A 21267 D

Mathematische Zeitschrift

Band 128 · Heft 1 · 1972

Springer-Verlag · Berlin · Heidelberg · New York

Math. Z. – MAZEAX 128(1) 1–92 (1972) 11. X. 1972

Die **„Mathematische Zeitschrift“** wurde im Jahre 1918 von *L. Lichtenstein* unter der Mitwirkung von *K. Knopp, E. Schmidt* und *I. Schur* gegründet und herausgegeben. Nach dem Tode Lichtensteins übernahm *K. Knopp* 1933 die Herausgabe. Die Schriftleitung ergänzte sich 1933 durch *E. Kampke* und *F. K. Schmidt*, 1936 durch *R. Nevanlinna* und 1950 durch *H. Wielandt*, der 1952 die Herausgabe übernahm.

„Mathematische Zeitschrift“ was founded in 1918 and edited by *L. Lichtenstein* in cooperation with *K. Knopp, E. Schmidt* and *I. Schur*; after Lichtenstein's death, 1933, it was edited by *K. Knopp*. The Editorial Committee was increased to include *E. Kamke* and *F. K. Schmidt* in 1933, *R. Nevanlinna* in 1936 and *H. Wielandt* in 1950. The latter became Managing Editor in 1952.

*

Die Mathematische Zeitschrift dient der Pflege der reinen Mathematik, doch werden auch Beiträge aus den Gebieten der theoretischen Physik und Astronomie Aufnahme finden, soweit sie mathematisch von Interesse sind. Besprechungen, Aufgaben u. dgl. werden nicht zugelassen.

Von jeder Arbeit werden 75 Sonderdrucke kostenlos zur Verfügung gestellt, weitere können zum Selbstkostenpreis bezogen werden.

"Mathematische Zeitschrift" is devoted primarily to pure mathematics; papers on theoretical physics and astronomy may be accepted if they present interesting mathematical results. Reviews, problems etc. will not be published.

75 reprints of each paper are provided free of charge; additional copies may be ordered at cost price.

Manuskripte nehmen entgegen / Manuscripts may be sent to:

*

Die Zeitschrift erscheint, um eine rasche Publikation zu ermöglichen, in einzelnen Heften, die zu Bänden vereinigt werden. Ein Band besteht im allgemeinen aus 4 Heften. Der Preis eines Bandes beträgt DM 112.—.

In the interest of speedy publications, this journal is issued at frequent intervals according to the material received. As a rule four numbers constitute one volume. The price is DM 112.— per volume.

Springer-Verlag

D-6900 Heidelberg 1
Postfach 1780
Tel. (06221) 49101
Telex 04-61723

D-1000 Berlin 33
Heidelberger Platz 3
Tel. (0311) 822001
Telex 01-83319

Springer-Verlag
New York Inc.
175 Fifth Avenue
New York, N.Y. 10010

Math. Z. 128, 1 – 14 (1972)

On Groups with Several Doubly-Transitive Permutation Representations

Peter J. Cameron

A group which has two inequivalent (finite) doubly transitive permutation representations with the same permutation character is an automorphism group of a projective design. So a group with more than two such representations is an automorphism group of each of the designs obtained by taking the representations two at a time, and preserves interrelations among these designs. The combinatorial object abstracted from this situation, tentatively called a system of linked projective designs, is studied, and the results applied to the groups.

In section 1, some basic properties of projective designs, and their construction from two doubly transitive representations of a group, are described. Section 2 defines systems of linked projective designs, and begins a combinatorial analysis of these systems, establishing relations connecting their parameters. In section 3, a simple consequence of these relations is noted, and some results on automorphisms of doubly transitive groups are deduced from it. (This section was inspired by Wielandt's paper [8] on the same topic.) Section 4 gives a collection of equations satisfied by the parameters of a system of three linked designs. I derive no consequences of these equations, but I give them in the hope that they may be useful in further work on this subject. Section 5 contains an account of an application to a conjecture of Wielandt (that a transitive permutation group of prime degree p has at most two inequivalent representations of degree p); while I cannot prove the conjecture, the equations are well suited to computation, and Dr. Peter Neumann has used them to show that the conjecture is true for primes less than two million.

I am grateful to Dr. Neumann for the large amount of time and energy he has devoted to this computation, and for many helpful discussions; and to Professor Wielandt, for his interest in this work, for many helpful suggestions which improved the clarity of this paper, and for several unpublished results. Also I acknowledge a debt to my wife (it was on our honeymoon that this line of attack on the problem occurred to me); and to Merton College, Oxford, where I hold a Junior Research Fellowship.

§ 1. Projective Designs

A *projective design* (see [4], pp. 57, 58) is a set Ω_1 of v elements called "points" and a set Ω_2 of v elements called "blocks" with an incidence relation between Ω_1 and Ω_2, such that for some integers k, λ with $1<k<v-1$,

(i) any block is incident with k points, and any point with k blocks;

(ii) any two blocks are incident with λ common points, and any two points with λ common blocks.

Put $n=k-\lambda$. Then we have ([4], pp. 57, 59)

$$(v-1)\lambda=k(k-1), \tag{1}$$

$$(v-1)n=k(v-k), \tag{2}$$

$$v\lambda=k^2-n. \tag{3}$$

We say the design has parameters (v, k, λ). The *dual* design (obtained by interchanging the labels "point" and "block") has the same parameters. The *complementary* design (obtained by replacing the incidence relation by its complement) has parameters (v, k', λ') where $k'=v-k$; by (2), $k-\lambda=k'-\lambda'$.

Let A be the incidence matrix of the design. (The rows of A are indexed by the points, and the columns by the blocks; the (α, β) entry is 1 if α and β are incident, and 0 otherwise.) The matrix A satisfies

$$AJ=A^TJ=kJ, \tag{4}$$

$$AA^T=A^TA=nI+\lambda J \tag{5}$$

where I is the identity matrix, J the matrix with every entry 1, and T denotes transpose ([4], p. 63). The incidence matrix of the dual design is A^T; that of the complementary design is $J-A$.

If v is even, then n is the square of an integer ([4], 2.1.8). There is a partial converse: if n is the square of an integer, then v is composite. (For suppose that v is prime and $n=t^2$ for some positive integer t. Then from (3), $v\lambda=k^2-t^2$, so v divides $k-t$ or $k+t$. This requires either $k-t=0$ or $k+t=v$. The first alternative contradicts $k-t\geqq k-t^2=\lambda>0$; the second implies $k-t=\lambda=k-t^2$, so $t=0$ or 1, $k=v$ or $v-1$, contradicting $k<v-1$.)

