset h(B) \subset B', \qquad A <_1 <_2 B, \qquad A,\ B \in \mathfrak{D},$$

c'est-à-dire

$$A <_1 C <_2 B$$

pour un ensemble C convenable. Il existe par hypothèse C_1, $D_1 \in \mathfrak{D}$ tels que $A \subset C_1 <_1 D_1 \subset C$, donc

$$A' \subset h(A) \subset h(C_1) \subset h(D_1), \qquad C_1 <_1 D_1$$

et par suite $A' <_1^h h(D_1)$. D'autre part, on a

$$h(D_1) \subset h(C) \subset h(B) \subset B', \qquad C <_2 B,$$

[2] Cf. [1], p. 28, proposition (2.16) pour le cas $<_1 = <_2$.

donc $h(D_1) <_2^h B'$ et

$$A' <_1^h <_2^h B'. \tag{2.10}$$

Cette formule est évidemment valable pour $A'=0$ ou $B'=E'$. On raisonne de la même façon si $\mathfrak{D}$ est dense pour $<_2$.

Réciproquement, supposons que h est bimonotone et admettons (2.10). On a

$$A' <_1^h C' <_2^h B',$$

donc pour $A' \neq 0$, $B' \neq E'$

$$A' \subset h(A) \subset h(C_1) \subset C' \subset h(C_2) \subset h(B) \subset B',$$

$$A <_1 C_1, \qquad C_2 <_2 B, \qquad A,\ C_1,\ C_2,\ B \in \mathfrak{D}.$$

De plus, $h(C_1) \subset h(C_2)$ implique $C_1 \subset C_2$, par conséquent $C_1 <_2 B$ et $A <_1 <_2 B$, d'où vient

$$A'(<_1 <_2)^h B',$$

valable aussi pour $A'=0$ ou $B'=E'$.

Dans ce qui suit, nous examinerons le comportement de l'opération $<^h$ vis-à-vis des opérations $^c, ^q, ^p, ^b$ et ^{s}.[3]

(2.11) *h étant une application monotone de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$, posons $\mathfrak{D}^* = \{X : E-X \in \mathfrak{D}\}$ et $h^*(X) = E' - h(E-X)$ pour $X \in \mathfrak{D}^*$. Alors $<^{ch^*} = <^{hc}$ pour un ordre semi-topogène quelconque sur E. Si $\mathfrak{D}$ et h satisfont à la conditoin suivante:*

$$X \in \mathfrak{D} \text{ implique } E-X \in \mathfrak{D} \text{ et } h(E-X) = E' - h(X), \tag{2.12}$$

alors $<^{ch} = <^{hc}$ et, si en outre $<$ est symétrique, $<^h$ l'est aussi.

En effet, $A' <^{hc} B'$ $(A' \neq 0,\ B' \neq E')$ équivaut à

$$E' - B' \subset h(A) \subset h(B) \subset E' - A', \qquad A,\ B \in \mathfrak{D}, \qquad A < B,$$

donc à

$$A' \subset h^*(E-B) \subset h^*(E-A) \subset B', \qquad E-B,\ E-A \in \mathfrak{D}^*,\ E-B <^c E-A$$

c'est-à-dire à $A' <^{ch^*} B'$. Si (2.12) est rempli, on a $\mathfrak{D}^* = \mathfrak{D}$ et $h^* = h$, d'où le reste de l'énoncé.

(2.13) *Si $\mathfrak{D}$ et h satisfont à la condition suivante:*

$$X, Y \in \mathfrak{D} \text{ implique } X \cup Y \in \mathfrak{D},\ X \cap Y \in \mathfrak{D} \text{ et} \tag{2.14}$$

$$h(X \cup Y) = h(X) \cup h(Y), \qquad h(X \cap Y) = h(X) \cap h(Y),$$

alors $<^{hq} \subset <^{qh}$ pour un ordre semi-topogène quelconque et, si $<$ est un ordre topogène $<^h$ l'est aussi. Si en outre $\mathfrak{D}$ est dense pour $<$, alors $<^{hq} = <^{qh}$.

Démonstration. La formule $<^{hq} \subset <^{qh}$ étant évidente si $<^h = <_{0,E'}$, on peut supposer que $<^h$ est déterminé par (2.3). Or, on a l'égalité

$$(\bigcup\{<_{h(A),\,h(B)} : A \in \mathfrak{D},\ B \in \mathfrak{D},\ A < B\})^q = \tag{2.15}$$

$$= \bigcup\{<_{C',\,D'} : (C', D') \in \mathfrak{C}'\}$$

[3] V. [1], pp. 27, 31, 41, 49 et 36.

où $\mathfrak{C}'$ désigne la famille des couples d'ensembles ayant la forme

$$C' = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} h(A_{ij}), \qquad D' = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} h(B_{ij}), \tag{2.16}$$

$$A_{ij} \in \mathfrak{D}, \quad B_{ij} \in \mathfrak{D}, \quad A_{ij} < B_{ij}, \quad m = 1, 2, \ldots, \quad n_i = 1, 2, \ldots$$ [4]

En vertu de (2.14), on peut écrire au lieu de (2.16)

$$C' = h\left(\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij}\right), \qquad D' = h\left(\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} B_{ij}\right)$$

où

$$\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij} \in \mathfrak{D}, \quad \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} B_{ij} \in \mathfrak{D},$$

$$\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij} <^q \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} B_{ij},$$

donc

$$<^{hq} = \bigcup \{<_{C', D'} : (C', D') \in \mathfrak{C}'\} \subset$$
$$\subset \bigcup \{<_{h(C), h(D)} : C \in \mathfrak{D}, D \in \mathfrak{D}, C <^q D\} = <^{qh}.$$

Si $< = <^q$, on a par suite $<^{hq} \subset <^h$ et $<^{hq} = <^h$.

Supposons maintenant que $\mathfrak{D}$ est dense pour $<$. On a alors

$$<^{qh} = \bigcup \{<_{h(C), h(D)} : C \in \mathfrak{D}, \ D \in \mathfrak{D}, \ C <^q D\} \tag{2.17}$$

et $C <^q D$ entraîne, d'autre part, l'existence de m, n_i, C_{ij} et D_{ij} tels que

$$C = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} C_{ij}, \qquad D = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} D_{ij}, \qquad C_{ij} < D_{ij}.$$

Il existe donc des ensembles $A_{ij} \in \mathfrak{D}$, $B_{ij} \in \mathfrak{D}$ vérifiant la formule

$$C_{ij} \subset A_{ij} < B_{ij} \subset D_{ij},$$

par conséquent

$$h(C) \subset h\left(\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} A_{ij}\right) = \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} h(A_{ij}) \subset$$

$$\subset \bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} h(B_{ij}) = h\left(\bigcup_{i=1}^{m} \bigcap_{j=1}^{n_i} B_{ij}\right) \subset h(D).$$

Eu égard à ce que $h(A_{ij}) <^h h(B_{ij})$, il vient

$$h(C) <^{hq} h(D), \quad <_{h(C), h(D)} \subset <^{hq},$$

et finalement $<^{qh} \subset <^{hq}$ d'après (2.17), c. q. f. d.

[4] V. [1], p. 57, proposition (5.54)

(2.18) *Si* $\mathfrak{D}$ *et* h *satisfont à la condition suivante:*

(2.19) $$X_i \in \mathfrak{D} \ (i \in I) \text{ implique } \bigcup_{i \in I} X_i \in \mathfrak{D} \text{ et } h(\bigcup_{i \in I} X_i) = \bigcup_{i \in I} h(X_i),$$

alors $<^{hp} \subset <^{ph}$ *est valable pour un ordre semi-topogène quelconque et, si* $<$ *est parfait,* $<^h$ *l'est aussi. Si en outre* $\mathfrak{D}$ *est dense pour* $<$, *alors* $<^{hp} = <^{ph}$.

(2.20) *Si* $\mathfrak{D}$ *et* h *satisfont à la condition suivante:*

(2.21) $$X_i \in \mathfrak{D} \ (i \in I) \text{ implique } \bigcup_{i \in I} X_i \in \mathfrak{D}, \ \bigcap_{i \in I} X_i \in \mathfrak{D} \text{ et}$$

$$h(\bigcup_{i \in I} X_i) = \bigcup_{i \in I} h(X_i), \qquad h(\bigcap_{i \in I} X_i) = \bigcap_{i \in I} h(X_i),$$

alors $<^{hb} \subset <^{bh}$ *est vrai pour un ordre semi-topogène quelconque et, si* $<$ *est biparfait,* $<^h$ *l'est aussi. Si de plus* $\mathfrak{D}$ *est dense pour* $<$, *on a* $<^{hb} = <^{bh}$.

On démontre (2.18) et (2.20) tout comme (2.13), seulement au lieu de l'égalité (2.15) on a besoin des formules

(2.22) $$(\mathbf{U}\{<_{h(A),h(B)} : A \in \mathfrak{D},\ B \in \mathfrak{D},\ A < B\})^p = \\ = \mathbf{U}\{<_{C',D'} : (C', D') \in \mathfrak{C}''\}$$

où $\mathfrak{C}''$ désigne la famille des couples

$$C' = \bigcup_{i \in I} h(A_i), \qquad D' = \bigcup_{i \in I} h(B_i),$$

$$A_i \in \mathfrak{D}, \quad B_i \in \mathfrak{D}, \quad A_i < B_i$$

et

(2.23) $$(\mathbf{U}\{<_{h(A),h(B)} : A \in \mathfrak{D},\ B \in \mathfrak{D},\ A < B\})^b = \\ = \mathbf{U}\{<_{C',D'} : (C', D') \in \mathfrak{C}'''\}$$

où $\mathfrak{C}'''$ désigne la famille des couples

$$C' = \bigcup_{i \in I} \bigcap_{j \in J_i} A_{ij}, \qquad D' = \bigcup_{i \in I} \bigcap_{j \in J_i} B_{ij},$$

$$A_{ij} \in \mathfrak{D}, \quad B_{ij} \in \mathfrak{D}, \quad A_{ij} < B_{ij},$$

les ensembles d'indices I et J_i étant arbitraires. Or, (2.22) s'ensuit de la formule, analogue à celle citée sous [4],

(2.24) $$(\mathbf{U}\{<_{A,B} : (A,B) \in \mathfrak{A}\})^p = \mathbf{U}\{<_{C,D} : (C,D) \in \mathfrak{B}\}$$

où $\mathfrak{B}$ désigne la famille des couples ayant la forme

$$C = \bigcup_{i \in I} A_i, \quad D = \bigcup_{i \in I} B_i, \quad (A_i, B_i) \in \mathfrak{A},$$

et (2.23) résulte de l'égalité

(2.25) $$(\mathbf{U}\{<_{A,B} : (A,B) \in \mathfrak{A}\})^b = \mathbf{U}\{<_{C,D} : (C,D) \in \mathfrak{C}\}$$

où $\mathfrak{C}$ est identique à la famille des couples

$$C = \bigcup_{i \in I} \bigcap_{j \in J_i} A_{ij}, \qquad D = \bigcup_{i \in I} \bigcap_{j \in J_i} B_{ij}, \qquad (A_{ij}, B_{ij}) \in \mathfrak{A},$$

ces égalités étant valables pour des ordres topogènes élémentaires sur un ensemble E quelconque.

On vérifie (2.25) de la même façon que l'égalité (5.56) de [1]. Quant à (2.24), en voici la démonstration.

Posons

$$< = \bigcup \{<_{A,B} : (A,B) \in \mathfrak{A}\}$$

et

$$<' = \bigcup \{<_{C,D} : (C,D) \in \mathfrak{B}\}.$$

$X <^p Y$ implique $X = \bigcup_{i \in I} X_i$, $X_i < Y$ pour $i \in I$, donc

$$X_i \subset A_i \subset B_i \subset Y, \qquad (A_i, B_i) \in \mathfrak{A}$$

pour $i \in I$, $X_i \neq 0$, $Y \neq E$. Les indices i tels que $X_i = 0$ peuvent être évidemment supprimés, si $X \neq 0$, admettons donc $X \neq 0$, $Y \neq E$, $X_i \neq 0$ pour $i \in I$. On a alors

$$X \subset \bigcup_{i \in I} A_i \subset \bigcup_{i \in I} B_i \subset Y$$

et en posant

$$C = \bigcup_{i \in I} A_i, \qquad D = \bigcup_{i \in I} B_i$$

il vient

$$X <_{C,D} Y, \qquad (C,D) \in \mathfrak{B}.$$

Il en résulte $X <' Y$ pour $X \neq 0$, $Y \neq E$. Cette formule étant évidemment valable pour $X = 0$ ou $Y = E$, on a donc $<^p \subset <'$. Réciproquement, si $X <' Y$, $X \neq 0$, $Y \neq E$, on a

$$X \subset C = \bigcup_{i \in I} A_i \subset \bigcup_{i \in I} B_i = D \subset Y, \quad (A_i, B_i) \in \mathfrak{A} \quad \text{pour} \quad i \in I.$$

Vu que $A_i < B_i$, on a donc $X <^p Y$ ce qui reste valable si $X = 0$ ou $Y = E$. Par suite $<' \subset <^p$, c. q. f. d.

(2.26) *Si $\mathfrak{D}$ et h satisfont aux conditions* (2.12) *et* (2.14), *on a $<^{hs} \subset <^{sh}$ et $< = <^s$ entraîne $<^h = <^{hs}$. Si en outre $\mathfrak{D}$ est dense pour $<$ (ou pour $< \cup <^c$), alors $<^{hs} = <^{sh}$.*

En effet, on a en conséquence de (2.12) et (2.14)

$$(2.27) \qquad <^{hs} = (<^h \cup <^{hc})^q = (<^h \cup <^{ch})^q =$$
$$= (< \cup <^c)^{hq} \subset (< \cup <^c)^{qh} = <^{sh}$$

conformément à (2.11), (2.6) et (2.13). Si de plus $< = <^s$, il en résulte $<^{hs} \subset <^{sh} = <^h$, donc $<^h = <^{hs}$. Si $\mathfrak{D}$ est dense pour $< \cup <^c$, l'égalité est valable au lieu de la relation $\subset$ dans (2.27) d'après (2.13). Enfin si $\mathfrak{D}$ est dense pour $<$, il l'est pour $<^c$ aussi par suite de (2.12): $A <^c B$ implique $E - B < E - A$, donc l'existence de $C, D \in \mathfrak{D}$ tels que $E - B \subset C \subset D \subset E - A$ et par conséquent

$$A \subset E - D <^c E - C \subset B, \qquad E - D \in \mathfrak{D}, \qquad E - C \in \mathfrak{D}.$$

Donc, dans ce cas, $\mathfrak{D}$ est dense pour $< \cup <^c$. Ceci achève la démonstration.

En ce qui concerne la dépendance du transposé $<^h$ du domaine de définition $\mathfrak{D}$ de h, on peut affirmer:

(2.28) *Soient h une application monotone de $\mathfrak{D}\subset 2^E$ dans $2^{E'}$, $\mathfrak{D}_1\subset\mathfrak{D}$, $\mathfrak{D}_2\subset\mathfrak{D}$, $h_1=h|\mathfrak{D}_1$, $h_2=h|\mathfrak{D}_2$* et $<$ *un ordre semi-topogène sur E. Si*

(2.29) *A, $B\in\mathfrak{D}_1$, $A<B$ impliquent l'existence de C, $D\in\mathfrak{D}_2$ tels que $A\subset C<D\subset B$, alors $<^{h_1}\subset<^{h_2}$.*

En effet, $A'<^{h_1}B'$, $A'\neq 0$, $B'\neq E'$ entraînent

$$A'\subset h_1(A)\subset h_1(B)\subset B', \qquad A, B\in\mathfrak{D}_1, \qquad A<B,$$

donc $A\subset C<D\subset B$ pour C, $D\in\mathfrak{D}_2$ convenables, par conséquent

$$A'\subset h_1(A)=h(A)\subset h(C)=h_2(C)\subset h_2(D)=h(D)\subset h(B)=h_1(B)\subset B',$$

c'est-à-dire $A'<^{h_2}B'$.

La condition (2.29) est vérifiée en particulier si $\mathfrak{D}_1\subset\mathfrak{D}_2$.

La proposition suivante étudie la composition de deux applications:

(2.30) *Soient h et h' des applications monotones de $\mathfrak{D}\subset 2^E$ dans $2^{E'}$ et de $\mathfrak{D}'\subset 2^{E'}$ dans $2^{E''}$ respectivement. Posons $\mathfrak{D}''=h^{-1}(\mathfrak{D}')$ et, pour $X\in\mathfrak{D}''$, $h''(X)=h'[h(X)]$. Alors h'' est une application monotone de $\mathfrak{D}''$ dans $2^{E''}$ et*

$$<^{h''}\subset(<^h)^{h'}.$$

Si $h(\mathfrak{D})\subset\mathfrak{D}'$, en outre $0\in\mathfrak{D}'$ entraîne soit $h'(0)=0$, soit $0\in\mathfrak{D}$ et $h(0)=0$, de plus, si $E'\in\mathfrak{D}'$ entraîne soit $h'(E')=E''$, soit $E\in\mathfrak{D}$ et $h(E)=E'$, alors

$$<^{h''}=(<^h)^{h'}.$$

Si h est bimonotone, la condition $h(\mathfrak{D})\subset\mathfrak{D}'$ peut être remplacée par $h(\mathfrak{D})\supset\mathfrak{D}'$.

Démonstration. L'application h'' étant évidemment monotone, posons $A''<^{h''}B''$, $A''\neq 0$, $B''\neq E''$. On a alors

$$A''\subset h''(A)\subset h''(B)\subset B'', \qquad A<B, \qquad A, B\in\mathfrak{D}'',$$

autrement dit

$$A''\subset h'[h(A)]\subset h'[h(B)]\subset B'', \qquad h(A)\in\mathfrak{D}', \qquad h(B)\in\mathfrak{D}',$$

$$h(A)<^h h(B),$$

c'est-à-dire $A''<^{hh'}B''$. Donc $<''\subset<^{hh'}$.

Réciproquement, soit $A''<^{hh'}B''$, $A''\neq 0$, $B''\neq E''$, c'est-à-dire

$$A''\subset h'(A')\subset h'(B')\subset B'', \qquad A', B'\in\mathfrak{D}', \ A'<^h B'.$$

En supposant $A'\neq 0$, $B'\neq E'$, on a encore

$$A'\subset h(A)\subset h(B)\subset B', \qquad A, B\in\mathfrak{D}, \qquad A<B$$

et si $h(\mathfrak{D})\subset\mathfrak{D}'$,

$$h'(A')\subset h'[h(A)]=h''(A)\subset h''(B)=h'[h(B)]\subset h'(B')$$

et par conséquent $A'' <^{h''} B''$. Si $A'=0$, on a donc $0 \in \mathfrak{D}'$ et ou bien $h'(0)=0$, ou bien $0 \in \mathfrak{D}$, $h(0)=0$. Dans le premier cas, $h'(A')=0$ impliquerait $A''=0$, et dans le second, on a

$$0 \in \mathfrak{D}, \quad 0<0, \quad A' \subset h(0) < h(0) \subset B',$$

$$A'' \subset h'[h(0)] = h''(0) \subset h''(0) = h'[h(0)] \subset B'',$$

d'où $A'' <^{h''} B''$. On raisonne de la même façon si $B' = E'$. Par conséquent, $<^{hh'} \subset <^{h''}$.

Si h est bimonotone et $A'' <^{hh'} B''$, $A'' \neq 0$, $B'' \neq E''$, on a de nouveau

$$A'' \subset h'(A') \subset h'(B') \subset B'', \quad A', B' \in \mathfrak{D}', \quad A' <^{h} B'$$

et, pour $A' \neq 0$, $B' \neq E'$,

$$A' \subset h(A) \subset h(B) \subset B', \quad A, B \in \mathfrak{D}, \quad A < B.$$

La condition $h(\mathfrak{D}) \supset \mathfrak{D}'$ entraîne $A' = h(A_1)$, $B' = h(B_1)$, $A_1, B_1 \in \mathfrak{D}''$, d'où vient $A_1 \subset A$, $B \subset B_1$ en conséquence de la bimonotonité de h, donc $A_1 < B_1$. Il en résulte

$$A'' \subset h'[h(A_1)] \subset h'[h(B_1)] \subset B'', \quad A_1, B_1 \in \mathfrak{D}'', \quad A_1 < B_1,$$

c'est-à-dire $A'' <^{h''} B''$. Les cas $A'=0$ et $B'=E'$ se traitent comme plus haut et on obtient cette fois aussi $<^{hh'} \subset <^{h''}$. C. q. f. d.

(2.31) *Soit h une application bimonotone de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$ telle que $0, E \in \mathfrak{D}$ et $h(0)=0$, $h(E)=E'$. Posons $\mathfrak{D}' = h(\mathfrak{D})$ et définissons l'application h' de $\mathfrak{D}'$ dans 2^E par l'égalité $h'[h(X)]=X$ pour $X \in \mathfrak{D}$. Alors*

$$<^{hh'} = (<|\mathfrak{D} \times \mathfrak{D})^e \subset <$$

pour un ordre semi-topogène quelconque sur E. Si de plus $\mathfrak{D}$ est dense pour $<$, on a donc $<^{hh'} = <$.

En effet, on peut appliquer (2.30) en y posant $E''=E$, $\mathfrak{D}''=\mathfrak{D}$, $h''(X)=X$ pour $X \in \mathfrak{D}$. On constate aisément que $<^{h''} = (<|\mathfrak{D} \times \mathfrak{D})^e$ et le reste résulte de (1.4).

3. *Transposition de familles d'ordres.* Soit, comme plus haut, h une application monotone de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$. Considérons une famille d'ordres $\mathscr{A}$ sur E, c'est-à-dire un ensemble non vide d'ordres topogènes sur E. Les ordres semi-topogènes $<^h$, où $< \in \mathscr{A}$, ne seront pas des ordres topogènes en général, c'est pourquoi nous définissons le *transposé* $\mathscr{A}^h$ de la famille d'ordres $\mathscr{A}$ par l'égalité

$$\mathscr{A}^h = \{<^{hq} : < \in \mathscr{A}\}; \tag{3.1}$$

c'est une famille d'ordres sur E'.

Les résultats du paragraphe précédent permettent de déduire les propositions qui vont suivre.

Il résulte de (2.5):

(3.2) *Si $\mathscr{A}$ et $\mathscr{B}$ sont deux familles d'ordres sur E, $\mathscr{A} \prec \mathscr{B}$ implique $\mathscr{A}^h \prec \mathscr{B}^h$ et $\mathscr{A} \sim \mathscr{B}$ entraîne $\mathscr{A}^h \sim \mathscr{B}^h$.*

On déduit de (2.11):

(3.3) *Si* $\mathfrak{D}$ *et* h *satisfont à la condition* (2.12), *on a* $\mathscr{A}^{hc} = \mathscr{A}^{ch}$ *pour une famille d'ordres quelconque et, si* $\mathscr{A}$ *est symétrique,* $\mathscr{A}^h$ *l'est aussi.*

(2.13) a pour conséquence:

(3.4) *Si* $\mathfrak{D}$ *et* h *satisfont à la condition* (2.14), *on a*

$$\mathscr{A}^h = \{<^h : < \in \mathscr{A}\} \tag{3.5}$$

pour une famille d'ordres $\mathscr{A}$ *quelconque.*

Il s'ensuit en vertu de (2.18):

(3.6) *Si* $\mathfrak{D}$ *et* h *satisfont à* (2.14) *et* (2.19), *on a* $\mathscr{A}^{hp} \leqslant \mathscr{A}^{ph}$ *pour une famille d'ordres* $\mathscr{A}$ *quelconque; de plus, si* $\mathscr{A}$ *est parfaite,* $\mathscr{A}^h$ *l'est aussi, si* $\mathfrak{D}$ *est dense pour tous les ordres* $< \in \mathscr{A}$, *on a* $\mathscr{A}^{hp} = \mathscr{A}^{ph}$.

Il résulte de (3.4) d'après (2.20):

(3.7) *Si* $\mathfrak{D}$ *et* h *satisfont à la condition* (2.21) [et à plus forte raison à (2.14)], *on a* $\mathscr{A}^{hb} \leqslant \mathscr{A}^{bh}$ *pour une famille d'ordres* $\mathscr{A}$ *quelconque et, si* $\mathscr{A}$ *est biparfaite,* $\mathscr{A}^h$ *l'est aussi; si* $\mathfrak{D}$ *est dense pour tous les ordres* $< \in \mathscr{A}$, *on a* $\mathscr{A}^{hb} = \mathscr{A}^{bh}$.

On déduit de (2.26):

(3.8) *Si* $\mathfrak{D}$ *et* h *satisfont à* (2.12) *et* (2.14), *on a* $\mathscr{A}^{hs} \leqslant \mathscr{A}^{sh}$ *pour une famille d'ordres* $\mathscr{A}$ *quelconque, et si* $\mathfrak{D}$ *est dense pour tous les ordres* $< \in \mathscr{A}$, *on a* $\mathscr{A}^{hs} = \mathscr{A}^{sh}$.

Si, dans les propositions (3.6) à (3.8), on se contente de remplacer le signe d'égalité par celui d'équivalence, la condition que $\mathfrak{D}$ doit être dense pour tous les ordres $< \in \mathscr{A}$ peut être remplacée par une condition moins restrictive. Pour cela, convenons de dire que $\mathfrak{D}$ est un *séparateur* pour $\mathscr{A}$ si à chaque ordre $< \in \mathscr{A}$ il correspond un ordre $<' \in \mathscr{A}$ tel que $A < B$ implique l'existence de $C, D \in \mathfrak{D}$ ayant la propriété

$$A \subset C <' D \subset B.^5$$

On voit aisément que $\mathfrak{D}$ est un séparateur pour $\mathscr{A}$ si et seulement si $0 \in \mathfrak{D}$, $E \in \mathfrak{D}$ et à chaque ordre $< \in \mathscr{A}$ il correspond $<' \in \mathscr{A}$ tel que

$$< \subset (<' \mid \mathfrak{D} \times \mathfrak{D})^e.$$

Or, on peut compléter (3.6) par la remarque suivante:

(3.9) *Si* $\mathfrak{D}$ *et* h *satisfont à* (2.14) *et* (2.19) *et si* $\mathfrak{D}$ *est un séparateur pour la famille d'ordres* $\mathscr{A}$, *on a* $\mathscr{A}^{hp} \sim \mathscr{A}^{ph}$.

Démonstration. La formule $\mathscr{A}^{hp} \leqslant \mathscr{A}^{ph}$ résulte de (3.6). Réciproquement, si $< \in \mathscr{A}$, $<' \in \mathscr{A}$ et $< \subset (<' \mid \mathfrak{D} \times \mathfrak{D})^e$, alors

$$<^{ph} \subset (<' \mid \mathfrak{D} \times \mathfrak{D})^{eph}$$

d'après (2.5). De plus, $\mathfrak{D}$ étant évidemment dense pour $(<' \mid \mathfrak{D} \times \mathfrak{D})^e$, on a en vertu de (2.18)

$$(<' \mid \mathfrak{D} \times \mathfrak{D})^{eph} = (<' \mid \mathfrak{D} \times \mathfrak{D})^{ehp}.$$

Enfin (1.4) et (2.5) entraînent

$$(<' \mid \mathfrak{D} \times \mathfrak{D})^{ehp} \subset <'^{hp}.$$

Par conséquent $<^{ph} \subset <'^{hp}$ et $\mathscr{A}^{ph} \leqslant \mathscr{A}^{hp}$.