If a finite group G acts transitively on two sets Ω_1 and Ω_2 with $|\Omega_1|=|\Omega_2|=v$, and G is doubly transitive on Ω_1, then in its componentwise action on $\Omega_1\times\Omega_2$, G has two orbits or is transitive (equivalently, for $\omega\in\Omega_2$, G_ω has two orbits or is transitive on Ω_1) according as the permutation characters of the representations of G on Ω_1 and Ω_2 are equal or unequal. (See Itô [5].) Suppose the first case occurs, but the two representations are inequivalent. Then G is doubly transitive on Ω_2. If O and O' are the orbits of G in $\Omega_1 \times \Omega_2$, we can define a projective design with point set Ω_1 and block set Ω_2 by letting O be the incidence relation. (We have $1<k<v-1$, since if $k=1$ or $k=v-1$ then the two representations would be equivalent.) G is an automorphism group of the design; it is doubly transitive on points and on blocks, and transitive on incident point-block pairs and on non-incident point-block pairs. Replacing O by O' yields the complementary design. Any projective design with an automorphism group doubly transitive on points arises in this way.

§ 2. Linked Projective Designs

We shall be concerned with combinatorial objects tentatively called *systems of linked projective designs*. Such a system is defined as a collection $\{\Omega_1, \ldots, \Omega_f\}$ of sets ($f \geqq 3$) and an incidence relation between each pair of sets such that

(i) each pair $\{\Omega_i, \Omega_j\}$ with its incidence relation is a projective design;

(ii) given three of the sets, say $\Omega_i, \Omega_j, \Omega_k$, the number of elements $\alpha \in \Omega_i$ incident with both $\beta \in \Omega_j$ and $\gamma \in \Omega_k$ depends only on whether or not β and γ are incident.

If G is a finite group with f pairwise inequivalent doubly transitive permutation representations (on sets $\Omega_1, \ldots, \Omega_f$) where $f \geqq 3$, all with the same permutation character, then $\{\Omega_1, \ldots, \Omega_f\}$ can be given the structure of a system of linked projective designs by defining the incidence relation between Ω_i and Ω_j to be an orbit of G in $\Omega_i \times \Omega_j$ for each pair $\{i, j\}$. Condition (i) follows from the remarks in the last section; (ii) from the fact that G preserves the incidence relations and is transitive on incident pairs and on non-incident pairs from $\{\Omega_j, \Omega_k\}$.

A *subsystem* of a system of linked projective designs is a subcollection of the sets (containing at least three sets) together with the appropriate incidence relations; clearly it forms a system in its own right. A system is *maximal* if it is not isomorphic to a proper subsystem of any system. A collection of sets with incidence relations is a system of linked projective designs if and only if every triple of sets is a system. So we first consider systems with $f=3$.

Theorem 1. *Suppose that in a system of linked projective designs with $f=3$, the designs have parameters (v, k_i, λ_i), and $n_i = k_i - \lambda_i$, $i=1,2,3$. Then there are integers u_1, u_2, u_3 such that*

(i) $n_1 = u_2 u_3$, $n_2 = u_3 u_1$, *and* $n_3 = u_1 u_2$;

(ii) v *divides* $k_2 k_3 + u_1 k_1$, $k_3 k_1 + u_2 k_2$, *and* $k_1 k_2 + u_3 k_3$.

Proof. Consider the designs as the edges in the following triangle, with the direction of the arrow from points to blocks:

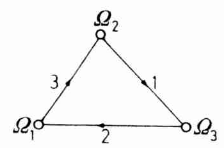


Fig. 1

Let the number of elements of Ω_1 incident with $\beta \in \Omega_2$ and $\gamma \in \Omega_3$ be $x_1 - u_1$ or x_1 according as β and γ are incident or not; define x_2, u_2, x_3, u_3 similarly. Let A_1, A_2, A_3 be the incidence matrices of the designs.

Consider $A_2 A_3$. This matrix has rows indexed by Ω_3 and columns indexed by Ω_2; the (γ, β) entry is $x_1 - u_1$ or x_1 according as β and γ are incident or not. So

$$A_2 A_3 = x_1 J - u_1 A_1^T. \tag{6}$$

Postmultiplying by J, using (4) and $J^2 = vJ$,

$$k_2 k_3 = x_1 v - u_1 k_1. \tag{7}$$

This and the two similar equations obtained by cyclic permutation of the subscripts prove part (ii) of the theorem. Also,

$$A_2^T (A_2 A_3) = (A_2^T A_2) A_3$$
$$x_1 A_2^T J - u_1 A_2^T A_1^T = n_2 A_3 + \lambda_2 J A_3.$$

By an analogue of (6), $A_1 A_2 = x_3 J - u_3 A_3^T$; so

$$(x_1 k_2 - u_1 x_3) J + u_1 u_3 A_3 = n_2 A_3 + \lambda_2 k_3 J.$$

Since A_3 and J are linearly independent, this implies

$$n_2 = u_1 u_3. \tag{8}$$

This and its two analogues are part (i) of the theorem. (The equation obtained from the coefficients of J in the above is a consequence of (3), (7), (8), and their analogues.)

The next result shows that the definition of a system of linked projective designs can be weakened.

Theorem 2. (a) $\{\Omega_1, \Omega_2, \Omega_3\}$ *(with incidence relations) is a system of linked projective designs if and only if*

(i) $\{\Omega_1, \Omega_2\}$ *and* $\{\Omega_1, \Omega_3\}$, *with their incidence relations, are projective designs;*

(ii) *the number of elements* $\alpha \in \Omega_1$ *incident with both* $\beta \in \Omega_2$ *and* $\gamma \in \Omega_3$ *depends only on whether or not* β *and* γ *are incident;*

(iii) *some element of* Ω_2 *is incident with at least two elements of* Ω_3, *and some element of* Ω_2 *is non-incident with at least two elements of* Ω_3.

(b) *For* $f \geqq 3$, $\{\Omega_1, \dots, \Omega_f\}$ *(with incidence relations) is a system of linked projective designs if and only if* $\{\Omega_1, \Omega_i, \Omega_j\}$ *is such a system for all* i, j *with* $1 < i < j \leqq f$.

Proof. (a) Conditions (i), (ii), and (iii) are clearly necessary. Suppose they hold. Let A_1, A_2, A_3 be the incidence matrices of the incidence relations (directed as in Theorem 1). We have

$$A_i J = J A_i = k_i J,$$
$$A_i A_i^T = A_i^T A_i = n_i I + \lambda_i J, \qquad (i=2,3)$$
$$A_2 A_3 = x_1 J - u_1 A_1^T,$$

for appropriate integers k_i, λ_i, n_i $(i=2,3)$, x_1, u_1. So

$$u_1 A_1 J = u_1 A_1^T J = (x_1 v - k_2 k_3) J,$$
$$u_1^2 A_1 A_1^T = u_1^2 A_1^T A_1 = t J + n_2 n_3 I,$$

for some integer t. If $u_1 = 0$ then $n_2 n_3 = 0$, which is false. So $A_1 J = A_1^T J = k_1 J$ and $A_1 A_1^T = A_1^T A_1 = n_1 I + \lambda_1 J$, for rational (and hence integral) k_1, λ_1, n_1. Condition (iii) gives $1 < k_1 < v - 1$, so $\{\Omega_2, \Omega_3\}$ is a projective design.

Now $A_3(x_1 J - u_1 A_1) = A_3 A_3^T A_2^T$,

$$x_1 k_3 J - u_1 A_3 A_1 = n_3 A_2^T + \lambda_3 k_2 J,$$
$$A_3 A_1 = x_2 J - u_2 A_2^T$$

for rational (and hence integral) x_2, u_2. Similarly

$$A_1 A_2 = x_3 J - u_3 A_3^T \qquad \text{for integral } x_3, u_3.$$

Thus $\{\Omega_1, \Omega_2, \Omega_3\}$ is a system of linked projective designs.

(b) The condition on triples of sets is clearly necessary. Suppose it holds; we must prove that $\{\Omega_i, \Omega_j, \Omega_k\}$ is a system, where $1 \notin \{i, j, k\}$; clearly we may assume $f = 4$. Let the designs be indexed and directed as shown.

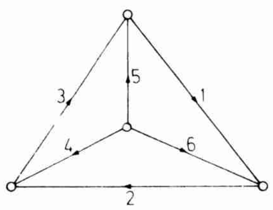


Fig. 2

We have

$$A_5^T A_6 = x_{56} J - u_{56} A_1,$$
$$A_6^T A_4 = x_{64} J - u_{64} A_2,$$
$$A_4^T A_5 = x_{45} J - u_{45} A_3,$$

together with the equations corresponding to (4) and (5) for each of the six designs. So

$$\begin{aligned}(A_5^T A_6)(A_6^T A_4) &= A_5^T(A_6 A_6^T) A_4\\ t J + u_{56} u_{64} A_1 A_2 &= A_5^T(n_6 I + \lambda_6 J) A_4\\ &= A_5^T A_4 (n_6 I + \lambda_6 J)\\ &= (x_{45} J - u_{45} A_3^T)(n_6 I + \lambda_6 J)\\ &= t' J - n_6 u_{45} A_3^T.\end{aligned}$$

So $A_1 A_2 = x_3 J - u_3 A_3^T$ for integral x_3, u_3. The other two equations are verified similarly. We have as a bonus

$$u_{56} u_{64} u_{12} = n_6 u_{45},$$

i..e. $u_{64} u_{12} = u_{16} u_{45}$ and $u_{56} u_{12} = u_{26} u_{45}$.

Examples. (i) If V is a $2m$-dimensional vector space over $GF(q)$, where $q=2^n$ and $n>1$, then the split extension of the translation group of V by the symplectic group $Sp(2m, q)$ has q pairwise inequivalent doubly transitive representations of degree q^{2m} with the same character. This gives a system of linked designs with $f=2^n$, $v=2^{2mn}$; each design has parameters $(2^{2mn}, 2^{2mn-1}-2^{mn-1}, 2^{2mn-2}-2^{mn-1})$. (The case $m=1$ was pointed out to me by Professor D.G. Higman; the general case follows from a result of Pollatsek [6].) The group of permutations of $\Omega_1 \cup \dots \cup \Omega_f$ preserving the incidence relation as a whole is doubly transitive on $\{\Omega_1, \dots, \Omega_f\}$; so all the designs are isomorphic.

(ii) Goethals (oral communication) has shown that a system of linked projective designs with $f=\frac{1}{2}v$, $v=2^{2m}$, can be constructed from the Kerdock codes (see Kerdock [2]) for any $m>1$. The designs have parameters $(2^{2m}, 2^{2m-1}-2^{m-1}, 2^{2m-2}-2^{m-1})$.

(iii) These two families of examples contain designs with the same parameters. To clarify the relation between them in the case $v=16$, take a Steiner system $\mathcal{S}=\mathcal{S}(5, 8, 24)$ on $\{1, \dots, 24\}$ containing the block $B_0=\{1, \dots, 8\}$. Let $\Omega_0=\{9, \dots, 24\}$. For $\{i, j\} \subseteq B_0$, let Ω_{ij} be the set of blocks B such that $B \cap B_0 = \{i, j\}$. (Recall that two blocks of $\mathcal{S}$ have 0, 2, or 4 common points.) A counting argument shows that $|\Omega_{ij}|=16$. The stabiliser (in Aut$(\mathcal{S})$) of B_0 and three points $i, j, k \in B_0$ is doubly transitive on Ω_0. It follows that $\{\Omega_0, \Omega_{ij}\}$ is a projective design with parameters (16, 6, 2), and that $\{\Omega_0, \Omega_{ij}, \Omega_{ik}\}$ is a system of linked projective designs (a block in Ω_{ij} and one in Ω_{ik} are incident if just one point of Ω_0 is incident with both). By Theorem 2(b), $\{\Omega_0, \Omega_{12}, \Omega_{13}, \Omega_{23}\}$ and $\{\Omega_0, \Omega_{12}, \dots, \Omega_{18}\}$ are systems; they turn out to be the case $v=16$ of examples (i) and (ii) respectively. There are, however, other systems with these parameters. We can find a set Ω' containing 8 blocks of Ω_{13} and 8 blocks of Ω_{24} such that $\{\Omega_0, \Omega_{12}, \Omega'\}$ is a system.

Theorem 2 motivates the following definition. Given a natural number v, let Ω be a set of v elements. Changing our point of view, regard a design with point set Ω as a collection of subsets of Ω satisfying appropriate conditions. Define a graph $\mathscr{G}_v$ whose vertices are the pairs of complementary projective designs on the point set Ω, with two vertices joined if the designs can be linked by a suitable incidence relation between their blocks. (If two designs can be linked, then just one pair of complementary relations links them.) By Theorem 2, systems of linked designs on f sets correspond to $(f-1)$-cliques in the graph, and maximal systems to maximal cliques. The symmetric group on Ω has a natural action on the graph; its orbits are in 1-1 correspondence with isomorphism classes of pairs of complementary projective designs with v points.