[5] Cf. la notion de i-séparateur, [1], p. 275.

On démontre de la même façon:

(3.10) *Si* $\mathfrak{D}$ *et* h *satisfont à* (2.21) *et si* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{A}$, *on a* $\mathscr{A}^{hb} \sim \mathscr{A}^{bh}$.

(3.11) *Si* $\mathfrak{D}$ *et* h *satisfont à* (2.12) *et* (2.14) *et si* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{A}$, *on a* $\mathscr{A}^{hs} \sim \mathscr{A}^{sh}$.

En ce qui concerne l'opération ^{t}, on peut dire:

(3.12) *Si* $\mathscr{A}$ *est une famille d'ordres sur* E, *on a* $\mathscr{A}^{ht} \leqslant \mathscr{A}^{th}$ *et, si* $\mathfrak{D}$ *et* h *satisfont à* (2.14) *et* $\mathfrak{D}$ *est dense pour*

$$<_0 = \bigcup\{< : < \in \mathscr{A}\}$$

(en particulier si $\mathfrak{D}$ *est un séparateur pour* $\mathscr{A}$*), on a même* $\mathscr{A}^{ht} = \mathscr{A}^{th}$.

En effet, $\mathscr{A}^{ht}$ est composé du seul ordre topogène

$$<' = (\bigcup\{<^{hq} : < \in \mathscr{A}\})^q = (\bigcup\{<^h : < \in \mathscr{A}\})^q =$$

$$= (\bigcup\{< : < \in \mathscr{A}\})^{hq} \subset (\bigcup\{< : < \in \mathscr{A}\})^{qhq} = <_1'$$

d'après (2.6) et (2.5), et $\mathscr{A}^{th}$ est composé du seul ordre $<_1'$. Si $\mathfrak{D}$ est dense pour $<_0$ et si (2.14) est valable, on a le signe d'égalité au lieu de $\subset$ dans cette formule. Enfin, si $\mathfrak{D}$ est un séparateur pour $\mathscr{A}$, il est évidemment dense pour $<_0$.

Quant à l'opération ^{g}, on peut affirmer:

(3.13) *Soit* $\mathscr{A}$ *une famille d'ordres sur* E. *On a* $\mathscr{A}^{hg} \leqslant \mathscr{A}^{gh}$ *et, si* $\mathscr{A}$ *est dirigé,* $\mathscr{A}^h$ *l'est aussi. De plus, si* (2.14) *est valable et* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{A}$, *on a* $\mathscr{A}^{hg} \sim \mathscr{A}^{gh}$.

DÉMONSTRATION. $\mathscr{A}^{hg}$ se compose des ordres topogènes de la forme

$$\left(\bigcup_1^n <_i^{hq}\right)^q \quad \text{où} \quad <_i \in \mathscr{A}.$$

Or, on a d'après (2.6) et (2.5)

$$\left(\bigcup_1^n <_i^{hq}\right)^q = \left(\bigcup_1^n <_i^h\right)^q = \left(\bigcup_1^n <_i\right)^{hq} \subset \left(\bigcup_1^n <_i\right)^{qhq}$$

et, $\mathscr{A}^g$ se composant des ordres ayant la forme

$$\left(\bigcup_1^n <_i\right)^q \quad \text{où} \quad <_i \in \mathscr{A},$$

la première assertion en résulte. Si $\mathscr{A}$ est dirigé, on a $\mathscr{A}^g \leqslant \mathscr{A}$, donc $\mathscr{A}^{hg} \leqslant \mathscr{A}^{gh} \leqslant \mathscr{A}^h$ en vertu de (3.2), c'est-à-dire $\mathscr{A}^h$ est dirigé.

Supposons enfin que (2.14) est rempli et que $\mathfrak{D}$ est un séparateur pour $\mathscr{A}$. Considérons un ordre quelconque de $\mathscr{A}^{gh}$, disons

$$\left(\bigcup_1^n <_i\right)^{qhq} = \left(\bigcup_1^n <_i\right)^{qh} \quad \text{où} \quad <_i \in \mathscr{A}$$

(cf. (2.13)). À chaque ordre $<_i$ il correspond un ordre $<'_i \in \mathscr{A}$ tel que $<_i \subset \subset (<'_i | \mathfrak{D} \times \mathfrak{D})^e = <'^j_i$ où j désigne l'application identique de $\mathfrak{D}$ sur soi-même. On a alors

$$\left(\bigcup_1^n <_i\right)^{qh} \subset \left(\bigcup_1^n <'^j_i\right)^{qh} = \left(\bigcup_1^n <'_i\right)^{jqh} = \left(\bigcup_1^n <'_i\right)^{jhq} \subset$$

$$\subset \left(\bigcup_1^n <'_i\right)^{hq} = \left(\bigcup_1^n <'^h_i\right)^{q} = \left(\bigcup_1^n <'^{hq}_i\right)^{q} \in \mathscr{A}^{hg}$$

en vertu de (2.5), (2.6), (2.13), (1.4) et du fait que $\mathfrak{D}$ est dense pour $\left(\bigcup_1^n <'_i\right)^j$.

Donc $\mathscr{A}^{gh} \prec \mathscr{A}^{hg}$.

(3.14) *Si $\mathfrak{D}$ est un séparateur pour la famille d'ordres $\mathscr{A}$, on a $\mathscr{A}^{2h} \prec \mathscr{A}^{h2}$.*

Démonstration. Posons $A' <^{2h} B'$, $< \in \mathscr{A}$, et soit $<' \in \mathscr{A}$ un ordre correspondant à $<$ d'après la définition du séparateur. Si $A' \neq 0$, $B' \neq E'$, on a

$$A' \subset h(A) \subset h(B) \subset B', \qquad A <^2 B,$$

donc $A < C < B$. Il existe par conséquent $A_1, C_1, C_2, B_1 \in \mathfrak{D}$ tels que

$$A \subset A_1 <' C_1 \subset C \subset C_2 <' B_1 \subset B,$$

d'où

$$A' \subset h(A) \subset h(A_1) \subset h(C_1) \subset h(C_2) \subset h(B_1) \subset h(B) \subset B',$$

$$A' <'^h h(C_1) <'^h B', \qquad A' <'^{h2} B'.$$

Il vient donc

$$<^{2h} \subset <'^{h2}$$

et par suite

$$<^{2hq} \subset <'^{h2q} \subset <'^{hq2} \in \mathscr{A}^{h2}.$$

Étant donné que $<^{2hq}$ désigne un ordre quelconque tiré de $\mathscr{A}^{2h}$, il s'ensuit $\mathscr{A}^{2h} \prec \mathscr{A}^{h2}$.

La proposition suivante se démontre de la même façon que (2.28):

(3.15) *Soient h une application monotone de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$, $\mathfrak{D}_1 \subset \mathfrak{D}$, $\mathfrak{D}_2 \subset \mathfrak{D}$, $h_1 = h|\mathfrak{D}_1$, $h_2 = h|\mathfrak{D}_2$ et $\mathscr{A}$ une famille d'ordres sur E. Si*

(3.16) *$<_1 \in \mathscr{A}$ entraîne l'existence de $<_2 \in \mathscr{A}$ tel que $A, B \in \mathfrak{D}_1$, $A <_1 B$ impliquent $A \subset C <_2 D \subset B$ pour $C, D \in \mathfrak{D}_2$ convenables,*

alors $\mathscr{A}^{h_1} \prec \mathscr{A}^{h_2}$.

De plus, on déduit de (2.30):

(3.17) *Soient h et h' des applications monotones de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$ et de $\mathfrak{D}' \subset 2^{E'}$ dans $2^{E''}$ respectivement. Posons $\mathfrak{D}'' = h^{-1}(\mathfrak{D}')$ et, pour $X \in \mathfrak{D}''$, $h''(X) = h'[h(X)]$. Alors, $\mathscr{A}$ étant une famille d'ordres sur E, on a $\mathscr{A}^{h''} \prec (\mathscr{A}^h)^{h'}$. Si $h(\mathfrak{D}) \subset \mathfrak{D}'$, en outre $0 \in \mathfrak{D}'$ entraîne soit $h'(0) = 0$, soit $0 \in \mathfrak{D}$ et $h(0) = 0$, si, de plus, $E' \in \mathfrak{D}'$ entraîne soit $h'(E') = E''$, soit $E \in \mathfrak{D}$ et $h(E) = E'$, si enfin $\mathfrak{D}$ et h satisfont à (2.14), alors $\mathscr{A}^{h''} = (\mathscr{A}^h)^{h'}$. Si h est bimonotone, la condition $h(\mathfrak{D}) \subset \mathfrak{D}'$ peut être remplacée par $h(\mathfrak{D}) \supset \mathfrak{D}'$.*

La proposition suivante correspond à (2.31):

(3.18) *Soit h une application bimonotone de $\mathfrak{D} \subset 2^E$ dans $2^{E'}$ telle que $0, E \in \mathfrak{D}$ et $h(0) = 0$, $h(E) = E'$ et satisfaisant à (2.14). Posons $\mathfrak{D}' = h(\mathfrak{D})$ et $h'[h(X)] = X$*

pour $X \in \mathfrak{D}$. *Alors* $\mathscr{A}^{hh'} \leqslant \mathscr{A}$ *pour une famille d'ordres quelconque sur* E. *Si de plus* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{A}$, *on a* $\mathscr{A}^{hh'} \sim \mathscr{A}$, *et si* $\mathfrak{D}$ *est dense pour tous les ordres* $< \in \mathscr{A}$, *on a* $\mathscr{A}^{hh'} = \mathscr{A}$.

Toutes les assertions résultent de (2.31), sauf celle qui concerne le cas où $\mathfrak{D}$ est un séparateur pour $\mathscr{A}$. Or, dans ce cas, on a pour $< \in \mathscr{A}$ la formule

$$< \subset (<' | \mathfrak{D} \times \mathfrak{D})^e$$

où $<' \in \mathscr{A}$ désigne un ordre qui correspond à $<$ suivant la définition du séparateur. Étant donné que

$$(<' | \mathfrak{D} \times \mathfrak{D})^e = <'^{hh'}$$

d'après (2.31), on en déduit $\mathscr{A} \leqslant \mathscr{A}^{hh'}$, c. q. f. d.

On déduit aisément de (3.13), (3.2) et (3.14):

(3.19) *Si* $\mathscr{S}$ *est une structure syntopogène sur* E *et* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{S}$, *alors* $\mathscr{S}^h$ *est une structure syntopogène sur* E'.

Il est utile de remarquer:

(3.20) $\mathfrak{D}$ *est un séparateur pour une structure syntopogène* $\mathscr{S}$ *si et seulement si*

(3.21) $A < B$, $< \in \mathscr{S}$ *impliquent l'existence de* $D \in \mathfrak{D}$ *tel que* $A \subset D \subset B$.

En effet, la nécessité de la condition (3.20) étant évidente, supposons qu'elle est vérifiée. Alors, $< \in \mathscr{S}$ étant donné, il existe $<' \in \mathscr{S}$ et $<'' \in \mathscr{S}$ tels que $< \subset <'^2, <' \subset <''^2$. Par conséquent, si $A < B$, on a

$$A <' C <' B, \qquad A <'' A_1 <'' C <'' B_1 <'' B,$$

et par hypothèse

$$A \subset D_1 \subset A_1 <'' B_1 \subset D_2 \subset B, \qquad D_1, D_2 \in \mathfrak{D}.$$

Ceci montre la validité de l'énoncé.

On en déduit:

(3.22) *Si* $\mathfrak{D}$ *est un séparateur pour une structure syntopogène* $\mathscr{S}$ *et si* $\mathscr{S}_1$ *est une structure syntopogène telle que* $\mathscr{S}_1 \leqslant \mathscr{S}$, *alors* $\mathfrak{D}$ *est un séparateur pour* $\mathscr{S}_1$ *aussi.*

Une proposition analogue serait en défaut pour des familles d'ordres quelconques.

4. *Cas particuliers.* Nous allons étudier certains cas particuliers de la construction générale de la transposition définie plus haut.

(4.1) Nous avons indiqué brièvement cette méthode de transposition pour le cas où

$$\mathfrak{D} = 2^E \qquad h(0) = 0, \quad h(E) = E' \tag{4.2}$$

dans notre article [2]. Là cette méthode est la base d'une construction de prolongement d'espaces syntopogènes, en particulier de la complétion et de la compactification de tels espaces. Dans ce cas, $<^h$ est donné par l'égalité (2.3) et $\mathfrak{D}$ est dense pour un ordre semi-topogène quelconque sur E. De plus, si c'est ce cas particulier qui se réalise pour les applications h et h' de la proposition (2.30), on a toujours $<^{h'} = <^{hh'}$; si (4.2) est valable pour une application bimonotone h, on a dans (2.31) $<^{hh'} = <$. On déduit encore de (4.2) que $\mathfrak{D}$ est un séparateur pour toutes les familles d'ordres sur E. D'ailleurs, une généralisation d'autre nature de ce cas particulier a été étudiée dans [1], pp. 231 à 238.

(4.3) Citons le cas particulier suivant du cas étudié sous (4.1): soient f une application de E' dans E, $\mathfrak{D}=2^E$ et

$$h(X)=f^{-1}(X) \quad \text{pour} \quad X\subset E. \tag{4.4}$$

On constate aisément que, dans ce cas,

$$<^h = f^{-1}(<) \tag{4.5}$$

pour un ordre semi-topogène quelconque sur E.[6] Dans ce cas, les conditions (2.12), (2.14), (2.19) et (2.21) sont remplies. Il est aisé de vérifier que l'application h de (4.4) est bimonotone si et seulement si $f(E')=E$; en effet, si h est bimonotone, $h(E)=h[f(E')]=E'$ entraîne $f(E')=E$ et, réciproquement, si $f(E')=E$, alors $x\in A-B$ implique $x=f(x')$, $x'\in E'$ et $x'\in f^{-1}(A)-f^{-1}(B)$, c'est-à-dire $f^{-1}(A)\subset f^{-1}(B)$ entraîne $A\subset B$ et par suite h est bimonotone. De là, on déduit aisément tous les résultats principaux de [1], pp. 60 à 65, 105 et 106.

(4.6) Admettons les hypothèses de (4.3) et supposons vérifiée la condition $f(E')=E$. Alors l'application h définie par (4.4) est bimonotone, de sorte qu'on peut considérer l'application h' de la proposition (2.31). Autrement dit, posons

$$\mathfrak{D}'=h(\mathfrak{D})=\{f^{-1}(X) : X\subset E\}, \tag{4.7}$$

$$h'[f^{-1}(X)]=X \quad \text{pour } X\subset E. \tag{4.8}$$

On voit aisément que

$$<'^{h'}=f(<') \tag{4.9}$$

pour un ordre semi-topogène quelconque $<'$ sur E'. En effet, si $A<'^{h'}B$, on a

$$A\subset h'(A')\subset h'(B')\subset B, \quad A', B'\in\mathfrak{D}', \ A'<'B', \tag{4.10}$$

donc

$$A'=f^{-1}(A_1),\ B'=f^{-1}(B_1),\ h'(A')=A_1,\ h'(B')=B_1,$$

et $f^{-1}(A_1)<'f^{-1}(B_1)$ entraîne $A_1\, f(<')B_1$, d'où résulte $A\, f(<')B$.[7] Réciproquement, si $Af(<')B$, on a $f^{-1}(A)<'f^{-1}(B)$, donc en posant $A'=f^{-1}(A)$, $B'=f^{-1}(B)$, on a (4.10) et $A<'^{h'}B$.

En échangeant les rôles de E et E', on peut donc dire: soient f une application de E *sur* E',

$$\mathfrak{D}=\{f^{-1}(X') : X'\subset E'\}, \tag{4.11}$$

$$h(X)=f(X) \quad \text{pour} \quad X\in\mathfrak{D}. \tag{4.12}$$

Alors $<^h=f(<)$ pour un ordre semi-topogène quelconque sur E. On a $0, E\in\mathfrak{D}$, $h(0)=0$, $h(E)=E'$, par suite (2.3) est toujours valable. De plus, l'application h de (4.12) est toujours bimonotone et elle satisfait aux conditions (2.12), (2.14), (2.19) et (2.21). Comme $h(\mathfrak{D})=2^{E'}$, on a dans (2.30) l'égalité $<^{h'}=<^{hh'}$ si h est de la forme (4.12) et h' est une application monotone quelconque de $\mathfrak{D}'\subset 2^{E'}$

[6] Cf. [1], p. 61, proposition (6.4).
[7] V. [1], p. 68, formule (6.29).

dans $2^{E'}$. On déduit aisément de (3.20) que le système $\mathfrak{D}$ de (4.11) est un séparateur pour une structure syntopogène $\mathscr{S}$ si et seulement si l'application f est compatible avec $\mathscr{S}$.[8] Tout cela fournit les résultats principaux de [1], pp. 68 à 70 et 112 à 114.

(4.13) Soit f une application de E dans E' et posons $\mathfrak{D}=2^E$, $h(X)=f(X)$ pour $X\subset E$. Alors $\mathfrak{D}$ est dense pour un ordre semi-topogène quelconque sur E et $<^h$ est défini par (2.3). Cette application h est bimonotone si et seulement si l'application f est biunivoque. La condition (2.19) est toujours vérifiée et (2.14) et (2.21) le sont si f est biunivoque. Dans (2.30), on a l'égalité $<^{h''}=<^{hh'}$ si h et h' sont du type envisagé ici et si l'on a pour les applications f et f' correspondantes soit $f(E)=E'$, soit $f'(E')=E''$. De plus, $\mathfrak{D}$ est un séparateur pour chaque famille d'ordres $\mathscr{A}$ sur E, donc, d'après (3.19), si $\mathscr{S}$ est une structure syntopogène sur E, $\mathscr{S}^h$ est une structure syntopogène sur E'.

(4.14) Soient de nouveau f une application de E dans E' et $\mathfrak{D}=2^E$, mais posons $k(X)=E'-f(E-X)$ pour $X\subset E$. Les mêmes remarques s'appliquent ici, en ce qui concerne la bimonotonité de k, les conditions (2.14) et (2.21) et la proposition (2.30), que pour l'application h étudiée sous (4.13). (2.19) est valable si f est biunivoque. De plus, il y a une connexion étroite entre ces deux méthodes de transposition, puisque l'égalité

$$<^k = <^{chc} \tag{4.15}$$

est valable pour un ordre semi-topogène quelconque sur E en vertu de (2.11).[9]

(4.16) Citons ce cas particulier de (4.13) où $E\subset E'$ et f est l'application canonique de E dans E'. Dans ce cas, h est bimonotone et par suite les conditions (2.14), (2.19) et (2.21) se trouvent vérifiées. Au moyen de cette méthode de transposition, la méthode (4.13) peut être réduite au cas particulier où $f(E)=E'$. En effet, si f est une application de E dans E' et si l'on définit h d'après (4.13), puis on considère f comme application de E sur $f(E)\subset E'$ et on définit h_1 d'après (4.13) à partir de cette application, enfin si h_2 désigne l'application de (4.16) construite à partir de l'application canonique de $f(E)$ dans E', on déduit de (2.30) l'égalité $<^h=<^{h_1h_2}$.

(4.17) Le raisonnement précédent peut être appliqué pour généraliser la méthode (4.6). En effet, soit f une application de E *dans* E' et définissons $\mathfrak{D}$ et h par les formules (4.11) et (4.12). Cette application h est encore bimonotone et satisfait toujours à (2.14), (2.19) et (2.21). En désignant par h_1 l'application construite suivant (4.11) et (4.12) à partir de f, considéré comme application de E sur $f(E)$, et par h_2 l'application déduite d'après (4.16) de l'application canonique de $f(E)$ dans E', on obtient $<^h=<^{h_1h_2}$ en vertu de (2.30).

(4.18) Soient f une application de E dans E' et $\mathscr{T}=\{<\}$ une topologie sur E. Posons

$$\mathfrak{D}=\{X : X=f^{-1}(X'),\ X<X\}, \tag{4.19}$$

$$h(X)=f(X) \quad \text{pour} \quad X\in\mathfrak{D}. \tag{4.20}$$

[8] V. [1], p. 112.

[9] Les transposés $<^h$ et $<^k$ de (4.13) et (4.14) sont considérés dans l'article [3] pour le cas où $\{<\}$ est une structure topogène sur E. En effet, le terme „proximity structure" de [3] équivaut à „structure topogène" et les symboles $f^0(<)$ et $f^c(<)$ de [3] correspondent à $<^{hq}$ et à $<^{kq}$ respectivement.

Ce système $\mathfrak{D}$ n'est pas en général un séparateur pour $\mathscr{T}$, mais on constate aisément que $\mathfrak{D}$ est une topologie classique sur E et en désignant par $\mathscr{T}_{\mathfrak{D}}$ la topologie associée à $\mathfrak{D}$, le système $\mathfrak{D}$ est évidemment un séparateur pour $\mathscr{T}_{\mathfrak{D}}$. On voit tout comme dans (4.17) que h est bimonotone et remplit les conditions (2.14) et (2.19). Par conséquent, $\mathscr{T}^h = \mathscr{T}^h_{\mathfrak{D}} = \{<^h\}$ est une topologie sur E' en vertu de (2.13), (2.18) et (3.19). Lorsque $f(E) = E'$, la topologie classique associée à $\mathscr{T}^h$ est identique à la „topologie quotient" correspondant à f et à la topologie classique associée à $\mathscr{T}$.[10] D'ailleurs, l'application de la méthode de (4.16) permet cette fois aussi de se borner à ce cas particulier dans lequel $f(E) = E'$.

Littérature

[1] Á. Császár, *Grundlagen der allgemeinen Topologie* (Budapest – Leipzig, 1963).
[2] Á. Császár, Complétion et compactification d'espaces syntopogènes, General Topology and its Relations to Modern Analysis and Algebra, *Proc. Symp. Prague* 1961, 133 – 137.
[3] C. H. Dowker, Mappings of proximity structures, General Topology and its Relations to Modern Analysis and Algebra, *Proc. Symp. Prague* 1961, 139 – 141.
[4] H. – J. Kowalsky, *Topologische Räume* (Basel und Stuttgart, 1961).

[10] V. p. ex. [4], p. 139.

RANKS OF MODULES

By

L. FUCHS

Department of Algebra of the Eötvös Loránd University, Budapest

(Received May 27, 1963)

The notion of rank plays an important role in the theory of vector spaces and abelian groups, since it is a useful invariant. It is known, however, that no invariant rank can be assigned to modules over arbitrary rings.[1] The objective of this note is to find out which rings R with identity have the property that each unitary left module over R admits an invariant rank.

The definition of rank is based on the concept of (linear) independence. This concept has been used in abelian group theory in two different senses. In the first sense independence can hold only for elements of infinite order; this is the definition of BAER introduced in his paper [1] and used e. g. by KUROSH [5]. In the second sense, elements of any order ($\neq 1$) can be independent, and so this seems to be more adequate from the general point of view.[2] In abelian group theory it is shown that if only maximal independent sets containing elements of orders infinity and powers of primes are considered, then the cardinal number of elements of infinite order as well as the cardinal numbers of elements of orders that are powers of fixed primes are the same for all maximal independent sets of the group, i. e., they are invariants of the group.

If we wish to carry over this last result to modules over certain rings R, then we have to impose on R certain conditions in order to exclude undesired phenomena. We require the existence of maximal independent sets where *1.* no element may be replaced by two or more elements without violating independence, and *2.* with every element a we can associate a quantity (which we have chosen to be the order of a – as in the case of abelian groups) so that if we replace the element a by some non-zero element in the cyclic module Ra, the associated quantity should remain the same. These rather natural conditions[3] lead us directly to hypotheses (i) and (ii) in **§ 4**, which enable us to establish a result analogous to the one mentioned for abelian groups.

The rings with the desired properties can be characterized intrinsically (Theorem 2). It turns out that a sufficiently large amount of factor rings of these rings must be Ore domains (just as the ring of integers has many factor rings which are fields). We shall see that the uniqueness of ranks is closely connected with the uniqueness of ranks of vector spaces over fields.[4]

[1] See for instance LEAVITT [6] or RÉDEI [9], p. 255.

[2] The first systematic use of this kind of independence has been made by SZELE [10]; cf. also [2].

[3] The first condition serves to exclude the possibility that a module of rank 1 is at the same time a module of higher rank, while the second condition is intended to ensure that if in a maximal independent set we replace an element by a (specified) other one, the associated quantity does not change. – Of course, we could choose another associated quantity, and then investigate the same problem for this.

[4] This will turn out from the proof of Theorem 1.

§ 1. Independence

Let R denote throughout an (associative) ring with identity 1. We shall exclusively consider unitary left modules over R which we shall call, for the sake of brevity, *R-modules.*

The non-zero elements $a_1, \ldots, a_n$ of an R-module M will be called linearly independent, or simply, *independent* if

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0 \qquad (\lambda_i \in R) \tag{1}$$

implies

$$\lambda_1 a_1 = 0, \ldots, \lambda_n a_n = 0. \tag{2}$$

If $O(a)$ denotes the *order* of $a \in M$, i. e. the left ideal of R consisting of all $\lambda \in R$ with $\lambda a = 0$, then the conclusion (2) can also be written in the form

$$\lambda_1 \in O(a_1), \ldots, \lambda_n \in O(a_n).$$

Naturally, we say that an infinite set of elements is independent if each of its finite subsets is independent. Hence the independence of a set a_λ $(\lambda \in \Lambda)$ is equivalent to the condition that the submodule $\{\ldots, a_\lambda, \ldots\}$ generated by all the a_λ is the direct sum of the cyclic submodules Ra_λ:

$$\{\ldots, a_\lambda, \ldots\} = \ldots \oplus Ra_\lambda \oplus \ldots$$

On the other hand, the elements $a_1, \ldots, a_n \in M$ are called *strongly independent* if (1) implies

$$\lambda_1 = 0, \ldots, \lambda_n = 0. \tag{3}$$

Hence in a strongly independent set every element must be of order 0; further, independence and strong independence are equivalent for elements of order 0. The elements a_λ $(\lambda \in \Lambda)$ form a strongly independent set if and only if the submodule $\{\ldots, a_\lambda, \ldots\}$ is a free R-module with the a_λ as free generators.

An R-module $N \neq 0$ is said to be *of rank* 1 if it does not contain two independent elements, or, in other words, if no two non-zero submodules of N have 0 intersection. It is readily checked that the R-module R/L where L is a left ideal of R is of rank 1 if and only if L is irreducible in the sense that it cannot be represented as the intersection of two left ideals, other than L, of R. By making use of this simple observation we verify:

LEMMA 1. *Let A be an ideal of R. Then R/A as an R-module is of rank* 1 *if and only if R/A as factor ring is a left Ore ring.*

Here we mean by a *left Ore ring*[5] a ring S (with identity) such that every pair of non-zero left ideals of S has non-zero intersection, that is, $S\lambda \neq 0$ and $S\mu \neq 0$ imply $S\lambda \cap S\mu \neq 0$. Since this is equivalent to the condition that the zero ideal is an irreducible left ideal, and since the irreducibility of 0 in R/A is the same thing as that of A in R, Lemma 1 follows at once from our previous observation.

[5] See ORE [8]. An Ore ring without divisors of zero will be called an *Ore-domain.*

Independence and strong independence are properties of finite character. Therefore, in every R-module M there exist maximal (strongly) independent sets. Moreover, every (strongly) independent set of M can be extended to a maximal (strongly) independent set of M.