§ 3. Automorphisms of Doubly Transitive Groups

In [8], Wielandt showed that if G is a doubly transitive permutation group in which the stabiliser of a point is a Hall subgroup (i.e. has coprime order and index), $A(G)$ is the automorphism group of G, and $P(G)$ is the group of permutation automorphisms of G (automorphisms induced by the normaliser of G in the symmetric group), then

(i) $P(G) \leqq O^2(A(G))$;

(ii) if G has prime degree p then $A(G)/G$ is cyclic of order dividing $p-1$.

After deriving a simple consequence of Theorem 1, I use Wielandt's methods to deduce extensions of his results.

Lemma. *Suppose the hypotheses of Theorem 1 hold, and $(v, k_1)=1$. Then $k_1 \neq k_2$ (and similarly $k_1 \neq k_3$).*

Proof. If $(v, k_1)=1$ and $k_1=k_2$, then $n_1=n_2$; so by Theorem 1, $u_1=u_2$, and $n_3=u_2^2$. v divides $k_1k_3+u_2k_2=k_1(k_3+u_2)$, so by hypothesis v divides k_3+u_2. But

$$v(k_3-u_2^2)=v\lambda_3=k_3^2-u_2^2=(k_3+u_2)(k_3-u_2)\neq 0,$$

so k_3-u_2 divides $k_3-u_2^2$. This implies $u_2=0$ or 1, $k_3=v$ or $v-1$, which is false.

Theorem 3. *Suppose G is a doubly transitive permutation group in which the stabiliser of a point is a Hall subgroup. Let $A(G)$ be the auto-*

morphism group of G, and $P(G)$ the group of permutation automorphisms of G. Then $P(G) \lhd A(G)$, and $A(G)/P(G)$ has exponent dividing 2.

Proof. Let $\mathscr{R}$ be a set of representatives of the equivalence classes of transitive permutation representations of G which have the same degree v as the given one. Since $v^2 \nmid |G|$, the stabiliser of a point in one such representation is intransitive in any other, and so all these representations have the same permutation character, and all are doubly transitive. We denote an element of $\mathscr{R}$ by a set Ω_i affording the appropriate representation of G. $A(G)$ acts as a permutation group on $\mathscr{R}$: the group of permutation automorphisms of a particular representation is the stabiliser of the corresponding element of $\mathscr{R}$. To prove that $P(G) \lhd A(G)$, I show that $A(G)$ acts semiregularly on $\mathscr{R}$; to show that the factor group has exponent 1 or 2, I show that the square of any element of $A(G)$ fixes every point of $\mathscr{R}$.

If $\theta \in A(G)$ does not act semiregularly on $\mathscr{R}$, then there are $\Omega_1, \Omega_2, \Omega_3 \in \mathscr{R}$ such that $\Omega_1 \theta = \Omega_1$, $\Omega_2 \theta = \Omega_3$. If $\theta \in A(G)$ and θ^2 does not fix $\Omega_2 \in \mathscr{R}$, then there are $\Omega_1, \Omega_3 \in \mathscr{R}$ such that $\Omega_2 \theta = \Omega_1$, $\Omega_1 \theta = \Omega_3$. In either case, θ carries the unordered pair $\{\Omega_1, \Omega_2\}$ to the unordered pair $\{\Omega_1, \Omega_3\}$. So the corresponding designs have the same parameters v, k, λ (replacing a design by its complement if necessary). By the lemma, $(v, k) > 1$.

But k is the length of an orbit of G_ω, where $\omega \in \Omega_2$, so $k \mid |G_\omega|$. G_ω is a Hall subgroup of G, so $|G_\omega|$ is coprime with $|G : G_\omega| = v$. Thus $(v, k) = 1$. This contradiction shows that the situations above cannot arise, and so the theorem is true.

In fact, a similar argument shows a more general result, bearing the same relation to Theorem 2 of [8] as the above theorem bears to Theorem 3 of that paper:

(i) If G is a doubly transitive permutation group on Ω with $|\Omega| = v$, and θ is an automorphism of G such that, for $\omega \in \Omega$, $(G_\omega)^\theta$ has an orbit of length k with $(v, k) = 1$, then θ^2 is a permutation automorphism.

(ii) If a group G has two doubly transitive representations on Ω_1 and Ω_2 with $|\Omega_1| = |\Omega_2| = v$, such that, for $\omega \in \Omega_2$, G_ω has an orbit of length k in Ω_1 with $(v, k) = 1$, then an automorphism of G is a permutation automorphism of G^{Ω_1} if and only if it is a permutation automorphism of G^{Ω_2}.

Lemma. *If p is an odd prime and v is a power of p, then a projective design on v points cannot have a symmetric incidence matrix with zero diagonal.*

Proof. Suppose there is such a design, with parameters (v, k, λ), where $v = p^a$, $k - \lambda = n$; let A be its incidence matrix, with $A = A^T$ and trace $(A) = 0$. Since $A^2 = nI + \lambda J$, A^2 has eigenvalues $n + \lambda v = k^2$ (with multiplicity 1) and n (multiplicity $v-1$). So A has eigenvalues k (with multiplicity 1), $n^{\frac{1}{2}}$ (multiplicity e) and $-n^{\frac{1}{2}}$ (multiplicity f), where e and f are non-negative integers whose sum is $v-1$. $0=\operatorname{trace}(A)=k+(e-f)\,n^{\frac{1}{2}}$; so $e-f\neq 0$, and hence $n=t^2$ for some integer t which divides k.

Let $(x)_p$ denote the exponent of the highest power of p dividing the integer x. Suppose $(t)_p=b$. Since from (3), $v\,|\,(k^2-t^2)$, and $a>2b$, we have $(k^2)_p=2b$, $(k)_p=b$. By (1), $k(k-1)/\lambda=v-1$, which is not divisible by p, so $(\lambda)_p\geqq b$ (with equality if $b>0$). Now $(k+t)_p(k-t)_p=(v\,\lambda)_p\geqq a+b$, and one at least of $(k+t)_p$ and $(k-t)_p$ is equal to b, since p is odd; so, say, $(k+t)_p\geqq a$, whence $k+t\geqq p^a=v$. Thus $k+t=v$, $k-t=\lambda=k-t^2$, and so $t=0$ or 1, $k=v$ or $v-1$, which is false.

Theorem 4. *Suppose G is a doubly transitive permutation group of degree $v=p^a$, where p is an odd prime, and suppose that a Sylow p-subgroup S of G acts regularly (equivalently, that the stabiliser of a point is a Hall p'-subgroup). Let $A(G)$, $P(G)$ be as in Theorem 2, and $A(S)$, $I(S)$ the automorphism group and inner automorphism group of S. Then*

(i) *the centraliser of S in $A(G)$ is contained in $P(G)$;*

(ii) *$A(G)/G$ is isomorphic to a quotient of an extension of a subgroup of $I(S)$ by a subgroup of $A(S)/I(S)$.*

Proof. Let $N(S)$, $C(S)$ denote the normaliser and centraliser of S in $A(G)$. The centraliser of a regular permutation group S in the symmetric group is a regular permutation group S^* isomorphic to S. (If S is regarded as its own right regular representation, then S^* is the left regular representation of S: Burnside [3], § 136.) Thus $C(S)\cap P(G)\leqq S^*$ and so $C(S)\cap P(G)$ is a p-group. Note that $S\cap S^*=Z(S)$, the centre of S.