§ 2. Modules over Ore domains

First we turn our attention to strong independence. We focus our attention on elements of order 0 and are looking for conditions on R to ensure, for all R-modules, the uniqueness of the number of elements of order 0 in maximal independent sets. These rings can be described easily as left Ore domains:

THEOREM 1. *A ring R has the property that, for all R-modules M, the cardinal number of elements of order 0 in maximal independent sets of M depends only on M, if and only if R is a left Ore domain.*

Assume that R enjoys the stated property. Then in the R-module R the element 1 forms a maximal independent set, and $O(1) = 0$. Thus in R every maximal independent set contains just one element of order 0. By way of contradiction suppose that R is not torsion-free[6]. Then it contains independent sets of elements of order $\neq 0$, and among them there exists a maximal one, say U. We claim that U is a maximal independent set in R. For if a is an element of order 0 which is independent of U, then in view of the R-isomorphism $Ra \cong R$, Ra certainly contains an element of order $\neq 0$ which is obviously again independent of U, contrary to the choice of U. Therefore R contains a maximal independent set without elements of order 0, a contradiction. Consequently, R is a torsion-free R-module, i. e. R has no divisors of zero. Independence and strong independence thus coincide in R.

Hence if λ, μ are arbitrary non-zero elements of R, then they are of order 0 and are not independent. Therefore, for some $\lambda_1, \mu_1 \in R$, different from 0, we have $\lambda_1\lambda + \mu_1\mu = 0$. This shows that $R\lambda \cap R\mu \neq 0$, and R is a left Ore domain.

To prove the converse, assume that R is a left Ore domain, Q is its field of left quotients, and M is an R-module. First we dispose of the case in which M is torsion-free. Let

$$\varphi : M \to Q \otimes_R M$$

denote the canonical map $x \to 1 \otimes x$ $(x \in M)$ where Q is to be considered as a right R-module. We show that φ is a monomorphism,[7] i. e. $1 \otimes x$ is never 0 unless $x = 0$. If $1 \otimes x$ vanishes in $Q \otimes_R M$, then it vanishes if regarded to belong to $R^* \otimes_R M$ for some suitable finitely generated right R-submodule R^* of Q. If R^* is generated by $\xi_1, \ldots, \xi_n (\in Q)$ and if $\lambda \in R$ is a common left denominator[8] of the ξ_i, $\xi_i = \lambda^{-1}\lambda_i$ $(\lambda_i \in R)$, then $R^* \subseteq \lambda^{-1}R$. Evidently, $1 \otimes x$ must vanish in

[6] A module is called *torsion-free* if each of its elements $\neq 0$ is of order 0. Evidently, R as an R-module is torsion-free exactly if it contains no divisors of zero.

[7] A similar proof for the monomorphism of φ has been given by LEVY [7]. This paragraph and the next can be replaced by a simple appeal to KERTÉSZ [4] where it is shown that two maximal independent sets of torsion-free R-modules have the same cardinal number if and only if R is a left Ore domain.

[8] If λ is an element $\neq 0$ common to all left ideals generated by the denominators, then it is a common left denominator.

$(\lambda^{-1}R)\otimes_R M$ too. Here $\lambda^{-1}R$ and R are isomorphic as right R-modules, and so the group isomorphisms $\lambda^{-1}R\otimes_R M \cong R\otimes_R M \cong M$ hold under the correspondences $\lambda^{-1}\mu\otimes y \to \mu\otimes y \to \mu y$ $(y\in M)$. Therefore $0 = 1\otimes x = \lambda^{-1}\lambda\otimes x \to \lambda x$ whence $\lambda x = 0$ and $x = 0$. Thus φ embeds M R-isomorphically in the left Q-vector space $Q\otimes_R M = M^*$.

We continue the proof with verifying that if $a_1, \ldots, a_n$ are independent elements in the torsion-free module M, then their images $1\otimes a_1, \ldots, 1\otimes a_n$ are independent elements in M^*. For, if

$$\xi_1(1\otimes a_1)+\ldots+\xi_n(1\otimes a_n) = 0 \qquad (\xi_i\in Q),$$

then writing $\xi_i = \lambda^{-1}\lambda_i$ $(\lambda, \lambda_i\in R)$ we have

$$\lambda^{-1}(\lambda_1\otimes a_1)+\cdots+\lambda^{-1}(\lambda_n\otimes a_n) = \lambda^{-1}[1\otimes(\lambda_1 a_1+\cdots+\lambda_n a_n)] = 0,$$

and hence $\lambda_1 a_1+\ldots+\lambda_n a_n = 0$; therefore all λ_i and all ξ_i vanish. Since the converse is also true and since the elements of M^* may be written in the form $\lambda^{-1}(1\otimes b)$ with $\lambda\in R$, $b\in M$, it follows that $\ldots, a_\lambda, \ldots$ form a maximal independent set in M if and only if $\ldots, 1\otimes a_\lambda, \ldots$ form a basis in M^*. But in vector spaces over fields the cardinalities of all bases are the same.[9] We conclude that the same holds for the maximal independent sets in M.

Turning to the case where M is arbitrary, first of all we show that the elements of order $\neq 0$ of M form a submodule T, the torsion submodule of M, such that the factor module M/T is torsion-free[10]. If $a, b\in M$ and $O(a)\neq 0$, $O(b)\neq 0$, then, R being a left Ore domain, $O(a)\cap O(b)\neq 0$, whence the obvious inclusion $O(a)\cap O(b)\subseteq O(a-b)$ implies $O(a-b)\neq 0$. Furthermore, for every $\lambda\in R$, $\lambda\neq 0$, we have $O(a)\cap R\lambda\neq 0$ which implies $\mu\lambda\in O(a)$ for some $\mu\in R$, $\mu\neq 0$, that is, $\mu(\lambda a) = 0$ and $O(\lambda a)\neq 0$. Thus the elements of order $\neq 0$ form in fact a submodule T. If[11] $a^*\in M/T$, $a^*\neq 0$ satisfies $\lambda a^* = 0$ for some $\lambda\in R$, then $\lambda a\in T$ for any representative a of a^*, and hence $\mu\lambda a = 0$ for some $\mu\in R$, $\mu\neq 0$. From $O(a) = 0$ we conclude $\mu\lambda = 0$ and $\lambda = 0$. Consequently, $O(a^*) = 0$, and M/T is torsion-free.

Now it is straightforward to prove that $a_1, \ldots, a_n$ are strongly independent elements in M if and only if their cosets $a_1^*, \ldots, a_n^*$ are (strongly) independent in M/T. Hence it follows that the cardinal numbers of maximal strongly independent sets in M are the same as those in M/T. But these have been shown to be uniquely determined by the torsion-free module M/T. Thus it remains only to verify that in a maximal independent set of M the elements of order 0 form a maximal strongly independent set. This is evident, since any dependence relation (1) can be multiplied by a non-zero $\lambda\in R$ to get a dependence relation containing only elements of order 0.

This completes the proof of Theorem 1.

COROLLARY. *In a module over a left Ore domain the cardinality of a maximal strongly independent set is an invariant of the module.* (KERTÉSZ [4].)

[9] For a proof of this we refer e. g. to JACOBSON [3].
[10] Compare LEVY [7].
[11] By asterisks we denote cosets.

§ 3. Homogeneous elements

A non-zero element a of an R-module M will be called *homogeneous* if

(4) for each $\lambda \in R$, either $\lambda a = 0$ or $O(\lambda a) = O(a)$.

This means that in the submodule Ra the non-zero elements have the same order. Obviously, if a is homogeneous and $\lambda a \neq 0$, then λa is homogeneous too.

Manifestly, in the case of vector spaces, every element $\neq 0$ is homogeneous, while in the case of abelian groups, an element is homogeneous if and only if its order is a prime or infinity. In the general case we have:

LEMMA 2. *An element of an R-module is homogeneous if and only if its order is a complete prime ideal of R.*[12]

Assume that a is a homogeneous element of the R-module M, and set $O(a) = P$. We know that P is a left ideal. Since P is the annihilator of the module Ra, P is a two-sided ideal. Next suppose that, for $\lambda, \mu \in R$, $\lambda\mu \in P$ holds where $\mu \notin P$. Then $\mu a \neq 0$ implies $O(\mu a) = P$, and therefore from $\lambda\mu a = 0$ we obtain $\lambda \in P$. Consequently, P is a complete prime ideal.

Conversely, let P be a complete prime ideal of R, and $a \in M$ of order P. Assume $\lambda \in R$ satisfies $\lambda a \neq 0$. Now $\mu\lambda a = 0$ if and only if $\mu\lambda \in P$, which holds true, because of $\lambda \notin P$, exactly if $\mu \in P$. In other words, $O(\lambda a) = P$ and we infer that a is homogeneous. Q. e. d.

In general, a module need not contain homogeneous elements. If N is an R-module of rank 1 and if it contains a homogeneous element a, then for every non-zero $b \in N$ there exists a $\lambda \in R$ such that λb is homogeneous. In fact, a and b are dependent, hence $\lambda b = \mu a \neq 0$ for some $\lambda, \mu \in R$, and so λb is homogeneous. Obviously, λb has the same order as a. Therefore: if an R-module N of rank 1 contains a homogeneous element, then every homogeneous element of N has the same order.

§ 4. The underlying rings

As noted in the introduction, we wish to have an invariant rank for R-modules, and to this end we have to guarantee the existence of maximal independent sets subject to conditions 1. and 2. mentioned in the introduction. These conditions mean that there must exist maximal independent sets where each element is homogeneous and generates a submodule of rank 1. This is implied by the milder condition that every non-zero R-module contains an element which generates a submodule of rank 1, i. e.

(i) *every R-module* $\neq 0$ *contains a submodule of rank* 1,

(ii) *every R-module of rank* 1 *contains a homogeneous element.*

If a ring R possesses these properties and if R' is a homomorphic image of R, then R' also shares these properties, for every R'-module can be made in the natural way into an R-module.

The next result characterizes the rings R satisfying (i) and (ii).

[12] By a complete prime ideal of R we mean a two-sided ideal P of R such that $\lambda\mu \in P$ implies that either $\lambda \in P$ or $\mu \in P$; we also assume that $P \neq R$. – Note that $O(a) = R$ is equivalent to $a = 0$.

THEOREM 2. *A ring R has properties* (i) *and* (ii) *if and only if*

(a) *for every complete prime ideal P of R, the factor ring R/P is a left Ore domain;*

(b) *to every left ideal L of R there exists a $\lambda \in R$ such that $L : \lambda$ is a complete prime ideal of R.*

Here $L : \lambda$ denotes the *right ideal-quotient of L* by λ, i. e. the set of all $\nu \in R$ such that $\nu\lambda \in L$. Evidently, $L : \lambda$ is a left ideal for all left ideals L and for all $\lambda \in R$; condition (b) requires that $L : \lambda$ should be two-sided in the mentioned case.

Let us begin the proof with assuming that R satisfies (i) and (ii). Let P be a complete prime ideal of R, and consider the factor ring $R/P = R^*$ as an R-module. Then every non-zero element of R^* is of order P, and so every element $\neq 0$ in R^* is homogeneous. If Ra^* is of rank 1 in R^*, then it is R-isomorphic to R/P; consequently, (i) implies that R^* must be of rank 1. From Lemma 1 we infer that (a) holds.

To verify the fulfilment of (b), let L be a left ideal of R. The R-module R/L contains, by (i) and (ii), a homogeneous element a of order, say P, which is by virtue of Lemma 2 a complete prime ideal. If $\lambda \in R$ is a representative of the coset a, then $\nu\lambda \in L$ holds for some $\nu \in R$ if and only if $\nu \in P$. This amounts to $L : \lambda = P$, establishing (b).

Conversely, suppose that R satisfies conditions (a) and (b). Let M be an R-module $\neq 0$, and $b \in M$, $b \neq 0$. The submodule Rb is R-isomorphic to R/L [where $L = O(b)$] under the mapping $\mu b \to \mu + L$ ($\mu \in R$). By (b) we can choose a $\lambda \in R$ with $L : \lambda$ equal to a complete prime ideal P. Then $\lambda + L$ is of order P in R/L. Because of Lemma 2, $\lambda + L$ is homogeneous in R/L, and therefore λb is homogeneous in M. In view of the R-isomorphism $R(\lambda b) \cong R/P$ and (a), Lemma 1 implies that $R(\lambda b)$ is of rank 1. The proof of Theorem 2 is thereby completed.

Note that condition (a) can be omitted if R is a commutative ring. In fact, in this case P is a complete prime ideal if and only if it is a prime ideal which is equivalent to the fact that R/P is a domain of integrity (i. e. a commutative Ore domain).

§ 5. Maximal independent sets

We are going to prove that the conditions of **§ 4** in fact guarantee the existence of maximal independent sets which contain homogeneous elements generating submodules of rank 1.

THEOREM 3. *If the ring R satisfies* (i) *and* (ii) *of* **§ 4**, *then in an arbitrary R-module M there exists a set $U = [a_\lambda]_{\lambda \in \Lambda}$ such that*

(a) *every a_λ is homogeneous;*

(b) *each submodule Ra_λ is of rank 1;*

(c) *U is a maximal independent set of M.*

Assume R satisfies (i) and (ii), and let M be an R-module. Define U as a maximal independent set of homogeneous elements of M; that is, U is an independent set of homogeneous elements such that if we adjoin any homogeneous element of M to U, the larger set will not be independent. If $a \in U$, then by Lemma 2, Ra is R-isomorphic to R/P for some complete prime ideal P of R. Theorem 2 and Lemma 1 imply that R/P is of rank 1. This establishes (b). If a $b \in M$ existed

independent of U, then because of (i) and (ii), Rb would contain a homogeneous element. This would clearly be independent of U, contrary to the choice of U. Hence U satisfies (c), and the proof is finished.

In particular, if R is the ring of integers (abelian groups evidently satisfy (i) and (ii)), then Theorem 3 asserts the existence of maximal independent sets (in abelian groups) containing elements of order infinity and primes.

§ 6. The ranks

With every subset U satisfying conditions of Theorem 3 we can associate, for every complete prime ideal P of R, a cardinal number as follows. Let r_P be the cardinality of the set of all those a_λ in U which are of order P. We shall call r_P the *P-rank* of the R-module M. That this definition is legitimate, i. e. that it depends only on M and not on the special choice of U in M, is shown by the next result.

THEOREM 4. *If the ring R has properties* (i) *and* (ii) *of* § 4, *then for every complete prime ideal P of R, the P-rank $r_P = r_P(M)$ of the R-module M is uniquely determined by M.*

Let M be an R-module where R satisfies (i) and (ii). For every complete prime P of M we consider the submodule of M consisting of all elements annihilated by P:

$$M(P) = [x \in M | \lambda x = 0 \text{ for all } \lambda \in P].$$

Since P is a two-sided ideal, $M(P)$ is actually a submodule of M which may also be regarded as a module over the ring R/P in the obvious manner. This $M(P)$ contains those elements of a subset U (satisfying conditions (a)–(c) of Theorem 3) whose orders are complete prime ideals containing P; we denote this subset of U by $U(P)$. We claim that $U(P)$ is a maximal independent set in the R/P-module $M(P)$. Each $a \in M(P)$ depends on U (over R), that is to say, there is a relation

$$\lambda a = \lambda_1 a_1 + \ldots + \lambda_n a_n \neq 0 \quad (\lambda, \lambda_i \in R,\ a_i \in U,\ \lambda_i a_i \neq 0). \tag{5}$$

Here no a_i may occur whose order P_i does not contain P. For otherwise there exists a $\mu \in P$, $\mu \notin P_i$, and if we multiply (5) by μ, the left member vanishes, and the independence of the a_i implies $\mu\lambda_i a_i = 0$, $\mu \in O(\lambda_i a_i) = P_i$, a contradiction. This shows that $a_1, \ldots, a_n \in U(P)$, and hence (5) is a dependence relation in $M(P)$ provided we replace the λ_i by their cosets mod P. We obtain that $U(P)$ is a maximal independent set in $M(P)$.

To conclude the proof, observe that just the elements of $U(P)$ which are of order P in the R-module M, are of order 0 in the R/P-module $M(P)$. Since, by Theorem 2, R/P is a left Ore domain, we conclude from Theorem 1 that the cardinal number of elements of order 0 in $U(P)$ (over R/P) is uniquely determined by $M(P)$, and hence by M. This completes the proof.

We have thus shown that if a ring R satisfies conditions of Theorem 2 then to every R-module M we can assign a cardinal number $r_P(M)$, one for each complete prime ideal P of R, which is an invariant of the module M.

§ 7. Noetherian rings

To illustrate our results we exhibit a non-trivial example for a class of rings satisfying the conditions stated above. We show that commutative Noetherian rings (in particular, polynomial rings in a finite number of indeterminates over commutative fields) satisfy these conditions, and so over them the modules admit an invariant definition of rank.

THEOREM 5. *A commutative Noetherian ring (with identity) satisfies conditions* (a)–(b) *of Theorem 2.*

We have noted that (a) is always satisfied in the commutative case. Hence we may restrict ourselves to verifying the fulfilment of (b). An arbitrary ideal L of R can be written as an intersection of a finite number of primary ideals Q_i:[13]

$$L = Q_1 \cap \ldots \cap Q_m$$

such that no Q_i can be omitted and the primes $P_1, \ldots, P_m$ associated with the primary components are all different. Select an isolated primary ideal, say Q_1; this means that the associated prime P_1 is a minimal one amongst $P_1, \ldots, P_m$. There certainly exists a $\mu \in Q_2 \cap \ldots \cap Q_m$ with $\mu \notin P_1$, for otherwise $Q_2 \cap \ldots \cap Q_m \subseteq P_1$, $Q_2 \ldots Q_m \subseteq P_1$ would imply $Q_i \subseteq P_1$, and so $P_i \subseteq P_1$ for some $i = 2, \ldots, m$. Then

$$L : \mu = Q_1 : \mu \cap (Q_2 \cap \ldots \cap Q_m) : \mu = Q_1.$$

Furthermore, there is an exponent r such that $P_1^r \subseteq Q_1 \subseteq P_1$. If r is chosen to be as small as possible, then $P_1^{r-1} \nsubseteq Q_1$, and thus a $\nu \in P_1^{r-1}$ exists which does not belong to Q_1. For this ν we have $Q_1 : \nu \supseteq P_1$ (because $P_1 \nu \subseteq P_1^r \subseteq Q_1$), while $Q_1 : \nu \subseteq P_1$ by virtue of the primary property of Q_1 and $\nu \notin Q_1$; therefore $Q_1 : \nu = P_1$. Put $\lambda = \mu\nu$, then

$$L : \lambda = (L : \mu) : \nu = Q_1 : \nu = P_1,$$

and (b) of Theorem 2 is satisfied. This establishes our assertion.

Literature

[1] R. BAER, Abelian groups without elements of finite order, *Duke Math. Journal*, **3** (1937), 68–122.
[2] L. FUCHS, *Abelian groups* (Budapest, 1958).
[3] N. JACOBSON, *Lectures in abstract algebra*, vol. 2 (New York, 1953).
[4] A. KERTÉSZ, On independent sets of elements in algebra, *Acta Sci. Math. Szeged*, **21** (1960), 260–269.
[5] (A. G. KUROSH) А. Г. КУРОШ, *Теория групп* (Москва, 1953).
[6] W. G. LEAVITT, Modules without invariant basis numbers, *Proc. Amer. Math. Soc.*, **8** (1957), 322–328.
[7] L. LEVY, Torsion-free and divisible modules over non-integral-domains, *Canadian Journ. Math.*, **15** (1963), 132–151.
[8] O. ORE, Linear equations in non-commutative fields, *Annals of Math.*, **32** (1931), 463–477.
[9] L. RÉDEI, *Algebra*, vol. 1 (Leipzig, 1959).
[10] T. SZELE, Ein Analogon der Körpertheorie für abelsche Gruppen, *Journ. reine u. angew. Math.*, **188** (1950), 167–192.
[11] B. L. VAN DER WAERDEN, *Moderne Algebra*, vol. 2 (Berlin, 1940).

[13] For the needed results on primary ideals we may refer e. g. to VAN DER WAERDEN [11]. Recall that Q is a primary ideal if $ab \in Q$ implies $a \in Q$ or $b^n \in Q$ for some natural number n.

ON RANKS OF MODULES

A REMARK TO THE PRECEDING PAPER OF L. FUCHS

By

A. KERTÉSZ

(Debrecen/Halle (S))

(Received June 29, 1963)

In the preceding paper [1] L. FUCHS introduces a rank-concept for R-modules over rings R satisfying certain conditions. The purpose of the present note is to show that this concept of rank can be extended to the case of arbitrary R-modules, this extension yielding at the same time a simpler and more natural treatment. As an additional remark we prove that the 0-rank of a module (definition see later) is equal to the sum of the 0-rank of an arbitrary submodule and of the 0-rank of the factormodule with respect to this submodule.

In this note we consider unitary modules only, on the basis of [2] however this does not imply any loss of generality.

A non-zero element g of the arbitrary R-module G is called *principal*, if Rg is of rank 1,[1]) and for all non-zero elements g_1 of Rg the relation $O(g_1)=O(g)$ holds.

THEOREM 1. *Let $U=[u_\nu]_{\nu\in\Gamma}$ be a maximal independent system of principal elements of the R-module G, and the left ideal $P(\subseteq R)$ the order of some principal element of the module G. In this case the power of the subset of elements of order P of the system U does not depend on the special choice of the set U, i. e. this power is uniquely determined by the module G.* — This invariant cardinality belonging to P is called the *P-rank* of the module G.

First of all we show that

(*) *the 0-rank of G is uniquely determined.*[2])

We say that a principal element x of order 0 depends on a set $A(\subseteq G)$ of principal elements of order 0 (and we denote this by $D[x, A]$), if λx, for some non-zero $\lambda(\in R)$, is a linear combination over R of a finite subset of the set A. It is clear that the relation $D[x, A]$ so defined on the set H of all principal elements of order 0 of G satisfies axioms (I), (II) and (IV) of [4] on abstract linear dependence. Since the subset in U of elements of order 0 is clearly a maximal independent subset of H,[3]) in order to prove our assertion it will be sufficient — by the Corollary to Theorem 1 in [4] — to show that the relation $D[x, A]$ satisfies also (II) of [4].

[1]) For the terminology and notations see [1].

[2]) This assertion is essentially the same as Theorem 2 of our paper [3]. Since, however, this paper is written in Hungarian, for completeness' sake we give a proof here.

[3]) Let indeed $h\in H$ and

$$(**)\qquad \tau h = \tau_1 u_{i_1} + \ldots + \tau_k u_{i_k} \qquad (\tau\neq 0,\ u_{i_j}\in U).$$

Multiplying this equality by non-zero elements taken in turn from the order of each element with non-zero order occurring on the right hand side, we again get an equality of the form (**), but with $u_{i_j}\in H$ throughout.

Let $D[h, A]$ and $D[a, B]$ be fulfilled for all elements $a \in A$. We show that $D[h, B]$ is also valid. $D[h, A]$ means that, for a suitable finite subset $a_1, \ldots, a_k$ of A, a relation of the form

$$(1) \qquad \varrho_0 h = \varrho_1 a_1 + \cdots + \varrho_k a_k \qquad (\varrho_0 \neq 0 \neq \varrho_i \in R)$$

holds, and in view of $D[a_i, B]$ for suitably chosen elements b_{ij} of B we likewise have the equalities

$$(2) \qquad \left\{\begin{array}{l} \varrho_{10} a_1 = \varrho_{11} b_{11} + \cdots + \varrho_{1m_1} b_{1m_1} \\ \quad \vdots \\ \varrho_{k0} a_k = \varrho_{k1} b_{k1} + \cdots + \varrho_{km_k} b_{km_k} \end{array}\right\} (\varrho_{i0} \neq 0 \neq \varrho_{ij} \in R).$$

By the principality of the elements of H, there exist elements σ_t $(\neq 0)$ and σ'_t $(\neq 0)$ $(t = 1, 2, \ldots, k)$ of R, for which

$$(3) \qquad \left\{\begin{array}{r} \sigma_1 \varrho_1 a_1 = \sigma'_1 \varrho_{10} a_1 \\ \sigma_2 \sigma_1 \varrho_2 a_2 = \sigma'_2 \varrho_{20} a_2 \\ \vdots \\ \sigma_k \cdots \sigma_2 \sigma_1 \varrho_k a_k = \sigma'_k \varrho_{k0} a_k. \end{array}\right.$$

Let us now multiply (1) (from the left) by σ_1, and the first equation of (2) again from the left by σ'_1; by virtue of (3) the element $\sigma_1 \varrho_1 a_1$ can be replaced in the multiplied equation (1) by a linear combination (over R) of the elements b_{ij}. Multiplying the expression so obtained by σ_2 and the second equation of (2) by σ'_2, we effect a further replacement, again on the basis of (3). By a continuation of this process we have finally that the element

$$\sigma_k \cdots \sigma_2 \sigma_1 \varrho_0 h \quad (\neq 0)$$

is a linear combination (over R) of a finite subset of B, giving $D[h, B]$. This completes the proof of the assertion (*).

Let now $g (\in G)$ be an arbitrary principal element and $O(g) = P \neq 0$. Since in this case P is the annihilator of the module Rg, P is an ideal of R. Let us consider the set T_P of all those elements of G which are annihilated by P. T_P is a submodule of G, which can be considered in a natural way as an R/P-module. Clearly, all elements of the set $S = U \cap T_P$ are principal elements of the R/P-module T_P. Let now a be a principal element in T_P. Then a depends on U (over R) i. e. a relation of the form

$$(4) \qquad 0 \neq \lambda a = \lambda_1 u_1 + \cdots + \lambda_n u_n \qquad (\lambda, \lambda_i \in R; u_i \in U; \lambda_i u_i \neq 0)$$

holds. If $\pi \in P$, then on multiplying equation (4) by π the left hand side will be 0, and by the independence of the system U we have $\pi \lambda_i u_i = 0$, i. e. $P \subseteq O(u_i)$ for each i. Thus $u_1, \ldots, u_n \in S$, and since (4) clearly remains valid if λ and the λ_i's are replaced by their cosets with respect to P, we see that S is a maximal independent system of principal elements of the R/P-module T_P. In S exactly those elements of T_P considered as R/P-module are of order 0, which are of order

P in G. Thus, on the basis of (*), the P-rank of G is uniquely determined. This completes the proof of Theorem 1.

Modules without principal elements are said to be of rank 0.

THEOREM 2. *If any submodule $\neq 0$ of an R-module G has rank different from 0, then any maximal independent system of principal elements of G is a maximal independent system of elements of the whole G.*

Let U be a maximal independent system of principal elements of the R-module G. Let $0 \neq g \in G$ and suppose the system of elements $[U, g]$ to be independent. Then for the principal element, say g_0, which exists by hypothesis in the submodule $Rg(\subseteq G)$ we would also have $[U, g_0]$ independent, and this contradicts the maximality of U.

THEOREM 3. *A ring R has the property that any R-module different from 0 has rank different from 0, if and only if*

(a) *for every complete prime ideal P of R, the factor ring R/P is a left Ore domain;*

(b) *to every left ideal L of R there exists a $\lambda \in R$ such that the set of all $\tau \in R$ with $\tau\lambda \in L$ is a complete prime ideal of R.*

Proof essentially as the proof of Theorem 2 in [1].