(i) Suppose $\theta\in C(S)$ but $\theta\notin P(G)$. By Theorem 3, $\theta^2\in P(G)$; since p is odd, we may assume without loss of generality that $\theta^2=1$. Suppose θ interchanges the given representation (on Ω_1) with a representation on Ω_2; let $\mathscr{D}$ be the design constructed from these two representations. θ induces a polarity of $\mathscr{D}$; so we may index the points and blocks of $\mathscr{D}$ with $1,2,\dots,p^a$ in such a way that θ interchanges points and blocks with the same index, and with this ordering the incidence matrix A of $\mathscr{D}$ is symmetric. Since θ commutes with every element of S, it follows easily that the correspondence between Ω_1 and Ω_2 induced by θ is an S-isomorphism. (Elements of S are represented by the same permutations of the index set on Ω_1 and Ω_2.)

If an automorphism z of $\mathscr{D}$ is represented by permutation matrices Z_1 and Z_2 on Ω_1 and Ω_2, then $Z_1A=AZ_2$. So elements of S are represented by permutation matrices which commute with A. So A is in the integral centraliser ring of (the right regular representation of) S. By

Wielandt [7], Proposition 28.6, this ring is $\mathbb{Z}S^*$. So A is a sum of elements of S^* (regarded as permutation matrices). By replacing $\mathscr{D}$ by its complement if necessary, we may assume that the identity does not occur in this sum. So A has zero diagonal; the lemma is contradicted.

(ii) Now we know $C(S) \leqq P(G)$, so $C(S) \leqq S^*$. Then $SC(S) \lhd N(S)$ and $SC(S)/S \cong C(S)/S \cap C(S) = C(S)/Z(S) \leqq S^*/Z(S) \cong S/Z(S) \cong I(S)$. It is well known that $N(S)/SC(S)$ is isomorphic to a subgroup of $A(S)/I(S)$. By the Frattini lemma, $A(G) = GN(S)$, so $A(G)/G \cong N(S)/N(S) \cap G$. Since $S \leqq N(S) \cap G$, this is a quotient of $N(S)/S$, and the result is proved. If S is abelian, (ii) may be replaced by the simpler statement that $A(G)/G$ is isomorphic to a quotient of a subgroup of $A(S)$.

Corollary. *Let G be a non-abelian finite simple group, and p an odd prime such that G contains a Hall p'-subgroup and a Sylow p-subgroup S which is an abelian B-group.* (See [7], Section 25, for the definition of a B-group.) *Then $A(G)/G$ is isomorphic to a quotient of a subgroup of $A(S)$.*

Proof. Represent G as a permutation group on the cosets of a Hall p'-subgroup. G has a regular abelian subgroup S, so by a result of Wielandt ([7], Theorem 27.4) G is primitive. Since S is a B-group, G is doubly transitive. Now the result follows from Theorem 4. Note that Bercov [1] has shown that if p is an odd prime, an abelian p-group $S = A \times B$, where A is cyclic of order p^a $(a > 1)$ and B has exponent p^b $(b < a)$, is a B-group provided it is not a direct product of n groups of the same order with $n > 1$.

§ 4. Further Numerical Results

Return to the notation of Section 2 and the hypotheses of Theorem 1. We have

$$v\lambda_1 = k_1^2 - u_2 u_3,$$

$$v x_1 = k_2 k_3 + u_1 k_1,$$

and four similar equations. From these we obtain

$$v(k_1 x_1 - \lambda_1 u_1) = k_1 k_2 k_3 + u_1 u_2 u_3, \qquad \text{etc.,}$$

and so $k_1 x_1 - \lambda_1 u_1 = k_2 x_2 - \lambda_2 u_2 = k_3 x_3 - \lambda_3 u_3 = Q$, say, with

$$vQ = k_1 k_2 k_3 + u_1 u_2 u_3.$$

Further calculation shows that

$$v(x_1 x_2 x_3 + \lambda_1 \lambda_2 \lambda_3) = Q(k_1 x_1 + k_2 x_2 + k_3 x_3 - Q).$$

Define the rational number w by

$$v w = k_1 x_1 + k_2 x_2 + k_3 x_3 - Q,$$

$$wQ = x_1 x_2 x_3 + \lambda_1 \lambda_2 \lambda_3.$$

Now we can show that

$$v^2 w = 2k_1 k_2 k_3 + u_1 k_1^2 + u_2 k_2^2 + u_3 k_3^2 - u_1 u_2 u_3,$$

$$v w^2 = 2x_1 x_2 x_3 + \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 - \lambda_1 \lambda_2 \lambda_3,$$

$$w u_1 = x_1^2 - \lambda_2 \lambda_3, \qquad \text{etc.},$$

$$w k_1 = x_2 x_3 + \lambda_1 x_1, \quad \text{etc.}$$

(Assume v is prime. Then v divides $Q(k_1 x_1 + k_2 x_2 + k_3 x_3 - Q)$. If $v \mid Q$, then $k_i x_i \equiv \lambda_i u_i \pmod v$ for $i = 1, 2, 3$, and so

$$k_1 k_2 k_3 x_1 x_2 x_3 \equiv \lambda_1 \lambda_2 \lambda_3 u_1 u_2 u_3 \pmod v.$$

But $k_1 k_2 k_3 \equiv -u_1 u_2 u_3 \not\equiv 0 \pmod v$; so $x_1 x_2 x_3 \equiv -\lambda_1 \lambda_2 \lambda_3 \pmod v$, and so $v^2 \mid Q(k_1 x_1 + k_2 x_2 + k_3 x_3 - Q)$. We cannot have $v^2 \mid Q$, and so

$$v \mid (k_1 x_1 + k_2 x_2 + k_3 x_3 - Q)$$

in any case. So, from its definition, w is an integer. w may be an integer in other situations as well; in all the examples in Section 2 except those with $v = 16$, w is an integer.)

The six further equations

$$k_1 x_3 - k_2 \lambda_1 - u_3 x_2 = 0, \qquad \text{etc.},$$

may be deduced from those given.

On interchanging v with w, k_i with x_i, and λ_i with u_i, all the equations in this section so far are transformed into valid equations. (Q is fixed by this interchange.) I have been unable to exploit or explain this remarkable "duality". Note, however, that the duals of

$$k_1(x_1 - u_1) = k_2(x_2 - u_2) = k_3(x_3 - u_3)$$

and

$$u_1(k_1 - \lambda_1) = u_2(k_2 - \lambda_2) = u_3(k_3 - \lambda_3)$$

are not valid, though the equations themselves are.

If some of the designs in the system are replaced by their complements, another system is obtained. In the table below, σ_i is the substitution on the parameters when the i^{th} design is replaced by its complement. $\langle \sigma_1, \sigma_2, \sigma_3 \rangle$ is an elementary abelian group of order 8; its elements which are not listed in Table 1 may be obtained from its automorphism of order 3 which permutes the subscripts 1, 2, 3 cyclically.

Table 1

1	σ_1	$\sigma_1\sigma_2$	$\sigma_1\sigma_2\sigma_3$
A_1	$J-A_1$	$J-A_1$	$J-A_1$
A_2	A_2	$J-A_2$	$J-A_2$
A_3	A_3	A_3	$J-A_3$
k_1	$v-k_1$	$v-k_1$	$v-k_1$
k_2	k_2	$v-k_2$	$v-k_2$
k_3	k_3	k_3	$v-k_3$
λ_1	$v-2k_1+\lambda_1$	$v-2k_1+\lambda_1$	$v-2k_1+\lambda_1$
λ_2	λ_2	$v-2k_2+\lambda_2$	$v-2k_2+\lambda_2$
λ_3	λ_3	λ_3	$v-2k_3+\lambda_3$
u_1	$-u_1$	u_1	$-u_1$
u_2	$-u_2$	u_2	$-u_2$
u_3	$-u_3$	u_3	$-u_3$
x_1	x_1-u_1	$k_3-x_1+u_1$	$v-(k_2+k_3)+x_1-u_1$
x_2	k_3-x_2	$k_3-x_2+u_2$	$v-(k_3+k_1)+x_2-u_2$
x_3	k_2-x_3	$v-(k_1+k_2)+x_3$	$v-(k_1+k_2)+x_3-u_3$
Q	k_2k_3-Q	$(v-k_1-k_2)k_3+Q$	$v^2-(k_1+k_2+k_3)v+(k_1k_2+k_2k_3+k_3k_1)-Q$
w	$2x_1-u_1-w$	$2k_3-2(x_1+x_2)$ $+u_1+u_2+w$	$2v-2(k_1+k_2+k_3)+2(x_1+x_2+x_3)$ $-(u_1+u_2+u_3)-w$

v, n_1, n_2, n_3 are unaltered by all the substitutions in $\langle\sigma_1, \sigma_2, \sigma_3\rangle$.

§ 5. Groups of Prime Degree

Wielandt has conjectured that a transitive permutation group of prime degree p has at most two conjugacy classes of subgroups of index p (equivalently, at most inequivalent transitive representations of degree p). It was hoped that the methods of this paper would provide a proof of the conjecture; they have not done so, but have provided evidence in its support. (In fact, the editor of this journal informs me that it was precisely the content of Theorem 1 and numerical evidence of the type to be discussed below which suggested this conjecture to Wielandt in 1955.)

If G is a transitive group of prime degree p having three pairwise inequivalent representations of degree p, then G is insoluble, and, by a theorem of Burnside ([3], § 251), doubly transitive. The first lemma of Section 3 shows that the three designs constructed from these representations have pairwise distinct parameters, these parameters being related as in Theorem 1. These relations (with $v=p$, a prime) appear to be difficult to satisfy. There are no sets of numbers satisfying them with $p \leqq 2000000$ (this was shown by computation by Neumann). Each of the designs has a polarity ([4], 1.2.13) which induces an automorphism of its full automorphism group; it follows from Theorem 4 (or from Wielandt's result mentioned in Section 3) that G can be the full automorphism group of at most one of the three designs. Thus the existence of G entails the existence of several as-yet-unknown permutation groups.

The procedure for Neumann's computations was as follows. First, for the prime p, all solutions of $p-1 \mid k(k-1)$ were obtained, by finding all factorisations $p-1=rs$ with r and s coprime, and then solving $k \equiv 0 \pmod r$, $k \equiv 1 \pmod s$, which have a unique solution mod $p-1$. The trivial solutions $k=1$ and $k=p-1$ are not required, and of each pair $(k, p-k)$, only one is required (for example, we may specify $r>s$). So if l is the number of distinct prime divisors of $p-1$, we obtain $2^{l-1}-1$ values of k. For each k, the value of $n=k-\lambda$ was found. The computer then performed a weak test: if one value of n has a prime divisor which divides no other value, that value is deleted. (By part (i) of Theorem 1, it cannot occur in a system of linked designs.) This process was continued until no further deletions occurred, and if three or more values of n remained, they were printed out. All primes less than two million were examined, and no sets of numbers satisfying Theorem 1 were found. So there is no system of linked projective designs with v a prime less than 2000000.

Using only the fact that the three designs have different parameters, and not the full content of Theorem 1, the following weak result can be proved. Let G be a transitive permutation group of prime degree p on Ω. Let S be a Sylow p-subgroup of G and N its normaliser. Let l be the number of distinct prime factors of $p-1$, m the number of distinct prime factors of $|N:S|$, h the number of non-principal irreducible constituents of the permutation character restricted to G_ω, for $\omega \in \Omega$, t the minimum number of fixed points of an involution in G, and f the number of conjugacy classes of subgroups of index p. Then $f \leqq 2^{l-m}$ and $f \leqq 2^{h-1}$. If $l \leqq m+1$ (in particular, if $l \leqq 2$), or if G_ω has rank 3 on $\Omega-\{\omega\}$, or if $t \leqq 3$, then $f \leqq 2$. (The last statement follows from the fact that if a projective design with parameters (v, k, λ) has an automorphism of order 2 fixing t points with $t \geqq 1$, then $t \geqq 1+(v-1)/k$; if k and λ are even then $t \geqq 1+(v-1)/(k-1)$.) Wielandt has further shown that $l \leqq 3$ implies $f \leqq 2$. The remaining results in this section are also due to Wielandt.

Proposition. *Let G be a doubly transitive permutation group of prime degree p. Suppose two subgroups H_1 and H_2 of index p in G have orbit lengths $k_1, p-k_1$, and k_2, $p-k_2$ respectively. Then H_1 and H_2 are conjugate if and only if $k_1(p-k_1)/k_2(p-k_2)$ is the square of a rational number (in which case it is* 1).

This is a consequence of Theorem 1 and the observation that in a projective design with parameters (v, k, λ) with v prime, $k-\lambda$ is not a square. Alternatively it follows from part (i) of the next result.