THEOREM 4. *Let H be a submodule of the R-module G. Then the 0-rank of G is equal to the sum of the 0-ranks of the modules H and G/H.*

Let $[h_\nu]_{\nu \in \Delta}$ be a maximal independent system of principal elements of order 0 of the module H, and $[\overline{g_\mu}]_{\mu \in \Gamma}$ a system of the same kind for the module G/H. In order to prove the theorem it will be sufficient to show that the system of elements

$$S = [h_\nu, g_\mu]_{\nu \in \Delta, \mu \in \Gamma} \qquad (g_\mu \in \overline{g}_\mu)$$

is a maximal independent system of principal elements of order 0 of G.

Since we clearly have $Rg_\mu \cong R\overline{g_\mu}$, the system S consists of principal elements of order 0. Suppose that for the elements of S there holds a relation of the form

$$\varrho_1 h_1 + \cdots + \varrho_k h_k + \sigma_1 g_1 + \cdots + \sigma_l g_l = 0 \qquad (\varrho_i, \sigma_j \in R). \tag{5}$$

Going over to the factor module G/H we get

$$\sigma_1 \overline{g}_1 + \cdots + \sigma_l \overline{g}_l = \overline{0}$$

and by the independence of the system $[\overline{g_\mu}]_{\mu \in \Gamma}$ this implies

$$\sigma_1 = \cdots = \sigma_l = 0.$$

Thus equality (5) reduces to

$$\varrho_1 h_1 + \cdots + \varrho_k h_k = 0$$

and making use of the independence of the system $[h_\nu]_{\nu \in \Delta}$ we obtain

$$\varrho_1 = \cdots = \varrho_k = 0.$$

We see that the system S is independent.

We have still to prove the maximality of the system S.

Let g be a principal element of order 0 of the module G. Then a relation of the form

$$\varrho g = \tau_1 g_1 + \cdots + \tau_p g_p + h \tag{6}$$

holds, where $\varrho \neq 0$ and $h \in H$. If h is not an element of order 0, i. e., if $\sigma h = 0$ for some element $\varrho (\neq 0)$, then

$$\sigma \varrho g = \sigma \tau_1 g_1 + \cdots + \sigma \tau_p g_p. \tag{7}$$

If h has order 0, then clearly $Rh \cong Rg$, so h is also principal and therefore a relation of the form

$$\varrho' h = \vartheta_1 h_1 + \cdots + \vartheta_q h_q \qquad (\varrho' \neq 0) \tag{8}$$

holds. Multiplying equation (6) by ϱ' and substituting on the basis of (8) into the equality so obtained we get the equation

$$\varrho' \varrho g = \vartheta_1 h_1 + \cdots + \vartheta_q h_q + \varrho' \tau_1 g_1 + \cdots + \varrho' \tau_p g_p. \tag{9}$$

Taking into account that g is a principal element of order 0, we get $\sigma \varrho g \neq 0$ and $\varrho' \varrho g \neq 0$, and so, by (7) and by (9), the system $[S, g]$ is in both cases non-independent. This completes the proof of Theorem 4.

Literature

[1] L. Fuchs, Ranks of modules, *Annales Univ. Sci. Budapest*, Sectio Math., **6** (1963), 71–78.

[2] A. Kertész, A remark on the general theory of modules, *Publ. Math. Debrecen*, **6** (1959), 86–89.

[3] A. Kertész, Vizsgálatok az operátormodulusok általános elméletében, I., *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.*, **8** (1958), 411–436.

[4] A. Kertész, On independent sets of elements in algebra, *Acta Sci. Math. Szeged*, **21** (1960), 260–269.

АСИМПТОТИЧЕСКАЯ ФОРМУЛА В ТЕОРИИ ЧИСЕЛ

И. КАТАИ*

Кафедра алгебры Университета им. Л. Этвеша, Будапешт

(Поступила 29. V. 1963.)

1. Обозначения. Примем обозначения:

n, m — натуральные числа

p, q — простое

(1,1) $$S(n) = \sum_{p<n} \frac{1}{n-p}$$

(1,2) $$\Lambda(n) = \begin{cases} \log p & \text{если} \quad n = p^k \\ 0 & \text{иначе} \end{cases}$$

(1,3) $$S_1(n) = \sum_{m=2}^{n-1} \frac{\Lambda(m)}{n-m}$$

$$s = \sigma + it$$

$\xi(s)$ — ξ — функция Римама.

$N(\sigma_0, T)$ — число нулеи $\xi(s)$ в прямоугольнике $\sigma_0 \leqq \sigma \leqq 1, |t| \leq T$.

$$z = x + iy, \qquad 0 < x < 1, \qquad -\pi \leq y \leq \pi$$

(1,4) $$f_1(z) = \sum_{n=1}^{\infty} \Lambda(n) e^{-nz}$$

(1,5) $$f_2(z) = \sum_{n=1}^{\infty} \frac{1}{n} e^{-nz}.$$

Легко убедиться, что

(1,6) $$f_2(z) = \log(1 - e^{-z}),$$

где log означает главное значение.

ϱ — переменное, пробегающее нуля функций $\xi(s)$.

$$T(z) = \sum z^{-\varrho} \Gamma(\varrho)$$

* I. Kátai

(1,7) $$\Delta = \frac{1}{\left(\log \frac{1}{x}\right)^7}$$

2. Нетурдно видеть, что

(2,1) $$\sum_{n \leq u} S(n) = u + o(u)$$

П. Эрдеш и Н. Г. де Брейн доказали, что

(2.2) $$c_1 u < \sum_{n \leq u} S^2(n) < c_2 u,$$

где c_1, c_2 положительные постоянные [1].

3. Имеет место следующая условная теорема.

ТЕОРЕМА. Если в случае $\frac{1}{2} \leq \sigma \leq 1$ имеет место $N(\sigma, T) = O(T^{2(1-\sigma)} \log^2 T)$, — то

(3,1) $$\sum_{n \leq u} S^2(n) = u + o(u).$$

Доказательство.
Применяя формулу-Парсеваля

(3,2) $$g(x) \overset{\text{def}}{=} \sum_{n=1}^{\infty} S_1^2(n) e^{-2nx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f_1(z) f_2(z)|^2 dy.$$

Непосредственно получаем из (1, 6), что имеет место

(3,3) $$|f_2(z)|^2 = \frac{1}{4} \log^2 (1 - 2e^{-x} \cos y + e^{-2x}) + O(1).$$

Сдругои стороны доказал Ю. В. Линник (см. [2], [3]), что

(3,4) $$f_1(z) = \frac{1}{z} - T(z) + O\left(\log^3 \frac{1}{x}\right).$$

Тогда

(3,5) $$|f_1(z)|^2 = \frac{1}{|z|^2} - \left(\frac{T(z)}{\bar{z}} + \frac{T(\bar{z})}{z}\right) + |T(z)|^2 + O\left(\log^3 \frac{1}{x} \cdot |f_1(z)|\right) + O\left(\log^6 \frac{1}{x}\right).$$

Рассмотрим теперь интеграл

$$g(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f_1(z) f_2(z)|^2 dy$$

и разобъем на сумму интегралов

$$g(x) = = I_1 + I_2 + I_3,$$

где

(3,6) $$I_1 = \frac{1}{2\pi}\int_{-\Delta}^{\Delta} |f_1(z)f_2(z)|^2 dy$$

(3,7) $$I_2 = \frac{1}{2\pi}\int_{\Delta}^{\pi} |f_1(z)f_2(z)|^2 dy$$

(3,8) $$I_3 = \frac{1}{2\pi}\int_{-\pi}^{-\Delta} |f_1(z)f_2(z)|^2 dy.$$

Если $|y| \geqq \Delta$, то $|f_2(z)| = O\left(\log\log\frac{1}{x}\right)$, откуда следует

$$I_2, I_3 = O\left(\log\log\frac{1}{x}\right)\int_{-\pi}^{\pi} |f_1(z)|^2 dy.$$

Из асимптотического закона простых чисел известно, что

$$\int_{-\pi}^{\pi} |f_1(z)|^2 dy \sim \frac{1}{2x}\log\frac{1}{x}.$$

Из этово соотношения следует

(3,9) $$I_2, I_3 = O\left(\frac{1}{x}\log\frac{1}{x}\cdot\log\log\frac{1}{x}\right).$$

Применяя соотношения (3,3) и (3,5) получаем

$$I_1 = O(J_1) + O\left(\log^6\frac{1}{x}\right)\cdot J_2 + O\left(\log^3\frac{1}{x}\right)\cdot J_3 + O\left(\log^3\frac{1}{x}\right)J_4 +$$

(3,10) $$+ O(J_5) + O(J_6) + J_7,$$

где

(3,11)–(3,12) $$J_1 = \int_{-\Delta}^{\Delta} |f_1(z)|^2 dy, \qquad J_2 = \int_{-\Delta}^{\Delta} |f_2(z)|^2 dy$$

(3,13)–(3,14) $$J_3 = \int_{-\Delta}^{\Delta} |f_1(z)||f_2(z)|^2 dy, \qquad J_4 = \int_{-\Delta}^{\Delta} |f_1(z)| dy$$

(3,15) $$J_5 = \int_{-\Delta}^{\Delta} |T(z)|^2 \log^2(1 - 2e^{-x}\cos y + e^{-2x}) dy$$

(3,16) $$J_6 = \int_{-\Delta}^{\Delta} \left|\frac{T(z)}{z}\right| \cdot \log^2(1-2e^{-x}\cos y + e^{-2x})dy$$

(3,17) $$J_7 = \frac{1}{8\pi}\int_{-\Delta}^{\Delta} \frac{1}{|z|^2}\log^2(1-2e^{-x}\cos y + e^{-2x})dy.$$

Очевидно имеет место

(3,18) $$J_1 = O\left(\frac{1}{x}\log\frac{1}{x}\right).$$

Так как

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|f_2(z)|^2dy = \sum\frac{1}{n^2}e^{-2nx},$$

то мы получаем

(3,19) $$J_2 = O(1).$$

Применяя неравенство-Буняковсково

$$\int_{-\Delta}^{\Delta}|f_1(z)||f_2(z)|^2dy \leq \left[\int_{-\Delta}^{\Delta}|f_1(z)|^2|f_2(z)|^2dy\right]^{1/2}\left[\int_{-\Delta}^{\Delta}|f_2(z)|^2dy\right]^{1/2} = O\left(\frac{1}{x^{1/2}}\cdot\log\frac{1}{x}\right)$$

откуда находим

(3,20) $$J_3 = O\left(\frac{1}{x^{1/2}}\log\frac{1}{x}\right).$$

Далее имеет место

(3,21) $$J_4 = O\left(\Delta^{1/2}\cdot\frac{1}{x^{1/2}}\log\frac{1}{x}\right) = O\left(\frac{1}{x^{1/2}}\right).$$

Ю. В. Линник доказал, что

(3,22) $$\int_{-\Delta}^{\Delta}|T(z)|^2dy = O\left(\frac{1}{x}\log^{-1}\frac{1}{x}\right).$$ [4]

откуда

(3,23) $$J_5 = O\left(\log^2\frac{1}{x}\right)\int_{-\Delta}^{\Delta}|T(z)|^2dy = O\left(\frac{1}{x}\log\frac{1}{x}\right).$$

Подобным образом получаем

(3,24) $$J_6 = \left[\int_{-\Delta}^{\Delta}\frac{1}{|z|^2}dy\right]^{1/2}\left[\int_{-\Delta}^{\Delta}|T(z)|^2dy\right]\cdot O\left(\log^2\frac{1}{x}\right) = O\left(\frac{1}{x}\cdot\log^{\frac{3}{2}}\frac{1}{x}\right).$$

Так как

$$\log^2(1-2e^{-x}\cos y + e^{-2x}) = \log^2(1-2e^{-x}+e^{-2x}) + \\ + O\left[\log\left(1+\frac{y}{x}\right)\cdot\log\frac{1}{x}\right] + O\left[\log^2\left(1+\frac{y}{x}\right)\right],$$

то

$$J_7 = \frac{\log^2(1-e^{-x})^2}{8\pi}\int_{-\infty}^{\infty}\frac{1}{x^2+y^2}\,dy + O\left(\int_{\Delta}^{\infty}\frac{1}{x^2+y^2}\,dy\cdot\log^2\frac{1}{x}\right)+$$

$$+O\left(\frac{1}{x}\log\frac{1}{x}\right) = \frac{1}{2x}\log^2\frac{1}{2x}+O\left(\frac{1}{x}\log\frac{1}{x}\right).$$

Применяя формулы (3,9), (3,10),(3,18)–(3,24) непосредственно получаем ассимптотическую формулу

$$g(x) \sim \frac{1}{2x}\log^2\frac{1}{2x}.$$

Из этово известным образом следует

$$\sum_{n\leq u} S_1^2(n) \sim u\log^2 u,$$

и отсюда непосредственно получаем теорему.

Если было бы доказать, что $\overline{\lim\limits_{n\to\infty}}\, S(n) = \infty$, то $\lim\limits_{n\to\infty}\dfrac{P_{n+1}-P_n}{\log P_n} = 0$, где P_n обозначает n-ное простое число.

Литература

[1] K. Prachar, *Primzahlverteilung*, Springer Verlag, 1957. 175.

[2] Ю. В. ЛИННИК, Некоторые условные теоремы, касающиеся бинарных задач с простыми числами, *Доклды Ак. Наук, СССР*, **77** (1951), 15–18.

[3] Ю. В. ЛИННИК, Простые числа и степени двойки, *Труды Мат. института Ак. Наук им. В. А. Стеклова*, **38** (1951), 152–169.

[4] Ю.В. ЛИННИК, Некоторые условные теоремы, касающиеся бинарной проблемы Гольдбаха, *Известия Ак. Наук СССР*, **16** (1952) 503–520.

GENERALIZED EULER SUMS AND THE SEMIGROUP PROPERTY OF INTEGRO-DIFFERENTIAL OPERATORS

By

M. MIKOLÁS

Department of Math. Analysis of the Eötvös Loránd University, Budapest

(Received June 8, 1963)

1. Some well-known attempts in finding such a common generalization of ordinary derivatives and repeated integrals which extends immediately the elementary limit definition of these fundamental ideas, failed mainly because of unsuitable treatment of the arising limit problems of more fine character and in consequence of a certain inevitable ambiguity.[1] Not long ago, STULOFF [10] and MOPPERT [8] gave rigorous considerations on the question and got interesting partial results.

In [10], a definition of fractional derivative (due to LIOUVILLE) is used in case of monotone and totally monotone functions, resp., which permits the application of recent theorems connected with fractional differences (KNOPP).[2] The starting-point of [8] is somewhat more general. The author defines the differential quotient of order α (arbitrary real) by

$$D^\alpha f(\xi)\Big|_{x_0}^{x_1}(\xi) = \lim_{n\to\infty}\left(\frac{n}{x_1-x_0}\right)^\alpha \sum_{\nu=0}^{n-1}(-1)^\nu\binom{\alpha}{\nu} f\left(x_1-\nu\frac{x_1-x_0}{n}\right) \tag{1.1}$$

as an interval function over $(x_0, x_1]$[3] and shows a. o. that 1. $D^\alpha f(\xi)\Big|_{x_0}^{x_1}(\xi)$ can be expressed for $\alpha < 0$ as an integral, for $\alpha > 0$ as an ordinary derivative, provided that $f(\xi)$ is differentiable or it has a continuous derivative of sufficiently high order in $x_0 < \xi \leqq x_1$ with an O-restriction for $\xi \to x_0$, respectively; 2. the so-called *semigroup property*[4] of the operator D^α, i. e.

$$D^\alpha D^\beta = D^\beta D^\alpha = D^{\alpha+\beta} \tag{1.2}$$

holds not only for $\alpha \leqq 0$, $\beta \leqq 0$, but also for $\alpha < 1$, $\beta < 1$, if $f(\xi)$ is in $[x_0, x_1]$ continuously derivable and vanishes for $\xi = x_0$.

Now, the aim of this paper is to investigate more closely the sums of the type (1.1) ("*generalized Euler sums*") and, on this basis, to sharpen and to develop further some of the last-mentioned results. First of all, we compare the idea of the generalized derivative (1.1) ("*integro-derivative*") for $\alpha < 0$ or $\alpha =$

[1] See e. g. Enzykl. der math. Wiss. II. 1/1, pp. 116–119; for the bibliography cf. [6], too.

[2] It may be remarked, however, that the definition of the generalized difference as a series $\left[\Delta^\alpha x_n = \sum_{\nu=0}^{\infty}(-1)^\nu\binom{\alpha}{\nu}x_{n+\nu}\right]$ leads to unnecessary restrictions of convergence.

[3] For $\alpha = 0, 1, 2, \ldots$, (1.1) depends only on the behaviour of $f(\xi)$ in an arbitrarily small neighbourhood of x_1.

[4] Cf. [3], pp. 439–443. – Another denomination of (1.2) is: "index law".

$= 0, 1, 2, \ldots$ with that of the Riemann integral and $f^{(r)}(x)$ $(r = 0, 1, \ldots)$; it can be pointed out that (1.1) is, in fact, a common extension of $\int_a^x f(t)dt$, of the Riemann–Liouville integral (2.9) and of the ordinary derivatives (Theorem I.). On the other hand, we find a group property of the generalized Euler sums as operators (Theorem II.); this renders possible to give the necessary and sufficient condition for the validity of (1.2) (Theorem III.) and hence a. o. a theorem of inclusion on the existence of integro-derivatives of arbitrary order (cf. Theorems IV.–V.). Though we limit ourselves to the case of real corder and of real functions, no obstacle is to the treatment in the complex field. Finally, some remarks will be taken concerning applications to generalized difference and integro-differential equations (cf. Theorem VI.); the detailed discussion will be published elsewhere.

2. In what follows, x, u, t will denote allways real variables, $a \neq x$ and μ real parameters.

Let $f(u)$ be a function defined at all points of the interval $a < u \leqq x$, for which the ratio $(u, x, a) = (u-a)/(x-a)$ is rational;[5] this enumerable set will be denoted by $S_{a,x}$. We write ($n > 0$, integer)

[5] We have plainly $0 < (u, x, a) < 1$.

$$(2.1) \qquad {}_a\Delta_n^\mu f(u)|x = {}_a\Delta_n^\mu f|x = \sum_{p=0}^{n-1} (-1)^p \binom{\mu}{p} f\left(x - p\frac{x-a}{n}\right),$$

so that, in particular,

$$(2.2) \qquad {}_a\Delta_n^0 f|x = {}_a\Delta_1^\mu f|x = f(x)$$

and

$$(2.3) \qquad \begin{cases} {}_a\Delta_n^r f|x = f(x) - \binom{r}{1} f(x-\omega_n) + - \cdots + (-1)^r \binom{r}{r} f(x - r\omega_n) \\ \qquad (n > r; \qquad r = 1, 2, \ldots) \end{cases}$$

with $\omega_n = (x-a)/n$.

Since (2.4) is of the type $\sum_{p=0}^{n-1} F(p)$, i. e. a sum which occurs in the Euler–Maclaurin summation formula, ${}_a\Delta_n^\mu f|x$ will be called a *generalized Euler sum* of order μ over $(a, x]$.[6] ((2.3) shows that it may be regarded simultaneously as a generalized difference of order μ.)

If for fixed a and μ

$$(2.5) \qquad {}_af^{(\mu)}(u)|x = {}_af^{(\mu)}|x = \lim_{n\to\infty} \left(\frac{x-a}{n}\right)^{-\mu} {}_a\Delta_n^\mu f|x$$

[6] The extensions in question are, of course, rather strong; e. g. the function of Dirichlet (taking 1 or 0 according to rational or irrational x, resp.) is for all x ${}_0D^1$-derivable and ${}_0D^{-1}$-derivable.

exists, we say that $f(u)$ is *${}_aD^\mu$-derivable* at the point x; for (2.5) we use the denomination: *integro-derivative of $f(u)$* of order μ over $(a, x]$ and apply the signs ${}_aD^\mu f(u)|x$, ${}_aD^\mu f|x$, ${}_aD^\mu f$, too. By (2.1)–(2.4), it is evident that ${}_af^{(\mu)}|x$ depends in general on the values of $f(u)$ between a and x; more precisely, ${}_af^{(\mu)}|x$ is determined by the values $f(u)$ over $S_{a,x}$ when $\mu \neq 0, 1, \ldots,$ while

$$(2.6) \qquad {}_af^{(0)}|x = f(x)$$

and

$$(2.7) \qquad \left\{ {}_af^{(r)}|x = \lim_{n\to\infty} \frac{f(x) - \binom{r}{1} f(x-\omega_n) + - \cdots + (-1)^r \binom{r}{r} f(x-\omega_n)}{\omega_n^r} \quad (r = 1, 2, \ldots) \right.$$

depend only upon the behaviour of the function f in $S_{a,x} \cap V_\delta$, where V_δ means an arbitrarily small neighbourhood of x.

As is to see, ${}_af^{(1)}|x = f'(x)$ and ${}_af^{(-1)}|x$ equals the *Riemann integral* of $f(u)$ from a to x, provided that $f(u)$ is differentiable at x (at least from the left), or it is integrable in Riemann's sense in the interval $(a, x]$, respectively.[6] More generally, we can state

THEOREM I. 1. *Let $r>0$, integer. The existence of $f^{(r)}(x)$ implies the ${}_aD^r$-derivability of $f(u)$ at x for all a in a neighbourhood of x and*

$$(2.8) \qquad {}_af^{(r)}|x = f^{(r)}(x).^7$$

2. *Let $\mu<0$. If the function $f(u)$ is defined in $(a, x]$, furthermore $f(u)(x-u)^{-(\mu+1)}$ is bounded and almost everywhere continuous there, then $f(u)$ is ${}_aD^\mu$-derivable at x, namely*

$$(2.9) \qquad {}_af^{(\mu)}|x = \Gamma(-\mu)^{-1} \int_a^x f(t)(x-t)^{-(\mu+1)} dt^8$$

with a Riemann integral on the right.

In particular, (2.9) *holds for any negative μ, when $f(u)$ is bounded and p. p. continuous in $(a, x]$.*

PROOF. 1. Suppose the existence of $f^{(r)}(x)$, so that $f^{(k)}(u)$ $(k = 0, 1, \ldots, r-1)$ are defined in a neighbourhood V of x.

Differentiating the function

$$(2.10) \qquad \Psi(t) = f(x) - \binom{r}{1} f(x+t) + - \cdots + (-1)^r \binom{r}{r} f(x+rt)$$

r-times, we obtain

$$(2.11) \qquad \Psi^{(k)}(0) = f^{(k)}(x) \cdot Q_{(k)}(-1) \qquad (k = 1, 2, \ldots, r)$$

[7] This can be regarded as an improvement of a well-known result, where the *continuity* of $f^{(r)}(x)$ is also required. Cf. e. g. [5], p. 89.

[8] The second term is the well-known *fractional integral of Riemann – Liouville.*

with

$$Q_{(k)}(u) = \binom{r}{1}u + 2^k\binom{r}{2}u^2 + \cdots + r^k\binom{r}{r}u^r.$$

By the recursion formula

$$Q_{(k+1)}(u) = uQ'_{(k)}(u),$$

induction shows that

$$Q_{(k)}(u) = \sum_{\nu=r-k}^{r} c_\nu(u+1)^\nu; \qquad c_k = r^k, \ldots, c_{r-k} = (-1)^k r(r-1)\cdots(r-k+1)$$

and hence

$$Q_{(k)}(-1) = \begin{cases} 0 & \text{for } k = 1, 2, \ldots, r-1; \\ (-1)^r r! & \text{for } k = r. \end{cases} \tag{2.12}$$

Using now (2.11), (2.12) and a well-known extension of L'Hospital's rule, we get

$$\lim_{t\to 0}\frac{\Psi(t)}{t^r} = \frac{\Psi^{(r)}(0)}{(t^r)^{(r)}_{t=0}} = (-1)^r f^{(r)}(x),$$

which implies plainly (2.8) (cf. (2.7)) for every $a \in V$.

2. The Gaussian definition fo the gamma-function gives

$$\begin{cases} \binom{u+n}{n} = \dfrac{(u+1)(u+2)\cdots(u+n)}{n!} \sim n^u\Gamma(u+1)^{-1} \qquad (n\to\infty); \\ \qquad (u \neq -1, -2, \ldots) \end{cases} \tag{2.13}$$

more precisely, by the Euler–Maclaurin sum-formula we have

$$\binom{u+n}{n} = \Gamma(u+1)^{-1}n^u\left(1+\frac{u}{n}\right)^{u+n+\frac{1}{2}} \exp\left[u\cdot\int_n^\infty \frac{t-[t]-\frac{1}{2}}{t(u+t)}dt - u\right]^{9} \tag{2.14}$$

and hence

$$\begin{cases} \binom{n+u}{n} = \Gamma(u+1)^{-1}n^u + \Gamma(u)^{-1}\left\{(1+\vartheta_n)u+\frac{1}{2}\right\}n^{u-1} + K_{n,u}\cdot n^{u-2} \\ \qquad (n = 1, 2, \ldots) \end{cases} \tag{2.15}$$

with $|\vartheta_n| < 1$ and $|K_{n,u}| \leqq K_u$ (a constant depending upon u only).

We need especially the asymptotic formula ($u = -\mu - 1$):

$$\begin{cases} (-1)^n\binom{\mu}{n} = \binom{n-(\mu+1)}{n} = \Gamma(-\mu)^{-1}n^{-(\mu+1)} + O(n^{-(\mu+2)}) \quad (n\to\infty), \\ \qquad (\mu \neq 0, 1, 2, \ldots) \end{cases} \tag{2.16}$$

where the constant implied by the O-symbol depend, of course, on u.

[9] Cf. e. g. [4], pp. 548–549. – Here exp x stands for e^x.

Let now $\mu < 0$ and suppose $|f(u)(x-u)^{-(\mu+1)}| \leqq B_\mu$ (constant) for $u \in (a, x]$; the condition that the function $f(u)(x-u)^{-(\mu+1)}$ is p. p. continuous being equivalent to its Riemann integrability (*Lebesgue*'s criterion), we have in case of an equidistant division of (a, x) $(\omega_n = (x-a)/n)$

$$(2.17)\qquad \int_a^x f(t)(x-t)^{-(\mu+1)}dt = \lim_{n\to\infty} \sum_{p=1}^{n-1} [f(t)(x-t)^{-(\mu+1)}]_{t=x-p\omega_n} \cdot \omega_n =$$

$$= \lim_{n\to\infty} \omega_n^{-\mu} \sum_{p=1}^{n-1} p^{-(\mu+1)} f(x-p\omega_n).$$

On the other hand, the corresponding "approximating sum" of ${}_af^{(\mu)}|x$ may be written [cf. (2.5), (2.16)]

$$\omega_n^{-\mu}\,{}_a\Delta_n f|x = \omega_n^{-\mu} f(x) + \Gamma(-\mu)^{-1}\omega_n^{-\mu} \sum_{p=1}^{n-1} p^{-(\mu+1)} f(x-p\omega_n) +$$

$$+ \omega_n^{-\mu} \sum_{p=1}^{n-1} A_{p,\mu}\, p^{-(\mu+2)} f(x-p\omega_n)$$

with $|A_{p,\mu}| \leqq A_\mu$ (constant) for $p = 1, 2, \ldots$.

Since the last term permits the estimation

$$(2.19)\qquad \left|\omega_n^{-\mu} \sum_{p=1}^{n-1} A_{p,\mu}\, p^{-(\mu+2)} f(x-p\omega_n)\right| \leqq A_\mu B_\mu \omega_n \sum_{p=1}^{n-1} \frac{1}{p} = O\left(\frac{\log n}{n}\right).$$

(2.9) follows from (2.17)–(2.18) at once.

3. Let us mention here an immediate corollary of (2.16), namely the calculation of the integro-derivative of a constant.

In fact, by the elementary formula $\binom{\mu}{p} = \binom{\mu-1}{p} + \binom{\mu-1}{p-1}$ $(p = 1, 2, \ldots)$ we get

$$(3.1)\qquad {}_a\Delta_n^\mu 1 = \sum_{p=0}^{n-1} (-1)^p \binom{\mu}{p} = 1 + \sum_{p=1}^{n-1} (-1)^p \binom{\mu-1}{p} -$$

$$- \sum_{p=0}^{n-2} (-1)^p \binom{\mu-1}{p} = (-1)^n \binom{\mu-1}{n-1}$$

and hence (2.16) yields

$$\lim_{n\to\infty} \left(\frac{x-a}{n}\right)^{-\mu} {}_a\Delta_n^\mu 1 = \lim_{n\to\infty} \left(\frac{x-a}{n}\right)^{-\mu} \cdot \Gamma(1-\mu)^{-1} n^{-\mu} = \Gamma(1-\mu)^{-1}(x-a)^{-\mu},$$

i. e. for the function $f(u) \equiv c$ (constant):

$$(3.2)\qquad \begin{cases} {}_a(c)^{(\mu)}|x = c\Gamma(1-\mu)^{-1}(x-a)^{-\mu} \\ (a \text{ and } \mu \text{ arbitrary, } x > a) \end{cases}$$

In particular, for $\mu = 1, 2, \ldots$ the Γ-factor vanishes, so that

$$ {}_a(c)^{(r)}|x \equiv 0 \qquad (r = 1, 2, \ldots) $$

in accordance with (2.8).

As far as the properties of the *operators* ${}_a\Delta_n^\mu$ and ${}_aD^\mu$ is concerned, we remark first of all that both are plainly *linear*. — Next we show, that every operator of the form ${}_a\Delta^{\mu_1+\mu_2}$ can be generated by successive application of ${}_a\Delta^{\mu_1}$ and ${}_a\Delta^{\mu_2}$; more exactly, it holds the following *group property of generalized Euler sums:*

THEOREM II. *For every function $f(u)$ defined in $S_{a,x}$ and for arbitrary μ_1, μ_2*

$$ {}_a\Delta_n^{\mu_1}({}_a\Delta_{[n(u,x,a)]}^{\mu_2} f|u)|x = {}_a\Delta_n^{\mu_1+\mu_2} f|x \qquad (n = 1, 2, \ldots)^{10} \tag{3.4} $$

PROOF. Owing to the definition (2.1), we have

$$ \left\{ \begin{aligned} & {}_a\Delta_n^{\mu}({}_a\Delta_{[n(u,x,a)]}^{\mu_2} f|u)|x = \\ & = \sum_{p=0}^{n-1} (-1)^p \binom{\mu_1}{p} \left\{ \sum_{q=0}^{[n(u,x,a)]-1} (-1)^q \binom{\mu_2}{q} f\left(u - q\,\frac{u-a}{[n(u,x,a)]}\right) \right\}_{u=x-p\frac{x-a}{n}} \end{aligned} \right. \tag{3.5} $$

and this, by

$$ [n(u,x,a)]_{u=x-p\frac{x-a}{n}} = \left[n\left(x - p\frac{x-a}{n} - a\right)(x-a)^{-1} \right] = n - p, \tag{3.6} $$

equals to

$$ \sum_{p=0}^{n-1} \sum_{q=0}^{n-p-1} (-1)^{p+q} \binom{\mu_1}{p} \binom{\mu_2}{q} f\left(x - (p+q)\frac{x-a}{n}\right). \tag{3.7} $$

Putting now $p+q = k$ and taking into consideration the so-called *addition formula of the binomial coefficients:*

$$ \binom{\mu_1}{0}\binom{\mu_2}{k} + \binom{\mu_1}{1}\binom{\mu_2}{k-1} + \cdots + \binom{\mu_1}{k}\binom{\mu_2}{0} = \binom{\mu_1+\mu_2}{k}, \tag{3.8} $$

(3.5) becomes

$$ \sum_{k=0}^{n-1} (-1)^k \binom{\mu_1+\mu_2}{k} f\left(x - k\frac{x-a}{n}\right) = {}_a\Delta_n^{\mu_1+\mu_2} f|x. \tag{3.9} $$

Qu. e. d.

(3.4) suggests for $n \to \infty$, that the operator ${}_aD^\mu$ satisfies the relation

$$ {}_aD^{\mu_1}({}_aD^{\mu_2} f|u)|x = {}_aD^{\mu_1+\mu_2} f|x \tag{3.10} $$

or, with other denotations,

$$ {}_a({}_af^{(\mu_2)}|u)^{(\mu_1)}|x = {}_af^{(\mu_1+\mu_2)}|x. \tag{3.11} $$

[10] Here and below, $[t]$ denotes the greatest integer $\leq t$.

This is indeed true for $\mu_1 \leqq 0$, $\mu_2 \leqq 0$, since, by (2.9), one obtains then the well-known *semigroup property* of the Riemann–Liouville integrals.[11]

Whereas besides, in case of μ_1, μ_2 positive integers, (3.11) corresponds to the forming of ordinary derivatives by repeated differentiation, it cannot hold in full generality and the operators ${}_aD^{\mu_1}$, ${}_aD^{\mu_2}$ must not be commutative, too. (The simplest counter-example is: $\mu_1 = 1$, $\mu_2 = -1$ and $\mu_1 = -1$, $\mu_2 = 1$.)

By (3.4), however, we can easily establish the exact condition of validity for (3.10)–(3.11).

THEOREM III. *Suppose that* ${}_af^{(\mu_2)}|u$ *is known for* $u \in S_{a,x}$ *and one of the integro-derivatives*

$$ {}_a({}_af^{(\mu_2)}|u)^{(\mu_1)}|x, \qquad {}_af^{(\mu_1+\mu_2)}|x \tag{3.12}$$

exists. In order that the other of these derivatives should also exist and both should be equal, it is necessary and sufficient that

$$\lim_{n\to\infty} (n^{\mu_1} \cdot {}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)|x) = 0 \tag{3.13}$$

with

$$\eta_m(u) = \left(\frac{u-a}{m}\right)^{-\mu_2} {}_a\Delta_m^{\mu_2} f|u - {}_af^{(\mu_2)}|u \qquad (m = 1, 2, \ldots),^{12} \tag{3.14}$$

PROOF. Considering the linearity of the operator ${}_a\Delta_n^u$, (3.4) is equivalent to the formula

$$\left(\frac{x-a}{n}\right)^{-(\mu_2+\mu_2)} {}_a\Delta_n^{\mu_1+\mu_2} f|x = \left(\frac{x-a}{n}\right)^{-\mu_1} {}_a\Delta_n^{\mu_1}({}_af^{(\mu_2)}|u)|x =$$

$$= \left(\frac{x-a}{n}\right)^{-\mu_1} {}_a\Delta_n^{\mu_1}\left\{\left(\frac{x-a}{n}\right)^{-\mu_2} {}_a\Delta_{[n(u,x,a)]}^{\mu_2} f|u - {}_af^{(\mu_2)}|u\right\}\Big|x$$

and we may substitute $\eta_{[n(u,x,a)]}(u)$ for the last difference in round brackets by (cf. (3.6))

$$\left(\frac{u-a}{[n(u,x,a)]}\right)_{u=x-p\frac{x-a}{n}} = \frac{x-a}{a} \qquad (p = 0, 1, \ldots, n-1).$$

Therefore

$$\left|\begin{aligned} &\lim_{n\to\infty}\left\{\left(\frac{x-a}{n}\right)^{-(\mu_1+\mu_2)} {}_a\Delta_n^{\mu_1+\mu_2} f|x - \left(\frac{x-a}{n}\right)^{-\mu_1} {}_a\Delta_n^{\mu_1}({}_af^{(\mu_2)}|u)|x\right\} = \\ &\qquad = \lim_{n\to\infty}\left(\frac{x-a}{n}\right)^{-\mu_1} {}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)|x \end{aligned}\right. \tag{3.15}$$

and the result follows from the definition of the integro-derivatives (3.12).

[11] For a simple proof see e. g. [7], p. 74. – Cf. also [2], p. 309 or [9], pp. 10–11.

[12] Observe that, by hypothesis, $\lim_{m\to\infty} \eta_m(u) = 0$ for any fixed $u \in S_{a,x}$.

4. Before going further, it is useful to introduce some shortening terms. We say that $f(u)$ is *uniformly* ${}_aD^\mu$*-derivable* in the interval $(a, x]$, if the sequence of functions

$$\left(\frac{u-a}{n}\right)^{-\mu} {}_a\Delta_n^\mu f|u \quad (n = 1, 2, \ldots)$$

is uniformly convergent for $u \in (a, x]$. Moreover, instead of writing a condition of the form

$$\left(\frac{u-a}{n}\right)^{-\mu} {}_a\Delta_n^\mu f|u - {}_af^{(\mu)}|u = o(n^\gamma) \qquad (\gamma < 0, \text{ constant}; \ n \to \infty)$$

uniformly in $(a, x]$, we use the expression: ${}_af^{(\mu)}|u$ *exists* for $u \in (a, x]$ *with the error-term* $o(n^\gamma)$.

It is clear (cf. (2.6)), that (3.11) is true for $\mu_1 = 0$ (or for $\mu_2 = 0$). For $\mu_1 \neq 0$, we deduce now from (3.13) quite general sufficient conditions.

THEOREM IV. 1. *Suppose that* (*) $\mu_1 < 0$ *and* $f(u)$ *is uniformly* ${}_aD^{\mu_2}$*-derivable in* $(a, x]$, *or* (**) $\mu_1 > 0$, *not integer and* ${}_af^{(\mu_2)}(u)$ *exists for* $u \in (a, x]$ *with an error-term* $o(n^{-\mu_1})$. *Then the existence of*

$$(4.1) \qquad {}_a({}_af^{(\mu_2)}|u)^{(\mu_1)}|x$$

implies that of

$$(4.2) \qquad {}_af^{(\mu_1+\mu_2)}|x$$

and inversely; furthermore (3.11) *is valid provided that at least one of* (4.1)–(4.2) *exists.*

2. *If* $\mu_1 = 1, 2, \ldots$ *and the existence of* ${}_af^{(\mu_2)}|u$ *with the error* $o(n^{-\mu_1})$ *is required only in a left-hand neighbourhood of the point* x, *then the same holds as before.*

PROOF. 1° Let us assume that the conditions (*) are fulfilled. By $\mu_1 < 0$, one has then

$$(4.3) \qquad (-1)^p \binom{\mu_1}{p} = (-\mu_1)\frac{1-\mu_1}{2}\cdots\frac{(p-1)-\mu_1}{2} > 0 \qquad (p = 1, 2, \ldots)$$

and the expression in (3.13) can be estimated as follows (cf. (3.6)):

$$(4.4) \qquad \left|\left(n^{\mu_1}{}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)|x\right)\right| \leqq n^{\mu_1}\sum_{p=0}^{n-1}(-1)^p\binom{\mu_1}{p}\left|\eta_{n-p}\left(x-p\frac{x-a}{n}\right)\right| = \Sigma_1 + \Sigma_2$$

with

$$(4.5) \qquad \begin{cases} \Sigma_1 = n^{\mu_1}\sum_{p=0}^{m}(-1)^p\binom{\mu_1}{p}\left|\eta_{n-p}\left(x-p\frac{x-a}{n}\right)\right| \\ \Sigma_2 = n^{\mu_1}\sum_{p=m+1}^{n-1}(-1)^p\binom{\mu_1}{p}\left|\eta_{n-p}\left(x-p\frac{x-a}{n}\right)\right| \end{cases} \qquad (0 < m < n).$$

Since, according to our hypothesis,

$$(4.6) \qquad \lim_{m\to\infty}\eta_m(u) = 0$$

uniformly in $(a, x]$, the least upper bound H of $|\eta_m(u)|$ for all m and $u \in (a, x]$ is finite; let $\varepsilon > 0$ be chosen such that

$$\varepsilon < \frac{4H}{\Gamma(1-\mu_1)}$$

and

$$\vartheta = \left(1 - \frac{\Gamma(1-\mu_1)}{4H}\varepsilon\right)^{-\frac{1}{\mu_1}} \tag{4.8}$$ [13]

Put now $m = [\vartheta n]$; by (4.6), there exists an integer $N_0 = N_0(\varepsilon) > 0$ for which

$$\left|\eta_{n-p}\left(x - p\frac{x-a}{n}\right)\right| < \frac{\Gamma(1-\mu_1)}{4}\varepsilon \quad (p = 0, 1, \ldots, [\vartheta n]; \quad n > N_0). \tag{4.9}$$

Thus the first sum under (4.5) (cf. (3.1), (2.16)):

$$\left\{\begin{aligned} \Sigma_1 &< n^{\mu_1}\frac{\Gamma(1-\mu_1)}{4}\varepsilon \sum_{p=0}^{[\vartheta n]}(-1)^p\binom{\mu_1}{p} = n^{\mu_1}\cdot(-1)^{[\vartheta n]}\binom{\mu_1-1}{[\vartheta n]}\cdot\frac{\Gamma(1-\mu_1)}{4}\varepsilon < \\ &< n^{\mu_1}\cdot 2\Gamma(1-\mu_1)^{-1}[\vartheta n]^{-\mu_1}\cdot\frac{\Gamma(1-\mu_1)}{4}\varepsilon = \frac{\varepsilon}{2} \end{aligned}\right.$$

whenever $n > N_1 = N_1(\varepsilon)\ (\geqq N_0)$.

On the other hand, we can write

$$\Sigma_2 < n^{\mu_1}\cdot H \sum_{p=[\vartheta n]+1}^{n-1}(-1)^p\binom{\mu_1}{p} = Hn^{\mu_1}\left\{(-1)^{n-1}\binom{\mu_1-1}{n-1} - (-1)^{[\vartheta n]}\binom{\mu_1-1}{[\vartheta n]}\right\} =$$

$$= H\left\{n^{\mu_1}\cdot(-1)^{n-1}\binom{\mu_1-1}{n-1} - \left(\frac{[\vartheta n]}{n}\right)^{-\mu_1}[\vartheta n]^{\mu_1}\cdot(-1)^{[\vartheta n]}\binom{\mu_1-1}{[\vartheta n]}\right\}.$$

The last difference in great brackets tends to $\Gamma(1-\mu_1)^{-1}(1-\vartheta^{-\mu_1})$ as $n \to \infty$ by (2.16), and so it does not exceed certainly the double of this value for n large enough; consequently

$$\Sigma_2 < 2H\Gamma(1-\mu_1)^{-1}(1-\vartheta^{-\mu_1}) = \frac{\varepsilon}{2}, \tag{4.11}$$

when $n > N_2 = N_2(\varepsilon)$.

Summing up, we get (cf. (4.4)–(4.5))

$$\left|\left(n^{\mu_1}{}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)\middle|x\right)\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{4.12}$$

for $n > N = \max(N_1, N_2)$, i. e. the limit relation (3.13); the assertion follows therefore by Theorem III.

2° In case of the conditions (**), one can argue similarly.

[13] Hence $0 < \vartheta < 1$. – As we know, $\Gamma(u) > 0$ for $u > 0$.

Next we remark that for $\mu_1 > 0$, not-integer, $p > \mu_1 + 1$

$$(4.13)\qquad \begin{cases} \operatorname{sg}(-1)^p\binom{\mu_1}{p} = (-1)^{[\mu_1]+1} \\ \left|(-1)^p\binom{\mu_1}{p}\right| = (-1)^{[\mu_1]+p+1}\binom{\mu_1}{p} \end{cases}$$

and apply the dissection

$$(4.14)\qquad \begin{cases} n^{\mu_1}{}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)|x = \Sigma' + \Sigma'', \\ \Sigma' = n^{\mu_1}\sum\limits_{p=0}^{m}(-1)^p\binom{\mu_1}{p}\eta_{n-p}\left(x - p\frac{x-a}{n}\right), \\ \Sigma'' = n^{\mu_1}\sum\limits_{p=m+1}^{n-1}(-1)^p\binom{\mu_1}{p}\eta_{n-p}\left(x - p\frac{x-a}{n}\right) \end{cases}$$

with $m = [\Theta n] > [\mu_1] + 1$, $0 < \Theta < 1$.

By the assumption that

$$(4.15)\qquad \lim_{m\to\infty} m^{\mu_1}\eta_m(u) = 0$$

uniformly for $a < u \leqq x$, the inequality

$$(4.16)\qquad |\eta_m(u)| < \delta m^{-\mu_1} \quad (\delta > 0, \text{ arbitrarily small})$$

holds in $(a, x]$ for sufficiently large m; hence

$$(4.17)\qquad \begin{cases} |\Sigma'| < n^{\mu_1}\sum\limits_{p=0}^{[\Theta n]}\left|\binom{\mu_1}{p}\right|\frac{\delta}{(n-p)^{\mu_1}} < \left(\frac{n}{n-[\Theta n]}\right)^{\mu_1} M\delta < (1-\Theta)^{-\mu_1}M\delta, \\ \qquad (n > N' = N'(\delta)) \end{cases}$$

where M means the sum of the series $1 + \left|\binom{\mu_1}{1}\right| + \left|\binom{\mu_1}{2}\right| + \ldots$[14]

As regards to $\sum''$, by (2.16), (3.1), (4.13) and

$$(-1)^{[\mu_1]}\Gamma(1-\mu_1) = |\Gamma(1-\mu_1)|$$

we obtain

$$(4.18)\qquad \begin{cases} |\Sigma''| \leqq n^{\mu_1}\sum\limits_{p=[\Theta n]+1}^{n-1}(-1)^{p+[\mu_1]+1}\left|\binom{\mu_1}{p}\right|\left|\eta_{n-p}\left(x - p\frac{x-a}{n}\right)\right| < \\ < H\cdot n^{\mu_1}(-1)^{[\mu_1]}\left\{(-1)^{[\Theta n]}\binom{\mu_1-1}{[\Theta n]} - (-1)^{n-1}\binom{\mu_1-1}{n-1}\right\} < \\ \qquad < 2H|\Gamma(1-\mu_1)|^{-1}(\Theta^{-\mu_1}-1) \end{cases}$$

for $n > N''$, H denoting $\sup|\eta_m(u)|$ over $(a, x]$ for all m.

[14] It is well-known, that the binomial series $\Sigma\binom{\alpha}{n}z^n$ converges absolutely for $\operatorname{Re}\alpha > 0$, $|z| = 1$. Cf. e. g. [4], p. 440.

Now, after being given a positive number ε, it is plainly possible to choose Θ in such a way (near to 1) that

$$2H|\Gamma(1-\mu_1)|^{-1}(\Theta^{-\mu_1}-1) < \frac{\varepsilon}{2}$$

and, for this fixed Θ, to find a $\delta > 0$ satisfying

$$(1-\Theta)^{-\mu_1}M\delta < \frac{\varepsilon}{2}.$$

As we see, (4.14), (4.17)–(4.18) yield again

$$\left|\left(n^{\mu_1}{}_a\Delta_n^{\mu_1}\eta_{[n(u,x,a)]}(u)|x\right)\right| < \varepsilon \quad (n > N^* = N^*(\varepsilon)).$$

3° If $\mu_1 = r > 0$, integer and $m^r\eta_m(u) \to 0$ as $m \to \infty$, uniformly in a neighbourhood $(x_0, x]$ of x with $a \leqq x_0 < x$, then for each $\delta > 0$, no matter how small, we shall have

$$|\eta_m(u)| < \frac{\delta}{m^r} \quad (u \in (x_0, x]; m > m_0 = m_0(\delta)). \tag{4,19}$$

Hence results the estimate:

$$\left\{\begin{aligned} &\left|\left(n^r{}_a\Delta_n^r\eta_{[n(u,x,a)]}(u)|x\right)\right| = n^r\left|\sum_{p=0}^{r}(-1)^p\binom{r}{p}\eta_{n-p}\left(x-p\frac{x-a}{n}\right)\right| < \\ &< \left(\frac{n}{n-r}\right)^r\delta\sum_{p=0}^{r}\binom{r}{p} \leqq (1+r)^r 2^r\delta, \end{aligned}\right. \tag{4.20}$$

whenever $n > \max\{m_0+r, r(x-a)/(x-x_0)\}$. Since the last bound in (4.20) may be also arbitrarily small, the condition (3.13) is certainly fulfilled. Qu. e. d.

5. The ordinary derivatives $f'(x), f''(x), \ldots$ taking their origin in iteration of the differentiation process, the existence of $f^{(r)}(x)$ implies, of course, that of $f^{(r-1)}(x)$ $(r = 1, 2, \ldots)$. But the integro-derivatives are defined independently of one another, so that the question may be raised: in what measure can we conclude from the existence of an ${}_aD^{\mu_0}$ to that of the ${}_aD^{\mu}$-derivatives with $\mu < \mu_0$? – The following result is a simple corollary of Theorem IV:

THEOREM V. *If $f(u)$ is uniformly ${}_aD^{\mu_0}$-derivable in the interval $(a, x]$, then it is at the same time ${}_aD^{\mu}$-derivable at the point x for every $\mu < \mu_0$.*

As a matter of fact: if ${}_af^{(\mu_0)}|u$ exists uniformly for $a < u \leqq x$, then also the existence of ${}_af^{(\mu)}|x = {}_af^{[\mu_0+(\mu-\mu_0)]}|x\,(\mu < \mu_0)$ is ensured by the first part of the previous theorem (case (*)); moreover, we have the connection:

$${}_af^{(\mu)}|x = {}_a({}_af^{(\mu_0)}|u)^{(\mu-\mu_0)}|x \quad (\mu < \mu_0). \tag{5.1}$$

Concerning the extension to complex variables and functions may be noted: if we define the integro-derivative of $f(w)$ of a *complex order* μ by

$${}_af^{(\mu)}(w)|z = {}_af^{(\mu)}|z = \lim_{n\to\infty}\left(\frac{z-a}{n}\right)^{-\mu}\sum_{p=0}^{n-1}(-1)^p\binom{\mu}{p}f\left(z-p\frac{z-a}{n}\right), \tag{5.2}$$

where the power to the exponent $(-\mu)$ is to be taken with its principal value and a means a given complex parameter, then ceases to exist any restriction on the mutual position of a and z. (In case of (2.5), the condition $a<x$ was needed because of the power $\{(x-a)/n\}^{-\mu}$.) For continuous functions, (5.2) can be written as a Riemann–Liouville integral (cf. Theorem I, part 2) and hence it is a regular function of the order μ. Therefore, ${}_af^{(\mu)}|z$ is in this case identical with the derivative ${}_aD^\mu f(z)$ of the paper [7], defined by the analytic continuation of the Riemann–Liouville integral and, for holomorphic $f(w)$, there exists also a contour integral representation of the integro-derivative of any order. —[15] As we pointed out in [7], for several types of functions it is suitable to consider the case $a = \infty$ (i. e. to submit ${}_aD^\mu$ to a limit process $|a| \to \infty$, $\arg(a-z) =$ constant), too.

6. Further application possibilities concerning the above results are furnished in the simplified treatment of certain *generalized difference, differential* or *integral equations.*

As an example, we mention some immediate consequences of Theorems II. and IV:

THEOREM VI. 1. *Let $\varphi_m(\xi)$ $(m = 1, 2, \ldots)$ be a given sequence of functions defined for $\xi \in (a, x]$. The solution of the system of equations*

$$\begin{cases} {}_a\Delta_m^\mu f|\xi = \varphi_m(\xi) & (a < \xi \leqq x) \\ & (m = 1, 2, \ldots) \end{cases} \tag{6.1}$$

can be given explicitly in the form

$$f(\xi) = {}_a\Delta_n^{-\mu}\varphi_{[n(u,\xi,a)]}(u)|\xi \quad (n = 1, 2, \ldots), \tag{6.2}$$

2. *Let $\Phi(u)$ be uniformly ${}_aD^\mu$-derivable with $\mu>0$ in the interval $(a, x]$. Then the equation*

$${}_af^{(-\mu)}|u = \Phi(u) \quad (a < u \leqq x) \tag{6.3}$$

is satisfied by

$$f(u) = {}_a\Phi^{(\mu)}|u. \tag{6.4}$$

Similar assertion holds for $\mu < 0$ and with a function $\Phi(u)$ for which ${}_a\Phi^{(\mu)}|u$ exists in $(a, x]$ with an error $o(n^\mu)$.

We remark still that functional equations involving fractional integrals and derivatives occur in connection with various problems of the literature, especially in the class of Volterra integral equations.[16] The most famous problem of this kind was studied by ABEL, namely the equation

$$\int_0^u \frac{f(t)}{(u-t)^\lambda}\,dt = \psi(u) \quad (\lambda < 1). \tag{6,5}$$

[15] Cf. [7], p. 71.
[16] Cf. e. g. [1].

If $f(t)$ is required to be continuous in $[0, u]$, (6.5) may be written in the form (cf. (2.9)):

$$(6.6) \qquad {}_0f^{(\lambda-1)}|u = \Gamma(1-\lambda)^{-1}\psi(u),$$

which is of (6.3) typ .

Literature

[1] H. T. Davis, The application of fractional operators to functional equations, *Amer. Journal of Math.*, **49** (1927), 123–142.

[2] É. Goursat, *Cours d'analyse*, I. (Paris, 1917).

[3] E. Hille, *Functional analysis and semigroups*, Amer. Math. Soc. Coll. Publ. XXXI (1948).

[4] K. Knopp, *Theorie und Anwendung der unendlichen Reihen*, 4. ed. (Berlin–Heidelberg, 1947).

[5] H. Mangoldt–K. Knopp, *Einführung in die höhere Mathematik*, II, 9. ed. (Leipzig, 1954).

[6] M. Mikolás, Differentiation and integration of complex order of functions represented by trigonometrical series and generalized zeta-functions, *Acta Math. Acad. Sci. Hung.*, **10** (1959), 77–124.

[7] M. Mikolás, Über die Begründung eines einheitlichen und erweiterten Infinitesimalkalküls im Komplexen, *Ann. Univ. Sci. Budapest*, Sectio Math., **5** (1962), 69–78.

[8] K.-F. Moppert, Über einen verallgemeinerten Ableitungsoperator, *Comm. Math. Helv.*, **27** (1953), 140–150.

[9] M. Riesz, L'intégrale de Riemann–Liouville et le problème de Cauchy, *Acta Math.*, **81** (1949), 1–223.

[10] N. Stuloff, Die Differentiation beliebiger reeller Ordnung, *Math. Annalen*, **122** (1951), 400–410.

PRINCIPAL COMPONENTS AND PRIME FACTORIZATION IN PARTIALLY ORDERED SEMIGROUPS

By

L. FUCHS and O. STEINFELD

Department of Algebra of the Eötvös Loránd University, and
Mathematical Institute of the Hungarian Academy of Sciences

(Received April 3, 1963)

Dedicated to T. GALLAI

This note is devoted to the problem of unique prime factorization in partially ordered semigroups S. Several authors (BIRKHOFF [2][1], ŠULGEĬFER [6], WEINERT [7] etc.) have dealt with this question in order to get suitable generalizations of the fundamental theorem of commutative ideal theory. There are known several necessary and sufficient conditions for S under which the unique prime factorization holds in S. Here we are giving two new sets of conditions so that 1. they are all, except for (vi), either trivially satisfied by the ideals of a ring or are usual conditions of ideal theory; 2. we do not assume the commutativity of S. We think that our discussion offers some novelty even in non-commutative ideal theory and in the ideal theory of semirings. Also, the concept of principal component in partially ordered semigroups seems to be new.