Proposition. *Let n_i be non-negative integers of the form $k_i(p-k_i)/(p-1)$, with p prime. Then*

(i) *if n_1/n_2 is a square then $n_1 = n_2$;*

(ii) *if $n_1 n_2 = n_3 n_4$ then $n_1 = n_3$ or $n_1 = n_4$.*

Corollary. *If a permutation group of prime degree p has four pairwise inequivalent representations of degree p, and the four designs shown in the diagram below satisfy $n_1 = n_3$, then $n_2 = n_4$.*

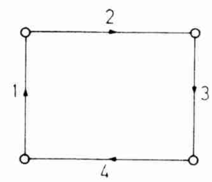


Fig. 3

References

1. Bercov, R.R.: The double transitivity of a class of permutation groups. Canadian J. Math. **17**, 480–493 (1965).
2. Kerdock, A.M.: A class of low-rate nonlinear binary codes. Inform. and Control **20**, 182–187 (1972).
3. Burnside, W.: Theory of groups of finite order. New York: Dover Publications (reprint) 1955.
4. Dembowski, P.: Finite geometries. Berlin-Heidelberg-New York: Springer 1968.
5. Itô, N.: Über die Gruppen $PSL_n(q)$, die eine Untergruppe von Primzahlindex enthalten. Acta Sci. Math. (Szeged) **21**, 206–217 (1960).
6. Pollatsek, H.: First cohomology groups of some linear groups over fields of characteristic two. Illinois J. Math. **15**, 393–417 (1971).
7. Wielandt, H.: Finite permutation groups. New York-London: Academic Press 1964.
8. Wielandt, H.: On automorphisms of doubly-transitive permutation groups. Proc. Int. Conf. Theory of Groups (ed. L.G. Kovacs and B.H. Neumann), 389–393. New York: Gordon and Breach 1967.

Dr. Peter J. Cameron
Merton College
Oxford OX 1 4 JD
England

(Received January 10, 1972)

Math. Z. 128, 15–22 (1972)

Verträglichkeit von konvergenztreuen Matrixverfahren

Johann Boos

Ausgehend vom Satz von Mazur-Orlicz (vgl. [4, Theorem 2]) wird in der folgenden Arbeit die Frage untersucht, unter welchen Voraussetzungen zwei auf den konvergenten Folgen verträgliche Matrixverfahren A und B bezüglich bestimmter Teilmengen der Wirkfelder von A und B verträglich sind.

Bezeichnen wir mit c_A bzw. c_B die Wirkfelder von A bzw. B, so heißen *A und B auf $M \subseteq c_A \cap c_B$ verträglich*, wenn $\lim_A x = \lim_B x$ für alle $x \in M$ gilt. Mit c bzw. m bezeichnen wir die konvergenten bzw. beschränkten Folgen. Weiter verwenden wir für ein Matrixverfahren $A = (a_{nk})$ folgende Bezeichnungen (vgl. [6, S. 330]):

$$a_k := \lim_{n \to \infty} a_{nk} \qquad (k = 1, 2, \ldots),$$

$$|A| \cap c_A := \left\{ x \in c_A : \sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} |a_{nk} x_k| < \infty \right\},$$

$$B_A := \left\{ x \in c_A : \sup_{m, n \in \mathbb{N}} \left| \sum_{k=1}^{m} a_{nk} x_k \right| < \infty \right\},$$

$$I_A := \left\{ x \in c_A : \sum_{k=1}^{\infty} a_k x_k \text{ existiert} \right\};$$

$$F_A := \left\{ x \in c_A : \sum_{k=1}^{\infty} f(e^k) x_k \text{ existiert für alle } f \in c'_A \right\},$$

wobei c'_A der Dualraum von c_A bezüglich der eindeutig bestimmten FK-Topologie von c_A ist (vgl. [7, S. 38ff.]).

Nach Wilansky in [6, 4.2 und 4.3] gilt $m \cap c_A \subseteq F_A = B_A \cap I_A$ für ein konvergenztreues Matrixverfahren A. Da

$$\sum_{k=1}^{\infty} |a_k x_k| < \infty \qquad \text{für alle } x \in |A| \cap c_A \tag{1}$$

gilt, erhält man

$$m \cap c_A \subseteq |A| \cap c_A \subseteq F_A = B_A \cap I_A. \tag{2}$$

Ist $A=(a_{nk})$ ein konvergenztreues Matrixverfahren, so heißt A *konullär*, falls

$$\chi(A) := \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk} - \sum_{k=1}^{\infty} a_k = 0$$

und *koregulär*, falls $\chi(A) \neq 0$ gilt.

Zunächst betrachten wir nur koreguläre Verfahren. Mazur und Orlicz haben in [4, Theorem 2] unabhängig von Brudno ([1, Theorem 8]) die gestellte Frage für koreguläre Verfahren auf dem beschränkten Wirkfeld durch folgenden Satz beantwortet:

Satz 1. *Seien A und B auf den konvergenten Folgen verträgliche Matrixverfahren, und sei A koregulär, dann folgt aus $m \cap c_A \subseteq c_B$ die Verträglichkeit von A und B auf $m \cap c_A$.*

Ersetzt man in Satz 1 das beschränkte Wirkfeld $m \cap c_A$ durch das absolut-A-beschränkte Wirkfeld $|A| \cap c_A$ und koregulär durch permanent, so erhält man den von Volkov in [5, Satz 1] mit der Beweismethode von Brudno bewiesenen Satz; wegen (1) erhält man mit der Beweismethode von Volkov den entsprechenden Satz für koreguläre Verfahren A auf $|A| \cap c_A$. Nun zeigen wir, daß der entsprechende Satz für koreguläre Verfahren A auf F_A gilt.

Satz 2. *Seien A und B auf den konvergenten Folgen verträgliche Matrixverfahren, und sei A koregulär, dann folgt aus $F_A \subseteq c_B$ die Verträglichkeit von A und B auf F_A.*

Beweisen werden wir den zu Satz 2 äquivalenten Satz:

Satz 3. *Seien A und B auf den konvergenten Folgen verträgliche Matrixverfahren, und sei A koregulär; gibt es eine Folge $x \in F_A \cap c_B$ mit $\lim_A x \neq \lim_B x$, dann gibt es eine Folge $y \in F_A$ mit $y \notin c_B$.*

Zum Beweis von Satz 3 verwenden wir folgendes Lemma, welches in dieser Form aus [8, Satz 3.2] folgt; Satz 3.2 in [8] liefert jedoch keine spezielle Folge $\varepsilon=(\varepsilon_k)$ mit den im folgenden Lemma geforderten Eigenschaften. Wählt man $x \in F_A \cap c_B$ wie in Satz 3, so besitzt $y := (x_k \varepsilon_k)$ die in Satz 3 verlangten Eigenschaften.