In **§ 1,** our objective is to prove the existence of a decomposition into the intersection of a finite number of (right) principal components. The conditions are all satisfied by ideals in rings with left identity except for (iii) which amounts to the restricted minimum condition. In **§ 2,** we impose further conditions on S to ensure that these principal components be prime powers and the intersection could be replaced by multiplication. In **§ 3,** we start with a new set of conditions which means in the special case of ideals that the maximum condition is satisfied and the ring is a multiplication ring (Multiplikationsring) in the sense of KRULL [5]. In this event we arrive rather easily at the unique prime factorization.

Throughout this note, S denotes a semigroup (i. e. a set with an associative multiplication) which is at the same time a partially ordered set under a relation $\leqq$ such that, for all a, b, c in S, $a \leqq b$ implies $ac \leqq bc$ and $ca \leqq cb$. We also assume that S is negatively ordered in the sense that $ab \leqq a$ and $ab \leqq b$. S may contain a minimum element[2] 0 which is then unique and satisfies $a0 = 0a = 0$ for all $a \in S$. Obviously, the set of all ideals of a ring, semiring etc. possesses this property.

§ 1. The principal component

In this section we are concerned with a negatively ordered semigroup with restricted minimum condition. We assume the existence of l. u. b. (but not that of g. l. b.) and residuals, and establish a meet decomposition into right principal components.

[1] Numbers in brackets refer to the bibliography given at the end of this note.

[2] We use the term „minimum element" to mean an element which is smaller than any other element. The meaning of „maximum element" is clear.

Let S be a partially ordered semigroup (with more than one element) satisfying the following conditions:

(i) S is negatively ordered and has a maximum element e which is a left identity:

$$ea = a \quad \text{for all } a \in S,$$

(ii) every pair of elements a, b in S has a l. u. b. which will be denoted by $a \vee b$;

(iii) for each $a \in S$, $a \neq 0$, the set of all $x \in S$ with $a \leqq x \leqq e$ satisfies the minimum condition;

(iv) for every pair of elements a, b in S there exists a right-residual $a : b \in S$ and a left-residual $a :: b \in S$ defined by

$$xb \leqq a \; [bx \leqq a] \text{ if and only if } x \leqq a : b \; [x \leqq a :: b].$$

Observe that the residuals obey the usual formal rules for which we refer to BIRKHOFF [1] or FUCHS [4]. We need to know that (ii) and (iv) imply the distributive laws (FUCHS [4], p. 190)

$$(1) \qquad a(b \vee c) = ab \vee ac, \qquad (b \vee c)a = ba \vee ca,$$

and the existence of g. l. b. for residuals:

$$(2) \qquad (a : b) \wedge (a : c) = a : (b \vee c), \qquad (a :: b) \wedge (a :: c) = a :: (b \vee c)$$

for all a, b, $c \in S$.

From hypotheses (i)–(iv) we conclude in turn:

a) *Given $a \in S$ with $0 < a < e$, there exists a prime*[3] *$p \in S$ such that, for some $b \in S$ $(b > a)$,*

$$p = a : b.$$

The set of all $x \in S$ satisfying $x > a$ is not empty, hence by (iii) there is a minimal element, say b, in this set. We show that $p = a : b$ is a prime. If $u, v \in S$ satisfy $uv \leqq p$, then by definition we have $uvb \leqq a$. Assume $v \not\leqq p$, that is, $vb \not\leqq a$. Hence $a < a \vee vb \leqq b$, and in view of the choice of b we obtain $a \vee vb = b$. By (1), $ua \vee uvb = ub$. Here the left member is $\leqq a$ whence $ub \leqq a$ and $u \leqq p$. Thus p is a prime, indeed.

b) *To every $a \in S$, $0 < a < e$, there exists a prime $p \in S$ with the properties:*

$$a \leqq p \qquad \textit{and} \qquad a :: p > a.$$

Let p be a prime of the form $p = a : b$ whose existence is guaranteed by a). Then obviously $a \leqq p$. Furthermore we have $pb \leqq a$ whence $a :: p \geqq b > a$.

c) *Let $a \in S$, $0 < a < e$, and p a prime in S. There exists a uniquely determined $a(p) \in S$ such that*

$$a : y \not\leqq p \quad \textit{if and only if} \quad y \leqq a(p).$$

This $a(p)$ will be called the *right principal component of a* belonging to p. *It satisfies: $a \leqq a(p)$; if $a \leqq p$ then $a(p) \leqq p$; if $a \not\leqq p$ then $a(p) = e$.*

[3] An element $p \in S$ $(0 < p < e)$ is said to be *prime* if $ab \leqq p$ implies either $a \leqq p$ or $b \leqq p$.

Consider the set T of all $y \in S$ satisfying $a : y \not\leqq p$. This T is not empty, for a belongs to it. If both y and z belong to T, then $a : (y \vee z) = a : y \wedge a : z \leqq p$ cannot hold, for this would imply $(a : y)(a : z) \leqq p$ and hence $a : y \leqq p$ or $a : z \leqq p$, in contradiction to the choice of y and z. Hence $y, z \in T$ implies $y \vee z \in T$. By (iii), there exists a $w \in T$ such that $a : w$ is a minimal element in the set of all $a : v$ with $v \in T$. Since T is closed under union, the set of all $a : v$ with $v \in T$ is closed under intersection, and thus $a : w$ is the minimum element in the set considered. Then

$$x = a :: (a : w) \tag{3}$$

has the property $a : x = a : w$. This x is the maximum element of S with this property, moreover, x is the maximum element of T: if $y \in T$ then $a : y \geqq a : w = a : x$ whence $(a : x)y \leqq (a : y)y \leqq a$ and $y \leqq a :: (a : x) = x$. That $y \leqq x$ implies $a : y \not\leqq p$ is evident because of $a : y \geqq a : x$. Thus the existence of $x = a(p)$ is established. Finally, since $a : a = e \not\leqq p$, therefore $a \leqq a(p)$. If $a \leqq p$, then $[a : a(p)]\, a(p) \leqq a \leqq p$, and the prime property of p implies $a(p) \leqq p$. If $a \not\leqq p$ then $a : e \not\leqq p$, and so $a(p) = e$. This proves c).

d), e) and f) are concerned with some properties of $a(p)$.

d) *If $b \in S$ is the right principal component of some $a \in S$, $b = a(p)$, then $b(p) = b$.*

By virtue of c), $a(p) \leqq b(p)$. Furthermore, $[b : b(p)]\, b(p) \leqq b = a(p)$ implies, by definition, that

$$a : \{[b : b(p)]\, b(p)\} \not\leqq p$$

or

$$a : b(p)] : [\, b : b(p)] \not\leqq p.$$

This and $b : b(p) \not\leqq p$ show that the product of $[a : b(p)] : [b : b(p)]$ and $b : b(p)$ cannot be $\leqq p$, and so $a : b(p) \not\leqq p$. Therefore $b(p) \leqq a(p)$, and we are led to $b(p) = a(p) = b$.

e) *For any $a, b \in S$ and any prime $p \in S$, the right principal component of $a : b$ belonging to p is $a(p) : b$.*

Let x denote the right principal component of $a : b$ belonging to p. Then $(a : b) : x \not\leqq p$. Here $(a : b) : x = a : xb$, and so $a : xb \not\leqq p$ whence $xb \leqq a(p)$ and $x \leqq a(p) : b$. On the other hand,

$$(a : b) : [a(p) : b] = a : \{[a(p) : b]\, b\} \geqq a : a(p).$$

Here $a : a(p) \not\leqq p$, thus $(a : b) : (a(p) : b) \not\leqq p$ and $a(p) : b \leqq x$. We obtain $x = a(p) : b$.

f) *If $a(p)$ is the right principal component of $a \in S$ belonging to the prime p, and if $a(p) :: b > a(p)$ for some $b \in S$, then $b \leqq p$.*

Set $x = a(p) :: b$. Then $bx \leqq a(p) = a :: (a : w)$ [see (3)], whence $(a : w)bx \leqq a$, $(a : w)b \leqq a : x$. Because of assumption, $x > a(p)$, and thus $a : x \leqq p$, and $(a : w)b \leqq p$. By the definition of w, $a : w \not\leqq p$, and therefore $b \leqq p$.

Now we are going to prove the main result of this section [in the proof we actually use only b) and c)].

THEOREM 1. *If S is a partially ordered semigroup satisfying conditions* (i)–(iv), *then to every $a \in S$ $(0 < a < e)$ there exist a finite number of primes $p_1, \ldots, p_r$, such that*

$$a = a(p_1) \wedge \cdots \wedge a(p_r). \tag{4}$$

By (3), every $a(p)$ is a left residual of a, and so our remark made in connection with (2) implies that any finite set of the $a(p_i)$ has an intersection in S. By b), primes do exist, and so the set of all finite intersections of right principal components of a is not empty. The elements of this set are all $\geqq a$, therefore this set contains a minimal element a^* which must be the minimum element, for the set is closed under intersection. Clearly $a \leqq a^*$, and what we need to show is that here equality holds. The minimum property of a^* shows that $a^* \leqq a(p)$ for all primes $p \in S$, and therefore, by c), $a : a^* \not\leqq p$. Consequently, $a : a^* \leqq p$ holds for no prime p in S, and b) implies $a : a^* = e$. Thus $a^* = ea^* \leqq a$, and we arrive at $a^* = a$, completing the proof of Theorem 1.

§ 2. Representation as product of primes

In the previous section we have obtained representations of the elements of S as intersections of right principal components. If we wish to get representations as products of commuting primes, then we have to impose on S further conditions.

Let S be a partially ordered semigroup satisfying in addition to (i)–(iv) also

(v)[4] if a, $b \in S$, $a \neq 0$ satisfy $ba = a$, then $b = e$; and if they satisfy $ba = 0$, then $b = 0$;

(vi) if a, $b \in S$ and the prime $p \in S$ satisfy $a < b \leqq p$, then

$$a :: p < b :: p.$$

Under conditions (i)–(vi) we can prove:

g) *If p, $q \in S$ are primes such that, for some $a \in S$, $a(p) \leqq q$, then $q \leqq p$.*

Assume that $a(p) \leqq q$ and $q \not\leqq p$. Then, by f), the left-residual $a(p) :: q$ cannot be greater than $a(p)$, i. e. $a(p) :: q = a(p)$. Therefore

$$a(p) \leqq [q \cdot a(p)] :: q \leqq a(p) :: q = a(p).$$

Apply (vi) to $q \cdot a(p) \leqq a(p)$ and q to obtain $q \cdot a(p) = a(p)$. From (v) we infer $q = e$, a contradiction to the primeness of q.

h) *For each prime $p \in S$,*

$$e > p > p^2 > \cdots > p^n > \cdots \tag{5}$$

is a properly descending maximal chain.

Since $p^n \geqq p^{n+1}$ and by (v) equality cannot hold here, (5) is a properly descending chain. In order to verify its maximality, let us first show that $e > q \geqq p$ implies $q = p$. Because of b), there is no loss of generality in assuming that q is a prime. Then the right principal component $p(p)$ is equal to p, for if $p :: x \not\leqq p$ then $(p :: x)x \leqq p$ implies $x \leqq p$. From $p(p) = p \leqq q$ it follows, in view of g), that $q \leqq p$, and so $q = p$, indeed.

[4] This condition replaces the condition of integral closure of commutative ideal theory. Here we do not consider „fractional ideals", and therefore we cannot use the notion of integral closure as in DUBREIL–JACOTIN [3].

To complete the proof, assume, as a basis of induction, that we have shown for some $k \geqq 1$:

1. $p^{k-1} > a \geqq p^k$ implies $a = p^k$, and
2. $p^k :: p = p^{k-1}$.

Let $p^k > b \geqq p^{k-1}$ for some $k \geqq 1$. Then from (vi) we infer

$$p^k :: p > b :: p \geqq p^{k+1} :: p.$$

Here $p^k :: p = p^{k-1}$ by assumption 2., and evidently $p^{k+1} :: p \geqq p^k$. Thus $p^{k-1} > b :: p \geqq p^{k+1} :: p \geqq p^k$, and assumption 1. implies $b :: p = p^{k+1} :: p = p^k$. Again from (vi), we get $b = p^{k+1}$, and therefore 1. and 2. hold true for $k+1$ in place of k. This establishes h).

j) *The primes of S commute.*

Suppose that p and q are different primes of S. If $r \in S$ is a prime of S such that $a = pq \leqq r$, then either $p \leqq r$ or $q \leqq r$, and hence, by what has been proved in h), either $r = p$ or $r = q$. By (v) certainly $a \neq 0$. On account of c), the only right principal components of a which differ from e are $a(p)$ and $a(q)$. From Theorem 1 we obtain

$$a = pq = a(p) \wedge a(q).$$

Obivously, $pq :: p \geqq q$ and $pq :: p \neq e$, thus, by h), $pq :: p = q$. The left-residual of a by p yields

$$q = [a(p) :: p] \wedge [a(q) :: p].$$

Since $p \not\leqq q$, from f) it follows that $a(q) :: p = a(q)$. The meet representation of q gives $q \leqq a(q)$, while $a(q) \leqq q$ by c); so $a(q) = q$. By h), there are two possibilities for $a(p) :: p$, namely q or e. It is not equal to q, for $a(p) \leqq q$ would contradict g). Consequently, $a(p) :: p = e$, $pe \leqq a(p)$. Here pe satisfies $p^2 \leqq pe \leqq p$. But $p^2 = pe$ would imply $p^3 = pep = p^2$, contrary to h). Hence $pe = p$ and $p \leqq a(p)$, which together with $a(p) \leqq p$ establishes the equality $a(p) = p$. We have thus shown:

$$pq = p \wedge q.$$

Changing the roles of p and q, we obtain $qp = p \wedge q$, and hence the desired commutability.

k) *To every $a \in S$, $a \neq 0$, and every prime $p \in S$, there is a positive integer n such that*

$$p^n \leqq a(p).$$

If $a \not\leqq p$ then $a(p) = e$, and the statement is obvious. Assume $a \leqq p$. In the set $a(p) \vee p^m$ $(m = 1, 2, \ldots)$ there is a minimal element, say, $a(p) \vee p^n$ [use (iii)]. If $a(p) \vee p^n \neq a(p)$, then $a(p) :: p^n = a(p) :: [a(p) \vee p^n] \neq e$, and from b) we obtain the existence of a prime $q \in S$ satisfying

$$[a(p) :: p^n] :: q > a(p) :: p^n.$$

We have $a(p) \leqq q$, and from g) we see that only $q = p$ is possible. Consequently,

$$a(p) :: p^n < [a(p) :: p^n] :: p = a(p) :: p^{n+1} = a(p) :: [a(p) \vee p^{n+1}]$$

whence $a(p)\vee p^n = a(p)\vee p^{n+1}$ is absurd. Thus $a(p)\vee p^n > a(p)\vee p^{n+1}$, in contradiction to our choice of n. Therefore $a(p)\vee p^n = a(p)$, and $p^n \leqq a(p)$, as we wished to prove.

l) If $0<a\leqq p$, then the right principal component $a(p)$ is a power of p.[5]

By k), we have $p^n \leqq a(p) \leqq p$ for some n which we may assume to be chosen as small as possible. Also, k can be chosen so as to satisfy

$$p^n \leqq a(p) \leqq p^k$$

and to be as large as possible. If $k=n$, we are ready. If $k<n$, then $p^n < a(p) < p^k$ and, because of (5), $n \neq k+1$. Clearly, $p^{k+1}\vee a(p) = p^k$ whence by multiplication by p^{n-k-1} we get $p^n \vee p^{n-k-1}a(p) = p^{n-1}$. Here the left member is $\leqq a(p)$, thus $p^{n-1} \leqq a(p)$, in contradiction to the choice of n. Thus $k<n$ is impossible.

We have arrived at the main result of this section.

THEOREM 2. *Let S be a partially ordered semigroup satisfying* (i)–(vi). *Then every $a\in S$, other than* 0 *and e, can be written uniquely in the form*

$$a = p_1^{k_1}\cdots p_r^{k_r}$$

with prime powers $p_i^{k_i}$ belonging to different bases p_i.

From Theorem 1 and from *l*) it results that every a $(0<a<e)$ is of the form

$$a = p_1^{k_1}\wedge\cdots\wedge p_r^{k_r}$$

with different primes p_i. BIRKHOFF [1] has shown that if $x\vee y = e$ and $xy = yx$, then $x\wedge y = xy$, further if $x\vee y = e$ and $x\vee z = e$, then $x\vee yz = e$. By making use of these, it is straightforward to verify that $a = p^{k_1}\ldots p^{k_r}$. Uniqueness follows in the following way. Taking into account the maximality of primes, by standard arguments we infer that it suffices to show that

$$a = p_1^{k_1}\cdots p_r^{k_r} = p_1^{l_1}\cdots p_r^{l_r} \qquad (p_i \text{ primes})$$

mplies $k_1 = l_1, \ldots, k_r = l_r$. By way of contradiction assume e. g. that $k_r > l_r$, that is to say, $p_r^{k_r} > p_r^{l_r}$. Since

$$p_r^{k_r} : p_r^{l_r} \geqq p_r^{k_r - l_r} \qquad \text{and} \qquad p_r^{k_r} : p_r^{l_r} \geqq a : p_r^{l_r} \geqq p_1^{l_1}\cdots p_{r-1}^{l_{r-1}},$$

therefore

$$p_r^{k_r} : p_r^{l_r} \geqq p_r^{k_r - l_r} \vee p_1^{l_1}\cdots p_{r-1}^{l_{r-1}} = e$$

because of the cited result of BIRKHOFF. Hence $p_r^{l_r} = e p_r^{l_r} \leqq p_r^{k_r}$, a contradiction. This proves uniqueness, completing the proof of Theorem 2.

Note that conditions (v) and (vi) can be replaced by the following two: (v*) S is commutative (or more simply that $a : b = a :: b$ for all $a, b\in S$); (vi*) for every prime $p\in S$, the chain (5) is maximal. Then we could omit the proofs of g), h) and j).

COROLLARY (see BIRKHOFF [2] Theorem 13). Let L be a commutative residuated lattice-ordered semigroup with identity e, in which

[5] Here we follow the argument of ŠULGEĬFER [6].

(α) $0<a<e$ implies that all chains between a and e are finite and

(β) $p(e:p)=e$ for every maximal element $p<e$.

Then every element $a\in L$ with $0<a<e$ has a unique factorization into maximal $p_i<e$.

It is easy to prove that the conditions of Birkhoff's Theorem imply conditions (i), (ii), (iii), (iv), (v*), (vi*) for the partially ordered semigroup S of the elements $a\in L$ with $0\leqq a\leqq e$.

§ 3. Divisibility partially ordered semigroups

Now we turn our attention to partially ordered semigroups in which the order relation is defined in terms of divisibility[6], and we wish to deal with conditions ensuring that unique prime factorization hold.

Assume that S is a partially ordered semigroup (with at least two elements) satisfying:

(I) S is negatively ordered and has a maximum element e which is a left identity in S;

(II) if, for some $a, b\in S$, $a=ba\neq 0$ then $b=e$, and if $ba=0$ then $b=0$ or $a=0$,

(III) if $a, b\in S$ and $a<b$, then there exists a $c\in S$ such that

$$a = bc,$$

(IV) S satisfies the maximum condition;

(V) for every pair $a, b\in S$, the right-residual $a:b$ exists in S.

It is evident that if every element $\neq 0$, e of S is uniquely factorizable into primes, then S satisfies conditions (I)–(V). We are going to show the converse. To this end, suppose that S satisfies (I)–(V). Then we have:

(A) *Every $a\in S$, $0<a<e$, is the product of a finite number of maximal elements $p_i\in S$:*

$$a = p_1p_2\cdots p_k.$$

Let $a\in S$, $0<a<e$. Because of (IV) there exists a maximal p_1 in the set of all $x\in S$ with $a\leqq x<e$. If $a=p_1$, there is nothing to prove. If $a<p_1$, we may apply (III) to conclude that, for some $a_2\in S$, $a=p_1a_2$. Here, by negative order, $a\leqq a_2$, but equality would lead to $a=p_1a$, contrary to (II). Hence $a<a_2$. Repeat this argument for a_2 in place of a to obtain that either a_2 is maximal or $a_2=p_2a_3$ with a maximal p_2 and $a_2<a_3$ etc. This process must terminate in a finite number of steps, for the sequence $a<a_2<a_3<\ldots$ cannot be infinite. Hence some a_k is a maximal element p_k, and substitution yields $a = p_1a_2 = p_1p_2a_3 = \ldots = p_1p_2\ldots p_k$.

(B) *An element $p\in S$ is maximal if and only if it is prime.*

Let $p\in S$ be a maximal element and $ab\leqq p$, $a\not\leqq p$. In view of (V), $p:b$ exists and $a\leqq p:b$. Obviously, $p\leqq p:b$ whence $a\not\leqq p$ implies $p<p:b$, and so $p:b=e$, $b=eb\leqq p$. Thus p is a prime.

[6] Divisibility in the sense of Birkhoff [2]; its precise meaning is given in (III).

Conversely, assume that p is a prime and $p<a$ for some $a\in S$. (III) guarantees the existence of a $b\in S$ with $p=ab$. The prime property of p shows that $b\leqq p$ whence $b=p$. Now $p=ap$ implies, by (II), that $a=e$ and p is maximal.

(C) *The prime elements of S commute.*

Let p, q be different primes of S. By (III), there exists an $a\in S$ such that $pq=qa$. In view of (B), $q\not\leqq p$, and thus $qa\leqq p$ implies $a\leqq p$. Therefore $pq\leqq qp$. Changing the roles of p and q, we get the converse inequality, whence $pq=qp$, in fact.

(D) *S is commutative.*

From (A) and (B) we know that the elements $\neq e$ are products of primes, and in (C) we have shown that the primes commute. Therefore $ab = ba$ whenever $a, b\in S$ are different from e. It remains to show that e is a right identity. It suffices to prove that $pe = p$ for primes $p\in S$. By (A), we may write pe as product of primes: $pe=p_1 \ldots p_k$. Here $pe\leqq p_i$ implies $p\leqq p_i$, i. e. $p=p_i$. Thus $pe = p^k$. Since e is a left identity, $p^2=pep=p^{k+1}$. Hence, in view of (II), it follows that $p^{k-1}=e$, that is, $k=1$. Therefore $pe=p$, as claimed.

(E) *S obeys the cancellation law:*

$$ax=bx \qquad \textit{implies} \qquad a=b \ (x\neq 0).$$

First we dispose of the case $a\leqq b$. If we had $a<b$, then by (II) there would exist a $c\in S$ such that $a=bc$. Thus we would have $bcx=bx$, whence, because of (D) and (II), $c=e$ which is absurd. Now assume $ax=bx$ with arbitrary $a, b\in S$. We have obviously $a\leqq(ax):x$ whence $ax\leqq(ax:x)x\leqq ax$ and $ax=(ax:x)x$. By the case disposed of we infer $a=ax:x$. Similarly, $b=bx:x$, and so $ax=bx$ in fact implies $a=b$.

Now we are in a position to verify:

THEOREM 3. *Let S be a partially ordered semigroup with properties* (I)–(V). *Then e is the identity of S and every element $a\in S$ $(0<a<e)$ may be represented, in one and only one way, as the product of pairwise commuting primes.*

By (A) and (D), every element $\neq 0$, e of S is the product of commuting primes which are, because of (B), maximal elements. We need still to prove the uniqueness. But this can be established on classical lines by making use of (E).

It might be of some interest to observe that the maximum condition (IV) could be replaced by the restricted minimum condition:

(IV*) the set of elements x of S satisfying $x\geqq a$ obeys the minimum condition, for each $a\in S$, $a\neq 0$.

In this case we also assume that e is a right identity and the dual of (II).

In order to prove that (IV*) replaces (IV) in the proof of Theorem 3, we show that (A) holds. Let $a\in S$, $0<a<e$ and a_2 a minimal element $>a$. There exists a $p_1\in S$ such that $a=a_2p_1$. This p_1 is maximal, for if $p_1<b$ then $p_1=bc$ for some $c\in S$, and $a=a_2bc$. Here either $a_2b=a_2$ or $a_2b=a$, that is either $b=e$ or $c=e$, and the latter case is impossible. Repeating this argument for a_2 etc., we obtain $a=a_kp_{k-1} \ldots p_1$. Since $p_1>p_2p_1>\ldots$, this process comes to an end, establishing (A).

Finally, let us mention that neither condition (II) nor (V) can be omitted. This is shown by the following two examples:

1. First let S consist of $e, a, a^2, \ldots, b, b^2 \ldots$ such that

$$e > a > a^2 > \cdots > b > b^2 > \cdots$$

with multiplication defined in this way: e is the identity, $a^k a^l = a^{k+l}$, $b^k b^l = b^{k+l}$, $a^k b^l = b^l a^k = b^l$. Here all conditions except for (II), are satisfied.

2. Let S consist of $e, b, a, a^2, \ldots$ such that

$$e \begin{array}{c} > a > \\ > b > \end{array} a^2 > a^3 > \cdots$$

where e is the identity, $a^k a^l = a^{k+l}$ and $b^2 = a^2$, $ba^k = a^k b = a^{k+1}$. Here there is no prime. Only condition (V) fails to hold.

Literature

[1] G. Birkhoff, *Lattice Theory*, New York, 1948.

[2] G. Birkhoff, Lattice-ordered semigroups, *Seminaire Dubreil-Pisot* (Algèbre et Théorie des nombres), **14** (1960/61), No. 19.

[3] M. L. Dubreil – Jacotin, Quelques propriétés des équivalences régulières par rapport à la multiplication et à l'union, dans un treillis à multiplication commutative avec élément unitée, *C. R. Acad. Sci. Paris*, **232** (1951), 1174 – 1176.

[4] L. Fuchs, *Partially ordered algebraic systems*, Budapest, 1963.

[5] W. Krull, *Idealtheorie*, Berlin, 1935.

[6] (E. G. Šulgeĭfer) Е. Г. ШУЛЬГЕЙФЕР, Разложение на простые множителы в структурах с умножениам, *Укранский Математический Журнал*, 2 (1950), 100—114.

[7] H. J. Weinert, *Über Halbringe und Halbkörper*, Habilitationsschrift, Potsdam, 1960.

ÜBER IN EINEM KREISRING ENTHALTENE POLYGONE

Von

K. BÖRÖCZKY

Lehrstuhl der Geometrie, L. Eötvös Universität, Budapest

(Eingegangen am 15. Juni 1963.)

Wir betrachten einen Kreisring vom Zentrum O, vom inneren Randkreis k und äußeren Randkreis K, ferner Polygone die die folgende Eigenschaft besitzen: sie enthalten den Kreis k und alle ihre Eckpunkte liegen bis auf höchstens zwei nebeneinanderliegende auf dem Randkreis K. Gibt es zwei solche Ausnahmseckpunkte, so sind sie durch die Ausnahmsseite verbunden. Für die betrachteten Polygone beweisen wir folgenden Satz, der von G. HAJÓS stammt, und eine Verallgemeinerung eines Lemmas von J. MOLNÁR[1] ist, welches er bei der Abschätzung der Dichte von Kreisunterdeckungen benützte.

Hat ein Polygon der betrachteten Eigenschaft einen minimalen Inhalt, so berühren alle seine Seiten den Kreis k und der Berührungspunkt seiner Ausnahmsseite halbiert den die Berührungspunkte der nebenanliegenden Seiten verbindenden Bogen.

Wir führen einige Operationen ein, welche Polygone der betrachteten Eigenschaft in wieder solche überführen und den Inhalt vermindern.

a) Wenn die Ausnahmsseite den Kreis k nicht berührt und wenn es nur einen bzw. keinen Ausnahmseckpunkt gibt, so kann man die Ausnahmseckpunkte bzw. einen beliebigen Eckpunkt durch eine Tangente des Kreises k abschneiden, welche durch keinen Eckpunkt des Polygons läuft (Fig. 1).

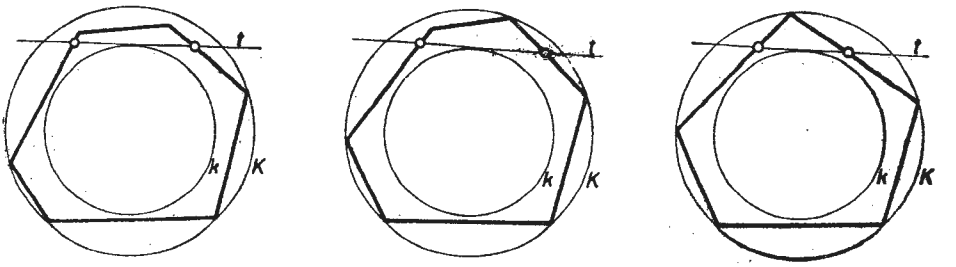


Fig. 1.

b) Berührt die Ausnahmsseite den Kreis k, eine nebenanliegende Seite aber nicht, so können wir diese um ihren auf K liegenden Endpunkt so verdrehen, daß auch sie k berühre (Fig. 2).

c) Den Streckenzug, der aus den die auf K liegenden Eckpunkte verbindenden Seiten besteht, kann man durch einen solchen ersetzen, dessen alle Seiten bis auf höchstens eine Endseite dieses Streckenzuges den Kreis k berühren.

[1] J. MOLNÁR, Körelhelyezések állandó görbületü felületeken, *MTA Mat. és Fiz. Oszt. Közleményei*, **12** (1962), 223–263, s. S. 227.

Diese Operation kann den Inhalt nicht vergrößern, was aus einem Lemma von J. MOLNÁR.[2] folgt (Fig. 3.).

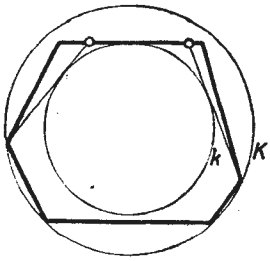


Fig. 2.

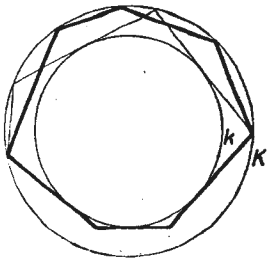


Fig. 3.

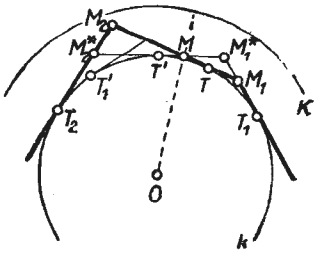


Fig. 4.

d) Wenn der Berührungspunkt T der Ausnahmsseite den die Berührungspunkte T_1, T_2 der Nachbarseiten verbindenden Bogen nicht halbiert, so können wir sie durch eine im Halbierungspunkt T' berührende Seite ersetzen. Diese Operation vermindert den Inhalt, was wir folgendermaßen einsehen können (Fig. 4). Liege T am Bogen $T'T_1$ und sei M der Schnittpunkt der Tangenten in T und T', welche die in T_1 und T_2 berührenden Tangenten in den Punkten M_1, M_2 bzw. M_1^*, M_2^* schneiden. Spiegelt man das hinzugefügte $\Delta MM_1M_1^*$ an der Geraden OM, so ist dieses Spiegelbild im abgeschnittenen $\Delta MM_2^*M_2$ enthalten, da der Punkt T_1 in einen Punkt T_1' des Bogens T_2T' gespügelt wird.

Nach dieser Vorbereitung beweisen wir unseren Satz folgendermaßen:

Wir dürfen voraussetzen, daß das betrachtete Polygon eine Ausnahmsseite besitzt, die den Kreis k berührt, da man sonst durch die Operation *a)* zu einem solchen Polygon gelangt. Ebenso darf man wegen der Operation *b)* voraussetzen, daß die neben der Ausnahmsseite liegenden Seiten k berühren. Endlich dürfen wir wegen der Operation *c)* voraussetzen, daß alle weitere Seiten k berühren mit Ausnachme höchtens einer Seite P_3P_4, die zweinächst der Ausnahmsseite P_1P_2 liegt (Fig. 5).

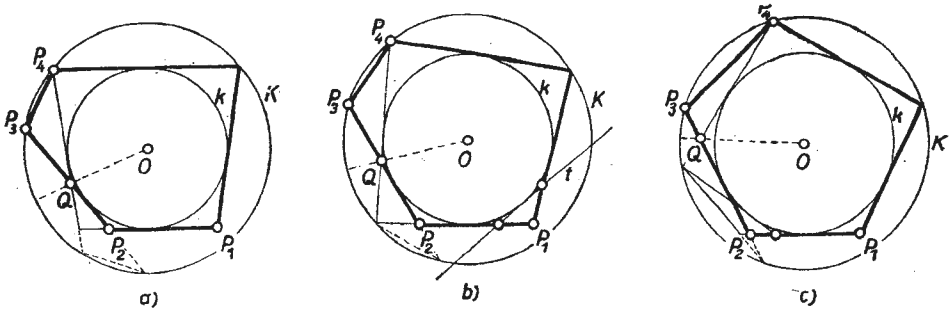


Fig. 5.

[2] J. MOLNÁR. Körelhelyezések állandó görbületű felületeken, *MTA Mat. és Fiz. Oszt. Közleményei*, 12 (1962) 226.

Wir schneiden durch die aus P_4 gezogene Tangente das ΔP_4P_3Q aus unserem Polygon ab, spiegeln dieses an der Geraden OQ, und fügen das erhaltene Dreieck hinzu. Danach schneiden wir vom hinzugefügten Dreieck jenen Teil ab, welcher jenseits der Tangente P_1P_2 liegt. Dadurch erhalten wir ein Polygon von der betrachteten Eigenschaft, welches eine Ausnahmsseite besitzt und dessen alle Seiten k berühren (Fig. 5a), oder gelangen zu einem solcher, wenn wir die Operation *a)* bzw. *b)* anwenden (Fig. 5b und 5c).

Endlich können wir die Operation *d)* anwenden, wodurch ein in unserem Satz gekennzeichnetes Polygon entsteht. Während der Operationen haben wir den Inhalt in jedem Fall vermindert, was nur dann nicht unmittelbar ersichtlich ist, wenn allein die Operation *c)* angewandt wurde. In diesem Fall vermindert aber auch diese Operation unbedingt den Inhalt.

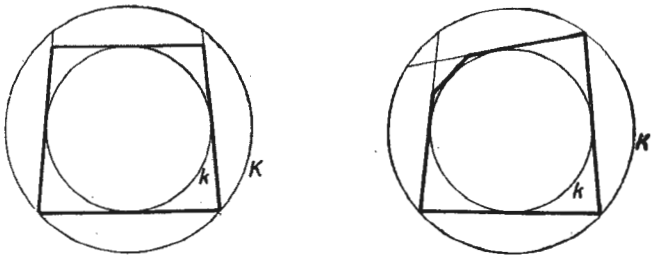


Fig. 6.

Wir bemerken noch, daß es in einem Kreisring im allgemeinen zwei solche Polygone gibt, die laut unseren Satzes einen minimalen Inhalt haben können (Fig. 6).

FILLING THE PLANE WITH CONGRUENT CONVEX HEXAGONS WITHOUT OVERLAPPING

By

BÉLA BOLLOBÁS

Department of Geometry of the Eötvös Loránd University, Budapest

(Received May 28, 1963)

Filling the plane with congruent figures (without overlapping and interstices)[1] seems to be a difficult unsolved problem of geometry. Not even the special case, where the congruent figures are convex polygons, has been settled; there are only partial results.[2]

In the present paper we shall discuss only the filling of the Euclidean plane with congruent convex hexagons, the so called convex hexagon-patterns. Only one of our results should be mentioned here, a corollary of Theorem 1: if the plane can be filled with congruent hexagons, then it can be filled with the same congruent hexagons in such a way that the graph of covering will be isomorphic to the graph of regular tesselation with hexagons.

DEFINITIONS. The term *pattern* will be used for any arrangement of polygons filling the plane.[3] A *tesselation* is a pattern in which no vertex of a polygon is the interior point of a side of another polygon.[4]

A *node* is a point of the plane coinciding with a vertex (consequently at least two vertices are in a node). A node will be called *first-rate*, if it is interior of a side of a polygon, *second-rate* otherwise.

THEOREM 1. *In a convex hexagon-pattern there is a square area of arbitrary side such that each interior node of this square is second-rate and it is surrounded by three hexagons.*

PROOF. Let a denote the area of the hexagon, and let $\frac{d}{2}$ denote the diameter of a circle, which contains the hexagon in its interior.

Let S be a square of side s on the plane of pattern ($s \geqq 3d$) and draw parallels to the sides of S of distance $\frac{d}{2}$. These parallels determine some squares, two of which, S_1 and S_2 have the side $s-d$ resp. $s+d$.

Let n_1 be the number of first-rate nodes in the square S, n_2 the number of interior second-rate nodes of the square surrounded by three hexagons, n_3 by at least four hexagons. Let L denote the number of hexagons having interior points in the square S, and K the number of hexagons contained by the interior of square S (Fig. 1.)

[1] These conditions will not be mentioned further.

[2] See e. g. Б. Н. Делоне [2], F. HAAG [3], [4] and F. LAVES [5].

[3] I. e. each point of the plane is inside or on the boundary of a polygon, but no two polygons have a common interior point. See e. g. COXETER [1].

[4] See e. g. J. MOLNÁR [6].

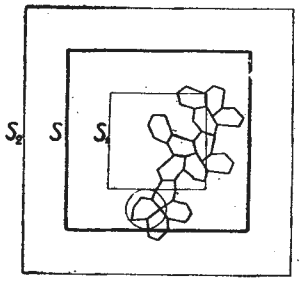


Fig. 1.

Since the square S_2 covers the hexagons, having interior points in S; the sum of the areas of these hexagons is smaller than the area of S_2; consequently

$$L < \frac{(s+d)^2}{a}. \tag{1}$$

We obtain similarly the inequality

$$K > \frac{(s-d)^2}{a}. \tag{2}$$

For the case, when $e = n_1 + n_2$ is limited for any square S, the theorem is trivial, so we may now suppose $e \geqq 5$. We shall show that to prove the the theorem it is sufficient to establish that the quotient $\frac{n_2}{n_1+n_3}$ will exceed an arbitrary number N.

a) If $\frac{n_2}{n_1+n_3} > N$, the square S contains at least $N(n_1+n_3) = Ne$ second-rate nodes, sorrounded by three hexagons. Being $e \geqq 5$, a positive integer f can be found satisfying $e < f^2 \leqq 2e$. Let us divide the square S into f^2 congruent squares. Since $f^2 > e$, at least one of these squares will contain in its interior only second-rate nodes, surrounded by three hexagons. We shall show that the side of this square will exceed any length, if N is sufficiently great.

For the area of this square we have the inequality:

$$\frac{s^2}{f^2} \geqq \frac{s^2}{2e}. \tag{3}$$

The square S contains in its interior at least n_2 nodes, each of them being surrounded by three hexagons, consequently in these nodes there are $3n_2$ vertices and this number can not be greater than $6L$ (the number of vertices of polygons having interior points in the square S):

$$3n_2 \leqq 6L,$$

in consequence

$$eN < 2\frac{(s+d)^2}{a}.$$

As

$$s \geqq 3d,$$

$$(s+d)^2 \leqq \left(\frac{4}{3}s\right)^2 < 2s^2,$$

therefore

$$s^2 > \frac{(s+d)^2}{2} > \frac{eNa}{4}$$

i. e.

$$\frac{s^2}{2e} > \frac{Na}{8}.$$

Considering (3) this means that we canchoose a square of area $\frac{Na}{8}$ such that the interior nodes of this square are second-rate nodes, and each of them is surrounded by three hexagons. Thus, for proving the theorem it is sufficient to show that N can exceed any number.

b) Now we prove that for a sufficiently great number s, $\frac{n_2}{n_1+n_3} > N$, where N is an arbitrary number.

In the first-rate nodes there are at least two vertices, and in the nodes surrounded by k hexagons there are k vertices; therefore in the square S there are at least $2n_1+3n_2+4n_3$ vertices. These vertices are contained by the vertices of hexagons covering the square (their number is $6L$), thus

$$2n_1+3n_2+4n_3 \leqq 6L$$

and, in consequence of (1),

$$2n_1+3n_2+4n_3 < 6\frac{(s+d)^2}{a}. \tag{4}$$

The angles of hexagons form in every first-rate node an angle π, and in every second-rate node an angle 2π. The vertices of the interior hexagons of the square S are in the interior nodes of S. Hence

$$n_1\pi+2n_2\pi+2n_3\pi > K4\pi,$$

where 4π is the sum of angles in a hexagon.

Owing to (2) we get

$$n_1+2n_2+3n_3 > 4\frac{(s-d)^2}{a}, \tag{5}$$

which can be written in the form

$$-\frac{3}{2}n_1-3n_2-3n_3 < -6\frac{(s-d)^2}{a}. \tag{6}$$

Adding the corresponding terms of inequalities (4) and (6), we have

$$\frac{n_1}{2}+n_3<\frac{24sd}{a}.$$

By this and by the inequality (5) we obtain:

$$n>2\frac{(s-d)^2}{a}-\frac{24sd}{a}=\frac{2s^2-28sd+2d^2}{a}.$$

Consequently

$$\frac{n_2}{\frac{n_1}{2}+n_3}>\frac{2s^2-2sd+2d^2}{24\,sd}=\frac{s}{12d}-\frac{7}{6}+\frac{d}{12s},$$

Therefore

$$\frac{n_2}{n_1+n_3}>\frac{s}{24d}-\frac{7}{12}+\frac{d}{24s}.$$

d being a fixed number, the value of the above expression will exceed any number if we choose s sufficiently large. This completes the proof of Theorem 1.

From this theorem it follows that, if there is a pattern with a convex hexagon, pattern can be made with the same hexagon the graph of which is isomorphic to the graph of regular tesselation with hexagons.*

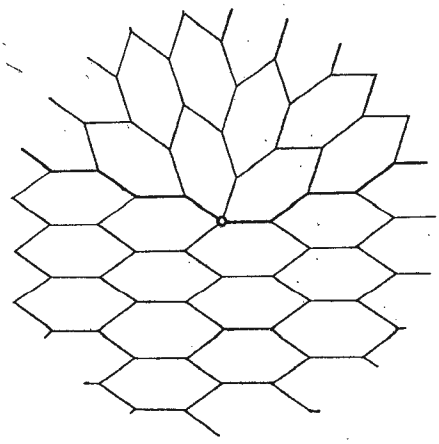


Fig. 2.

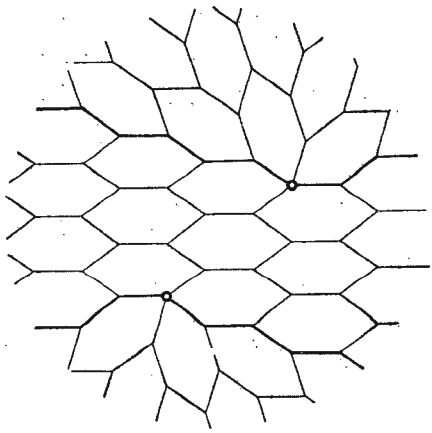


Fig. 3.

On the other hand these patterns are obviously not the only patterns with convex hexagons, as it is shown by figures 2, 3 and 4.

Using Theorem 1. we can prove the following theorems:

* From this it follows that, in the case of a given convex hexagon, to settle whether or not we can make a pattern with this hexagon, it is enough to examine the existence of a pattern the graph of which is isomorphic to the graph of the regular tesselation with hexagons.

THEOREM 2. *If there is a pattern with a convex hexagon, then at least two of the sides (of this hexagon) are equal.*

PROOF. Let the sides of hexagon $I \equiv ABCDEF$, implying the pattern, be $AB = a$, $BC = b$, $CD = c$, $DE = d$, $EF = e$ and $FA = f$, and let II and III be the hexagons joining to the edges AB and BC respectively. (Fig. 5.)

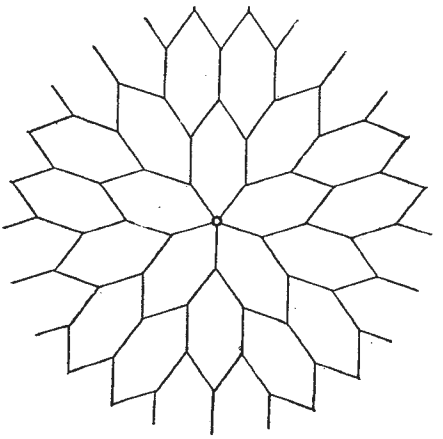


Fig. 4.

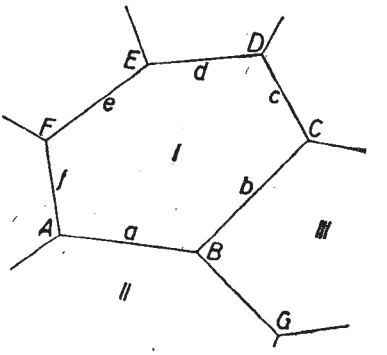


Fig. 5.

On account of theorem 1, we may suppose that the graph of this pattern which will be examined is isomorphic to the regular tesselation with hexagons; therefore the hexagons II and III are adjacent, their common edge being BG.

Supposing that the hexagon has no two equal sides, the edge AB is the side a of hexagon II, and the edge BC is the side of hexagon III. From this it follows that the edge BG in hexagon II is adjacent to a, and in hexagon III is adjacent to b. Consequently at least one of the adjacent sides of a (b or f) is equal to one the adjacent sides of b (a or c), and the hexagon has indeed two equal sides.

THEOREM 3. *If a convex hexagon has not more than two equal sides, there is a pattern with this hexagon if and only if these equal sides are opposite and parallel.*

a) The condition is necessary.

If a hexagon has only two equal sides, then there are two adjacent sides, that neither of them is equal to another side of the hexagon. Let a and b denote these sides, and let the others be c, d, e and f respectively (Fig. 5.). We may suppose that the graph of the pattern with this hexagon is the graph of Figure 6. As in the previous proof, we get that one of b and f must be equal to one of a and c. Because $a \neq b$, $a \neq f$ and $b \neq c$, this can be realized if and only if $c = f$. In this case the angles in the point B are adjacent angles of the hexagon, their sum is 2π, and consequently $c \| f$, i. e. the equal sides are opposite and parallel.

b) The condition is sufficient.

Let us reflect the hexagon to the midpoint of side a, and translate these two hexagons in such a way that side f coincides with c resp. e with f. Repeating these translations we get a ribbon, the length of which can exceed any value. Since the two broken lines forming the boundary of this ribbon are congruent with each other, we can rejoin this ribbon by translations from both sides to itself; and by repeating this we can cover an arbitrarily large part of the plane. (Fig. 6.)

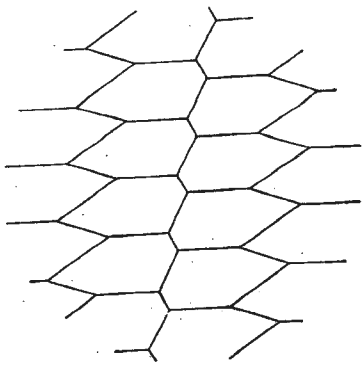


Fig. 6.

THEOREM 4. *If we can make a pattern with a hexagon, then the hexagon has three angles, the sum of which is 2π.*

PROOF: Suppose the contrary, then there is a hexagon, having no three angles with the sum 2π, suitable for making a pattern.

Let $\alpha_1, \alpha_2, \ldots \alpha_6$ denote the angles of the hexagon, and let us suppose, that $\left|\alpha_1 - \frac{2\pi}{3}\right| \geqq \left|\alpha_i - \frac{2\pi}{3}\right|$, where $i = 1, 2, \ldots 6$. It is obvious that $\alpha_1 \neq \frac{2\pi}{3}$, because otherwise every angle of the hexagon must be $\frac{2\pi}{3}$.

On account of theorem 1, we may consider an arbitrarily large part of the pattern, in which every node is second-rate, and three angles form an angle of 2π. It is impossible that the angles of a node should be different angles of the hexagon; consequently, two or three of them are corresponding. Taking a node, one angle of which is α_1, the angles in this node cannot be equal; hence we may suppose, that they are $\alpha_1, \alpha_1, \alpha_2$ or $\alpha_1, \alpha_2, \alpha_2$. From the first case it would follow, that $2\alpha_1 + \alpha_2 = 2\pi$, $2\left(\alpha_1 - \frac{2\pi}{3}\right) = \frac{2\pi}{3} - \alpha_2$; therefore $\left|\alpha_2 - \frac{2\pi}{3}\right| = 2\left|\alpha_1 - \frac{2\pi}{3}\right| > \left|\alpha_1 - \frac{2\pi}{3}\right|$, which is impossible. The hexagon has no angle equal to α_2, because in this case the sum of α_1, α_2 and of this angle (which naturally cannot be α_1) would be 2π. Consequently, every node of an angle α_1 is formed by two angles α_2 and by one angle α_1 only; that is to say, every domain of the chosen part of the pattern contains at least twice as many angles α_2 as α_1.

But this is impossible since it can be proved easily that in a sufficiently large part of the pattern the different angles are approximately of equal number.

This proves our theorem.

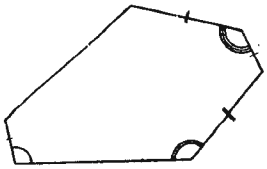


Fig. 7.

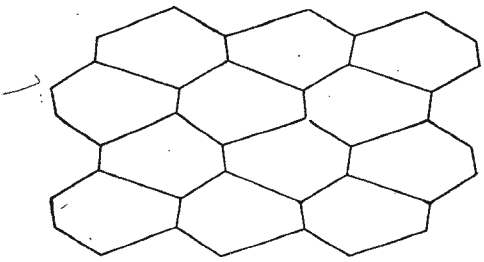


Fig. 8.

It can be proved, that if in a pattern with the hexagon $ABCDEF$ every node is second-rate, and surrounded by three hexagons, and in the nodes there are only two types of systems of angles (in one type of nodes the system $\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}$ forms the complete angle, and in the other nodes: $\alpha_{j_1}, \alpha_{j_2}, \alpha_{j_3}$, where $\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}, \alpha_{j_1}, \alpha_{j_2}, \alpha_{j_3}$ are the angles of hexagon), then either $AB \nparallel DE$ (Fig. 6.), or $FAB\sphericalangle + ABC\sphericalangle + CDE\sphericalangle = 2\pi$ with $FA = CD$ and $BC = DE$ (Fig. 7 and 8.).

It can be proved, too, as in theorems 2 and 3, that *if a hexagon has two sides, the lengths of which are different from the length of any other side, a pattern can be made with this hexagon if, and only if, the hexagon belongs to one of the above mentioned types of hexagon* (Fig. 6 and 7.).

We do not know as yet of pattern composed of another type of hexagon.

Literature

[1] H. S. M. Coxeter, *Introduction to Geometry*, New-York – London 1961.

[2] Б. Н. Делоне, Теория планигонов, *Изв. Акад. Наук СССР*, **23** (1959), 365-386.

[3] F. Haag, Die regelmässigen Planteilungen, *Z. Kristallogr.*, **49** (1911), 360 – 369.

[4] F. Haag, Die regelmässigen Planteilungen und Punktsysteme, *Z. Kristallogr.*, **58** (1924), 478 – 489.

[5] F. Laves, Ebenenteilung in Wirkungsbereiche, *Z. Kristallogr.*, **76** (1931) 277 – 284.

[6] J. Molnár, Körelhelyezések állandó görbületű felületeken, *MTA, III. Oszt. közl.*, **12** (1962), 223 – 263.

A CONTRIBUTION TO THE EXTENSIONS OF ABELIAN GROUPS

By

G. VILJOEN[1]

Department of Mathematics, Delft Technical University

(Received April 15, 1963)

1. Introduction

Throughout this paper, group will always denote an additively written abelian group. This paper is devoted to considering the following problems:[2])

1. Which are the torsion free groups L, such that Ext $(L, T)=0$ for every torsion group T?

2. If Ext $[L, C(\infty)]=0$, is L necessarily free?

The first problem was proposed by R. BAER [1] and he showed that a group L which satisfies the stated condition is $\aleph_1$-free. The second problem was proposed by J. H. C. WHITEHEAD. (See [4]). A. EHRENFEUCHT [4] showed that for countable L the answer is in the affirmative; J. ROTMAN [7] showed that groups L for which Ext $[L, C(\infty)]=0$, are $\aleph_1$-free.

We shall study the above mentioned groups L and derive some of their properties. We shall also discuss some problems related to (1) and (2).

I am greatly indebted to Professor L. FUCHS for valuable advice and critical comments on this paper.

2. Homological methods

In this section we give a brief outline of some results of homological algebra. For the proofs we refer to [2].

A sequence of groups and homomorphisms

$$L_1 \longrightarrow L_2 \longrightarrow \ldots\ldots \longrightarrow L_n \qquad (n \geqq 3)$$

is called *exact* if the image of L_i in L_{i+1} under $L_i \longrightarrow L_{i+1}$ coincides with the kernel of $L_{i+1} \longrightarrow L_{i+2}$ $(i=1, 2, \ldots, n-2)$. Note that $L \longrightarrow M$ is an isomorphism of L into M if and only if $0 \longrightarrow L \longrightarrow M$ is exact, whereas $M \longrightarrow N$ is a homomorphism of M onto N if and only if $M \longrightarrow N \longrightarrow 0$ is exact.

If L and K are two arbitrary groups, the set of all homomorphic mappings of L into K forms, under a suitable operation, a commutative group Hom (L, K), the *homomorphism group* of L into K. If L is the direct sum of groups $L_\lambda (\lambda \in \Lambda)$ then

$$\mathrm{Hom}(\sum_\lambda L_\lambda, K) \cong \sum_\lambda{}^* \mathrm{Hom}(L_\lambda, K)$$

[1]) The author gratefully acknowledges financial assistance received from the South African Council for Scientific and Industrial Research during his stay in Delft.

[2]) The notation and terminology is that of [5]. For the basic concepts concerning abelian groups we refer to [5].

and if K is the complete direct sum of the groups $K_\mu (\mu \in M)$ then

$$\mathrm{Hom}(L, \sum_\mu{}^* K_\mu) \cong \sum_\mu{}^* \mathrm{Hom}(L, K_\mu).$$

We have also

$$\mathrm{Ext}(\sum_\lambda L_\lambda, K) \cong \sum_\lambda{}^* \mathrm{Ext}(L_\lambda, K).$$

and

$$\mathrm{Ext}(L, \sum_\mu{}^* K_\mu) \cong \sum_\mu{}^* \mathrm{Ext}(L, K_\mu).$$

Let

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

be an exact sequence. This sequence gives rise to the following exact sequences

$$0 \longrightarrow \mathrm{Hom}(C, K) \longrightarrow \mathrm{Hom}(B, K) \longrightarrow \mathrm{Hom}(A, K) \longrightarrow$$
$$\longrightarrow \mathrm{Ext}(C, K) \longrightarrow \mathrm{Ext}(B, K) \longrightarrow \mathrm{Ext}(A, K) \longrightarrow 0$$

and

$$0 \longrightarrow \mathrm{Hom}(L, A) \longrightarrow \mathrm{Hom}(L, B) \longrightarrow \mathrm{Hom}(L, C) \longrightarrow$$
$$\longrightarrow \mathrm{Ext}(L, A) \longrightarrow \mathrm{Ext}(L, B) \longrightarrow \mathrm{Ext}(L, C) \longrightarrow 0$$

for all groups K and L. For any group L, the sequence

$$L \otimes A \longrightarrow L \otimes B \longrightarrow L \otimes C \longrightarrow 0 \tag{1}$$

is also exact. If L is torsion free, we may add $0 \longrightarrow$ to the left of (1).

R will always denote the additive group of rational numbers and $C(n)$ will denote the cyclic group of order $n = 1, 2, \ldots$ or ∞. $C(p^\infty)$ is the quasi-cyclic group where of course p is a prime number. We have $R/C(\infty) \cong \sum C(p^\infty)$ where the direct sum is taken over all primes p. If S is any set, $|S|$ will denote its cardinality.

3. The main theorems[3])

Quite recently CHASE [3] has shown that if $\mathrm{Ext}[L, C(\infty)] = 0$ and L is of infinite rank then $\mathrm{Hom}[L, C(\infty)] = 2^{|L|}$. We shall now give another simple proof of this fact for those groups L which satisfy $|L|^{\aleph_0} = |L|$. First we observe that if L is a group of cardinality $\mathfrak{m}$ and K is a group of cardinality $\mathfrak{n}$, then

$$|\mathrm{Ext}(L, K)| \leqq \mathfrak{n}^{\mathfrak{m}} \quad \text{and} \quad |\mathrm{Hom}(L, K)| \leqq \mathfrak{n}^{\mathfrak{m}}. \tag{A}$$

In the sequel we shall frequently make use of these inequalities. We shall need the following:

LEMMA 1. $|\mathrm{Ext}[R, C(\infty)]| = \aleph$. ($\aleph$ denotes the power of the continuum.)

PROOF. We have the exact sequence

$$0 \longrightarrow C(\infty) \longrightarrow R \longrightarrow R/C(\infty) \longrightarrow 0$$

[3]) We shall assume that the Generalized Continuum Hypothesis holds.

from which we obtain the exact sequence

$$0 \longrightarrow \mathrm{Hom}\,[C(\infty), C(\infty)] \longrightarrow \mathrm{Ext}\,[R/C(\infty), C(\infty)] \longrightarrow \mathrm{Ext}\,[R, C(\infty)] \longrightarrow 0$$

since R and $R/C(\infty)$ are divisible and $C(\infty)$ is free. We next make use of the following result [5, p. 244]: If B is a torsion free and A a torsion group, then

$$\mathrm{Ext}\,(A, B) \cong \mathrm{Hom}\,(A, D/B) \tag{2}$$

where D is a minimal divisible group containing B. From (2) it follows that

$$0 \longrightarrow C(\infty) \longrightarrow \mathrm{Hom}\,[R/C(\infty), R/C(\infty)] \longrightarrow \mathrm{Ext}\,[R, C(\infty)] \longrightarrow 0$$

is exact. Hence our assertion follows since Hom $[R/C(\infty), R/C(\infty)]$ is of the power $\aleph$.

THEOREM 1. *If* $\mathrm{Ext}\,[L, C(\infty)]=0$ *and* $|L|=\mathfrak{m}>\aleph_0$, $\mathfrak{m}^{\aleph_0}=\mathfrak{m}$, *then* $|\mathrm{Hom}\,[L, C(\infty)]|=2^{\mathfrak{m}}$.

REMARK. If L is countable, then it is free, and consequently $\mathrm{Hom}[L, C(\infty)] \cong L$ if L has finite rank. If L has countably infinite rank, then obviously we have $|\mathrm{Hom}\,[L, C(\infty)]|=2^{\mathfrak{m}}$. Consequently we may assume that $|L|=\mathfrak{m}>\aleph_0$.

PROOF. Now $|\mathrm{Hom}\,[L, C(\infty)]| \leqq \aleph_0{}^{\mathfrak{m}}=2^{\mathfrak{m}}$. Assume that $|\mathrm{Hom}\,[L, C(\infty)]| = \mathfrak{n}<2^{\mathfrak{m}}$, i. e. $\mathfrak{n} \leqq \mathfrak{m}$. We have the following isomorphism [2, p. 116]:

$$\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\} + \mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\} \cong$$

$$\cong \mathrm{Ext}\,[R \otimes L, C(\infty)] + \mathrm{Hom}\,[\mathrm{Tor}\,(R, L), C(\infty)]. \tag{3}$$

that is, we have

$$\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\} \cong \mathrm{Ext}\,[R \otimes L, C(\infty)]$$

since $\mathrm{Ext}\,[L, C(\infty)]=0$ and $\mathrm{Tor}\,(R, L)=0$[4]) because R and L are torsion free (see [4, chapter VII]). Let us consider $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}|$. We have $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| \leqq \mathfrak{n}^{\aleph_0}$. If $\mathfrak{n}=\aleph_0$, then $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| \leqq \aleph$ and if $\mathfrak{n}>\aleph_0$, we have $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| \leqq \mathfrak{n}^{\aleph_0} \leqq \mathfrak{m}^{\aleph_0} = \mathfrak{m} < 2^{\mathfrak{m}}$. At any rate, we see that our assumption on $|\mathrm{Hom}\,[L, C(\infty)]|$ implies that $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| < 2^{\mathfrak{m}}$.

However, $R \otimes L$ is a minimal divisible group containing L [5, p. 256], that is $R \otimes L \cong \sum_{\mathfrak{m}} R$ and

$$\mathrm{Ext}\,[R \otimes L, C(\infty)] \cong \mathrm{Ext}\left(\sum_{\mathfrak{m}} R, C(\infty)\right) \cong \sum_{\mathfrak{m}}{}^{*} \mathrm{Ext}\,[R, C(\infty)].$$

From Lemma 1 and our assumption on $|L|=\mathfrak{m}$ we obtain $|\mathrm{Ext}\,[R \otimes L, C(\infty)]|=2^{\mathfrak{m}}$.

The contradiction thus obtained shows that $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$. This establishes the proof of the theorem.

From Theorem 1 we obtain the following (see [3, p. 698]):

COROLLARY 1.1. *If* $\mathrm{Ext}\,[L, C(\infty)] = 0$, $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, *then* $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$[5]) *for every prime* p.

[4]) For the definition and properties of the torsion product we refer to [2].

[5]) If G is a group, $\mathfrak{r}(G)$ denotes its rank.

PROOF. Let $R^{(p)}$ denote the subgroup of R generated by $p^{-1}, p^{-2}, \ldots, p^{-n}, \ldots$. Then $R^{(p)}/C(\infty) \cong C(p^\infty)$. From the exact sequence

$$0 \longrightarrow C(\infty) \longrightarrow R^{(p)} \longrightarrow C(p^\infty) \longrightarrow 0$$

we obtain the exact sequence

$$0 \longrightarrow L \otimes C(\infty) \cong L \longrightarrow L \otimes R^{(p)} \longrightarrow L \otimes C(p^\infty) \longrightarrow 0.$$

Since $pR^{(p)} = R^{(p)}$, clearly $p(L \otimes R^{(p)}) = L \otimes R^{(p)}$. This exact sequence induces exactness of

$$0 \longrightarrow \mathrm{Hom}\,[L, C(\infty)] \longrightarrow \mathrm{Ext}\,[L \otimes C(p^\infty), C(\infty)] \longrightarrow$$
$$\longrightarrow \mathrm{Ext}\,[L \otimes R^{(p)}, C(\infty)] \longrightarrow 0,$$

that is, (by (2) and [5, p. 255])

$$0 \longrightarrow \mathrm{Hom}\,[L, C(\infty)] \longrightarrow \mathrm{Hom}\,[L \otimes C(p^\infty), R/C(\infty)] \longrightarrow$$
$$\longrightarrow \mathrm{Ext}\,[L \otimes R^{(p)}, C(\infty)] \longrightarrow 0$$

is exact. But $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$ and hence it follows from the above exact sequence that $|\mathrm{Hom}\,[L \otimes C(p^\infty), R/C(\infty)]| \geqq 2^{\mathfrak{m}}$. However, from the definition of the tensor product it follows that $|L \otimes C(p^\infty)| \leqq \mathfrak{m}$ and from (A) it follows that $|\mathrm{Hom}\,[L \otimes C(p^\infty), R/C(\infty)]| \leqq 2^{\mathfrak{m}}$. Consequently

$$|\mathrm{Hom}\,[L \otimes C(p^\infty), R/C(\infty)]| = 2^{\mathfrak{m}}.$$

Moreover, it follows now also that $|L \otimes C(p^\infty)| = \mathfrak{m}$.

On the other hand, it is known that $L \otimes C(p^\infty) \cong \sum_{\mathfrak{n}} C(p^\infty)$ where $\mathfrak{n}$ denotes the rank of L/pL, $\mathfrak{n} = \mathfrak{r}(L/pL)$ [5, p. 255]. Thus we have

$$|L \otimes C(p^\infty)| = \left|\sum_{\mathfrak{n}} C(p^\infty)\right| = \mathfrak{n}\aleph_0 = \max(\mathfrak{n}, \aleph_0) = \mathfrak{n}.$$

Hence $\mathfrak{n} = \mathfrak{m}$. This completes the proof since $\mathfrak{r}(L) = \mathfrak{m}$.

COROLLARY 1.2. *If L is torsion free, $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, and if $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$, then $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$ and $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$ for every prime p.*

REMARK. It is known [3, p. 694, Theorem 4.4] that if L is torsion free and $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$, then L is $\aleph_1$-free. Hence we may assume that $|L| > \aleph_0$.

PROOF. Because R and L are torsion free, $\mathrm{Tor}\,(R, L) = 0$. Hence by (3) we have

$$\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\} + \mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\} \cong \mathrm{Ext}\,[R \otimes L, C(\infty)].$$

It is known (see the proof of Theorem 1) that

$$|\mathrm{Ext}\,[R \otimes L, C(\infty)]| = \left|\sum_{\mathfrak{m}}{}^{*} \mathrm{Ext}\,[R, C(\infty)]\right| = 2^{\mathfrak{m}}.$$

Put $|\mathrm{Hom}\,[L, C(\infty)]| = \mathfrak{n}$. Since $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$ it follows from (A) that $|\mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\}| \leqq \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0}$. From the above isomorphism it follows that

$$|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| + |\mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\}| = 2^{\mathfrak{m}}.$$

Since we have assumed that $\mathfrak{m} > \aleph_0$, it is also clear that $|\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\}| = 2^{\mathfrak{m}}$. We conclude [by (A)] that

$$|\mathrm{Hom}\,[L, C(\infty)]| = \mathfrak{n} = 2^{\mathfrak{m}}.$$

From the exact sequence (see Corollary 1.1)

$$0 \longrightarrow L \longrightarrow L \otimes R^{(p)} \longrightarrow L \otimes C(p^{\infty}) \longrightarrow 0$$

we obtain the exact sequence

$$0 \longrightarrow \mathrm{Hom}\,[L, C(\infty)] \longrightarrow \mathrm{Ext}\,[L \otimes C(p^{\infty}), C(\infty)] \cong$$
$$\cong \mathrm{Hom}\,[L \otimes C(p^{\infty}), R/C(\infty)].$$

From this exact sequence we obtain $|\mathrm{Ext}\,[L \otimes C(p^{\infty}), C(\infty)]| = 2^{\mathfrak{m}}$. The proof that $\mathfrak{r}(L/pL) = \mathfrak{r}(L) = \mathfrak{m}$ follows now in exactly the same way as in the corresponding part of the proof of Corollary 1.1. This establishes the proof.

REMARK. If L is torsion free, $|L| > \aleph_0$, and $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$, let U_f denote a pure subgroup of L of finite rank. Then U_f is free and L/U_f is torsion free. The exact sequence

$$0 \longrightarrow U_f \longrightarrow L \longrightarrow L/U_f \longrightarrow 0$$

implies the exactness of

$$\mathrm{Hom}\,[U_f, C(\infty)] \longrightarrow \mathrm{Ext}\,[L/U_f, C(\infty)] \longrightarrow \mathrm{Ext}\,[L, C(\infty)] \longrightarrow 0.$$

Since $\mathrm{Hom}\,[U_f, C(\infty)]$ as well as $\mathrm{Ext}\,[L, C(\infty)]$ is countable, it follows that $|\mathrm{Ext}\,[L/U_f, C(\infty)]| < 2^{\aleph_0}$.

This remark and Corollary 1.2 lead to the following plausible conjecture:

If L is torsion free and $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$, then $\mathrm{Ext}\,[L, C(\infty)] = 0$.

Suppose that every torsion free group L such that $|\mathrm{Ext}\,[L, C(\infty)]| < 2^{\aleph_0}$ satisfies the following condition:

Every pure subgroup of L of finite rank is a direct summand of L.

Then clearly $\mathrm{Ext}\,[L, C(\infty)] = 0$. (See also [7, p. 251]).

The following statements are easy consequences of Theorem 1 and its corollaries:

(i) If L is torsion free and $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, then either $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$, or $|\mathrm{Ext}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$ or both.

(ii) If L is an $\aleph_1$-free group, $|L| = \mathfrak{m} > \aleph_0$, and $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$, then $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$ for every prime p.

(iii) If $L = \sum_{\aleph_0}{}^{*} C(\infty)$ then L is $\aleph_1$-free but not free, [5, p. 168]. However,

$\mathrm{Ext}\,[L, C(\infty)] \neq 0$. (See [7]). This follows also easily from our observations:

As a matter of fact, it is known that $|\mathrm{Hom}\,[L, C(\infty)]| = \aleph_0$ [5, p. 228], and from the isomorphism (see (3))

$$\mathrm{Ext}\,\{R, \mathrm{Hom}\,[L, C(\infty)]\} + \mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\} \cong \mathrm{Ext}\,[R \otimes L, C(\infty)]$$

and the fact that $|\mathrm{Ext}\,[R \otimes L, C(\infty)]| = 2^{\aleph}$, we deduce

$|\mathrm{Hom}\,\{R, \mathrm{Ext}\,[L, C(\infty)]\}| = 2^{\aleph}$ whence $|\mathrm{Ext}\,[L, C(\infty)]| = 2^{\aleph}$, indeed.

COROLLARY 1.3. *If L is reduced, $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, and* $\mathrm{Ext}\,[L, C(\infty)]$ *is torsion free and if, moreover, $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$ for at least one prime p, then* $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$. *Conversely, if L is reduced, $|L| = \mathfrak{m} > \aleph_0$, and* $\mathrm{Ext}\,[L, C(\infty)]$ *is torsion free and if* $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$, *then $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$ for every prime p.*

PROOF. It is known [3, p. 693] that if L is reduced and $\mathrm{Ext}\,[L, C(\infty)]$ is torsion free, then L is $\aleph_1$-free and every pure subgroup of L of finite rank is a direct summand of L. Assume that $\mathfrak{r}(L/pL) = \mathfrak{r}(L)$ for a prime p. Then it follows from the isomorphism (3) that

$$\mathrm{Ext}\,\{C(p^\infty), \mathrm{Hom}\,[L, C(\infty)]\} \cong \mathrm{Ext}\,[C(p^\infty) \otimes L, C(\infty)]$$

since $\mathrm{Ext}\,[L, C(\infty)]$ and L are torsion free. By [5, p. 255] and our assumption it follows that $C(p^\infty) \otimes L \cong \sum_{\mathfrak{m}} C(p^\infty)$ and hence $|\mathrm{Ext}\,[\sum_{\mathfrak{m}} C(p^\infty), C(\infty)]| = |\mathrm{Hom}\,[\sum_{\mathfrak{m}} C(p^\infty), R/C(\infty)]| = 2^{\mathfrak{m}}$. Thus $|\mathrm{Ext}\,\{C(p^\infty), \mathrm{Hom}\,[L, C(\infty)]\}| = 2^{\mathfrak{m}}$, whence $|\mathrm{Hom}\,[L, C(\infty)]| = 2^{\mathfrak{m}}$.

The converse of the corollary follows from (ii) above, completing the proof.

Now we turn our attention to the first problem mentioned in section 1. We prove:

THEOREM 2. *If $\mathrm{Ext}\,(L, T) = 0$ for every torsion group T, $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, then $\mathfrak{r}(L/pL) = \mathfrak{r}(L) = \mathfrak{m}$ for every prime p.*

For the proof we need the following:

LEMMA 2. *If* $\mathrm{Ext}\,(L, T) = 0$ *for every torsion group T and S is a reduced and unbounded torsion group, $|S| \leq |L| = \mathfrak{m}$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, then* $|\mathrm{Hom}\,(L, S)| = 2^{\mathfrak{m}}$.

REMARK. Acutally we prove much more than we need.

If S is not reduced, then it contains a direct summand $C(p^\infty)$. Let G be an arbitrary torsion free group and $|G| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$. Then $|\mathrm{Hom}\,(G, S)| = 2^{\mathfrak{m}}$. This follows from (A) and the following observation [2, p. 28]:

$$\mathrm{Hom}\,\{R, \mathrm{Hom}\,[G, C(p^\infty)]\} \cong \mathrm{Hom}\,[R \otimes G, C(p^\infty)] \cong \sum_{\mathfrak{m}}{}^{*}\, \mathrm{Hom}\,[R, C(p^\infty)]$$

and $|\sum_{\mathfrak{m}}{}^{*}\, \mathrm{Hom}\,[R, C(p^\infty)]| = 2^{\mathfrak{m}}$.

PROOF OF LEMMA 2. First we prove that $\mathrm{Ext}\,(R, S) \neq 0$. To this end, let S be an arbitrary reduced and unbounded torsion group which satisfies $|S| \leq \mathfrak{m}$. Because of the $\aleph_1$-freeness of L we have assumed $\mathfrak{m} > \aleph_0$.

Let $S = \sum_{i\in I} S_{p_i}$ be a decomposition of S into its primary components (S_{p_i} denotes the p_i-component of S). Since S is unbounded, either $|I| = \aleph_0$, or S contains an unbounded p-component S_p. In the first case each p-component S_{p_i} contains a cyclic direct summand $C(p_i^{k_i})$ [5, p. 80] and hence, if we put $S_{p_i} = C(p_i^{k_i}) + S'_{p_i}$, we have $S = \sum_{i\in I} C(p_i^{k_i}) + \sum_{i\in I} S'_{p_i}$, $|I| = \aleph_0$. Let $T_1 = \sum_{i\in I} C(p_i^{k_i})$ and write $T_1^* = \sum_{i\in I}^* C(p_i^{k_i})$. Then clearly T_1 is the maximal torsion subgroup of T_1^* and T_1^*/T_1 is divisible and torsion free. T_1^*/T_1 does not split since both T_1^* and T_1 are reduced. Consequently $\mathrm{Ext}\,(R, S) \neq 0$. Moreover, $|\mathrm{Ext}\,(R, T_1)| = \aleph$ for the exact sequence

$$0 \longrightarrow T_1 \longrightarrow T_1^* \longrightarrow T_1^*/T_1 \longrightarrow 0$$

implies the exactness of

$$0 \longrightarrow \mathrm{Hom}\,(R, T_1^*/T_1) \longrightarrow \mathrm{Ext}\,(R, T_1) \longrightarrow \sum_{i\in I}^* \mathrm{Ext}\,[R, C(p_i^{k_i})] = 0$$

since a bounded pure subgroup is a direct summand [5, p. 80]. Here we have $\mathrm{Hom}\,(R, T_1^*/T_1) \cong T_1^*/T_1$ for the exact sequence

$$0 \longrightarrow C(\infty) \longrightarrow R \longrightarrow R/C(\infty) \longrightarrow 0$$

gives rise to the exact sequence

$$0 \longrightarrow \mathrm{Hom}\,(R, T_1^*/T_1) \longrightarrow \mathrm{Hom}\,[C(\infty), T_1^*/T_1] \cong T_1^*/T_1 \longrightarrow 0$$

since T_1^*/T_1 is torsion free and divisible. Because $|T_1^*/T_1| = \aleph$ it follows that $|\mathrm{Ext}\,(R, T_1)| = \aleph$ and hence $|\mathrm{Ext}\,(R, S)| \geqq \aleph$.

If S contains an unbounded p-component S_p, put $S = S_p + A$. Let B denote a basic subgroup[6]) of S_p and write $B = \sum_{n=1}^{\infty} B_n$ where $B_n = \sum C(p^n)$ ($n = 1, 2, \ldots$).

Let $\overline{T}$ denote the maximal torsion subgroup of $B^* = \sum_{n=1}^{\infty}{}^* B_n$. Then $\overline{T}/B$ is divisible [5, p. 100] and thus $B^*/B = \overline{T}/B + M/B$ ($M \cap \overline{T} = B$, $\{M, T\} = B^*$), where M/B is torsion free. It can easily be shown that M/B contains a non zero divisible subgroup. However, M/B does not split since both B^* and B are reduced. It can also be shown that $|\mathrm{Ext}\,(R, B)| \geqq \aleph_0$. But from the exact sequence [5, p. 106] $S_p \to B \to 0$ we obtain the exact sequence

$$\mathrm{Ext}\,(R, S_p) \longrightarrow \mathrm{Ext}\,(R, B) \longrightarrow 0$$

whence $\mathrm{Ext}\,(R, S_p) \neq 0$ and $|\mathrm{Ext}\,(R, S_p)| \geqq \aleph_0$. Hence $\mathrm{Ext}\,(R, S) \neq 0$ and $|\mathrm{Ext}\,(R, S)| \geqq \aleph_0$.

By making use of the isomorphism (3) we obtain

$$\mathrm{Ext}\,[R, \mathrm{Hom}(L, S)] \cong \mathrm{Ext}\,(R \otimes L, S)$$

[6]) For the definition and properties of basic subgroups see [5].

since $\mathrm{Ext}(L, S) = 0$ and $\mathrm{Tor}(R, L) = 0$. Moreover, $|\mathrm{Ext}(R \otimes L, S)| = |\sum_{\mathfrak{m}}{}^{*}\mathrm{Ext}(R, S)| = 2^{\mathfrak{m}}$. From (A) it follows that $|\mathrm{Ext}(R, \mathrm{Hom}(L, S)| \leqq |\mathrm{Hom}(L, S)|^{\aleph_0}$ whence $|\mathrm{Hom}(L, S)| = 2^{\mathfrak{m}}$. This completes the proof of the lemma.

PROOF OF THEOREM 2. Let $T = \sum_{n=1}^{\infty} C(p^n)$. Then it follows from the above lemma that $|\mathrm{Hom}(L, T)| = 2^{\mathfrak{m}}$. We assume of course that $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$. Then from the exact sequence (see Corollary 1.1)

$$0 \longrightarrow C(\infty) \longrightarrow R^{(p)} \longrightarrow C(p^{\infty}) \longrightarrow 0$$

we derive the exact sequence

$$0 \longrightarrow L \otimes C(\infty) \cong L \longrightarrow L \otimes R^{(p)} \longrightarrow L \otimes C(p^{\infty}) \longrightarrow 0$$

from which we obtain the exact sequence

$$0 \longrightarrow \mathrm{Hom}(L, T) \longrightarrow \mathrm{Ext}[L \otimes C(p^{\infty}), T] \longrightarrow \mathrm{Ext}(L \otimes R^{(p)}, T) \longrightarrow 0.$$

Clearly $\mathrm{Hom}(L \otimes R^{(p)}, T) = 0$ because of $p(L \otimes R^{(p)}) = L \otimes R^{(p)}$. From this exact sequence it follows that $|\mathrm{Ext}[L \otimes C(p^{\infty}), T]| = 2^{\mathfrak{m}}$. That $\mathfrak{r}(L/pL) = \mathfrak{r}(L) = \mathfrak{m}$ follows now in the same way as in the corresponding part of the proof of Corollary 1.1, completing the proof.

COROLLARY 2.1. *If* $\mathrm{Ext}(L, T) = 0$ *for every torsion group* T, $|L| = \mathfrak{m} > \aleph_0$, $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$, *then* $\mathfrak{r}[(L \otimes L)/p(L \otimes L)] = \mathfrak{r}(L \otimes L) = \mathfrak{m}$ *for every prime* p.

PROOF. Since $\mathfrak{r}(L/pL) = \mathfrak{r}(L) = \mathfrak{m}$, the corollary is a consequence of the following theorem [6, p. 13, Theorem 2.3]:

If U and V are torsion free groups, then $U \otimes V$ is again torsion free and for any prime p we have

$$\mathfrak{r}[(U \otimes V)/p(U \otimes V)] = \mathfrak{r}(U/pU)\mathfrak{r}(V/pV).$$

Literature

[1] R. BAER, The subgroup of the elements of finite order of an abelian group, *Ann. of Math.*, **37** (1936), 766–781.

[2] H. CARTAN, and S. EILENBERG, *Homological Algebra*, Princeton (1956).

[3] S. U. CHASE, Locally free modules and a problem of Whitehead, *Ill. Journ. of Math.*, **6** (1962), 682–699.

[4] A. EHRENFEUCHT, On a problem of J. H. C. Whitehead concerning abelian groups, *Bull. Acad. Polon. Sci.*, Cl. III, **3** (1955), 127–128.

[5] L. FUCHS, *Abelian gropus*, Budapest (1958).

[6] L. FUCHS, Notes on abelian groups, I., *Annales Univ. Sci. Budapest*, Sectio Math., **2** (1959), 5–23.

[7] J. ROTMAN, On a problem of Baer and a problem of Whitehead in abelian groups, *Acta Math. Acad. Sci. Hung.*, **12** (1961), 245–254.

INDEX

TOMUS VI

ACZÉL, J., Remarks on probable inference 3
BOLLOBÁS, B., Filling the plane with congruent convex hexagons without overlapping 117
BÖRÖCZKY, K., Über in einem Kreisring enthaltene Polygone 113
CSÁSZÁR, Á., Transposition de structures syntopogènes 55
CSÁSZÁR, Á. et CZIPSZER, J. Sur des critères généraux d'approximation uniforme 17
CZIPSZER, J. et CSÁSZÁR, Á., Sur des critères généraux d'approximation uniforme 17
FUCHS, L., Ranks of modules 71
FUCHS, L. and STEINFELD, O., Principal components and prime factorization in partially ordered semigroups 103
JUHÁSZ, I., On extremal values of mappings. I 39
KÁTAI, I., On certain sets of integers 43
КАТАИ, И., Асимптотическая формула в теории чисел 83
KERTÉSZ, A., On ranks of modules 79
MIKOLÁS, M., Generalized Euler sums and the semigroup property of integro-differential operators 89
RÉNYI, A. and RÉNYI, C., On „small" coefficients of the power series of entire functions 27
RÉNYI, C. and RÉNYI, A., On „small" coefficients of the power series of entire functions 27
STEINFELD, O. and FUCHS, L., Principal components and prime factorization in partially ordered semigroups 103
STROHMAJER, J., Über die Verteilung von Punkten auf der Kugel 49
SZÁSZ, P., On generalized quasi-step and almost-step parabolas, respectively 13
VILJOEN, G., A contribution to the extensions of abelian groups 125

Technikai szerkesztő: Scharnitzky Viktor