Lemma. *Seien $C=(c_{nk})$ und $D=(d_{nk})$ Matrixverfahren mit den folgenden Eigenschaften:*

(Zr) $$\sup_{m,n\in\mathbb{N}} \left| \sum_{k=1}^{m} c_{nk} \right| < \infty,$$

(Sp$_0$) $$c_k := \lim_{n\to\infty} c_{nk} = 0 = \lim_{n\to\infty} d_{nk} =: d_k \quad (k=1,2,\ldots),$$

(Zs) $$\lim_{n\to\infty} \sum_{k=1}^{\infty} c_{nk} = 0 \quad \textit{und} \quad \lim_{n\to\infty} \sum_{k=1}^{\infty} d_{nk} = 1,$$

wobei die einzelnen Grenzwerte bzw. Reihen existieren. Unter diesen Voraussetzungen gibt es ein $\varepsilon=(\varepsilon_k)$ *mit* $\varepsilon\in c_C$ *und* $\varepsilon\notin c_D$.

Beweis. Die Matrixverfahren $G=(g_{nl})$ und $H=(h_{nl})$ seien definiert durch

$$g_{nl}:=\sum_{k=l}^{\infty} c_{nk} \quad \text{und} \quad h_{nl}:=\sum_{k=l}^{\infty} d_{nk} \qquad (n, l=1, 2, \ldots).$$

Da C und D die Bedingungen (Zr), (Sp_0) und (Zs) erfüllen, verifiziert man leicht, daß G und H folgende Eigenschaften besitzen:

$$|g_{nl}|<M \qquad (n, l=1, 2, \ldots), \tag{3}$$

$$\lim_{l\to\infty} g_{nl}=0=\lim_{l\to\infty} h_{nl} \qquad (n=1, 2, \ldots), \tag{4}$$

$$\lim_{n\to\infty} g_{nl}=0 \quad \text{und} \quad \lim_{n\to\infty} h_{nl}=1 \qquad (l=1, 2, \ldots). \tag{5}$$

Weiter sei (α_j) mit

$$\alpha_j>0 \quad (j=1, 2, \ldots), \qquad \sum_{j=1}^{\infty} \alpha_j<\infty \quad \text{und} \quad \lim_{j\to\infty} j\alpha_j=0 \tag{6}$$

vorgegeben; nach (3)–(5) kann α_1 derart gewählt werden, daß

$$\left.\begin{aligned} &|g_{nl}|<\alpha_1 && (n, l=1, 2, \ldots), \\ &|h_{nl}|<\alpha_1 && \text{für } n=1 \text{ und } l=1, 2, \ldots \\ \text{und}\quad &|h_{nl}-1|<\alpha_1 && \text{für } l=1 \text{ und } n=1, 2, \ldots \end{aligned}\right\} \tag{7}$$

gilt. Wählt man $q_1=1=n_1$, so können nach (4) und (5) eine streng isotone Indexfolge (q_j) und eine streng isotone Hilfsindexfolge (n_j) induktiv derart bestimmt werden, daß für alle $j=2, 3, \ldots$

$$|g_{nl}|<\alpha_j, \qquad |h_{nl}|<\alpha_j \qquad \text{für alle } l\geqq q_j \text{ und } n\leqq n_{j-1} \tag{8}$$

und

$$|g_{nl}|<\alpha_j, \qquad |h_{nl}-1|<\alpha_j \qquad \text{für alle } l\leqq q_j \text{ und } n\geqq n_j \text{ gilt.} \tag{9}$$

Im folgenden konstruieren wir analog zu [5, Satz 1] eine Folge $\varepsilon=(\varepsilon_k)$ mit $\varepsilon\in c_C$ und $\varepsilon\notin c_D$.

Die Glieder der Folge (F_j) seien definiert durch die Partialsummen der Reihe

$$1-1+\frac{1}{2}+\frac{1}{2}-\frac{1}{2}-\frac{1}{2}+\cdots+\frac{1}{n}+\cdots+\frac{1}{n}-\frac{1}{n}-\cdots-\frac{1}{n}+\cdots,$$

wobei das Glied $\frac{1}{n}$ n-mal mit positivem und n-mal mit negativem Vorzeichen auftritt. Laut Definition besitzt (F_j) folgende Eigenschaften:

$$0 \leqq F_j \leqq 1 \quad (j=1,2,\ldots), \tag{10}$$

$$F_j=0 \quad \text{für } j=1,2,\ldots \text{ mit } j=q(q+1) \text{ und } q=1,2,\ldots, \tag{11}$$

$$F_j=1 \quad \text{für } j=1,2,\ldots \text{ mit } j=q^2 \text{ und } q=1,2,\ldots. \tag{12}$$

Weiter definieren wir die Folge $\varepsilon=(\varepsilon_k)$ und die Hilfsfolge $u=(u_k)$ durch

$$\varepsilon_k := F_j \quad \text{für } k=1,2,\ldots \text{ mit } q_j \leqq k < q_{j+1} \quad (j=1,2,\ldots) \tag{13}$$

und

$$u_k := \varepsilon_k - \varepsilon_{k-1} \quad \text{mit } \varepsilon_0=0 \quad (k=1,2,\ldots). \tag{14}$$

Insbesondere gilt

$$u_k=0 \quad \text{für } k \neq q_j \ (j=1,2,\ldots) \quad \text{und} \quad \lim_{k\to\infty} u_k=0. \tag{15}$$

a) Zunächst zeigen wir $u \in c_G$ und $\lim_G u=0$. Aus (15), (8), (9) und (3) folgt für $j=2,3,\ldots$ und $n=1,2,\ldots$ mit $n_{j-1}<n\leqq n_j$

$$\begin{aligned}\left|\sum_{l=1}^{\infty} g_{nl}\,u_l\right| &\leqq \sum_{l=1}^{\infty} |g_{nl}\,u_l| = \sum_{v=1}^{\infty} |g_{nq_v}\,u_{q_v}| \\ &= \sum_{v=1}^{j-1} |g_{nq_v}\,u_{q_v}| + |g_{nq_j}\,u_{q_j}| + \sum_{v=j+1}^{\infty} |g_{nq_v}\,u_{q_v}| \\ &\leqq \alpha_{j-1}(j-1)\sup_{l\in\mathbb{N}} |u_l| + M\,|u_{q_j}| + \sup_{l\in\mathbb{N}} |u_l| \sum_{v=j+1}^{\infty} \alpha_v < \infty.\end{aligned}$$

Wegen (6) und (15) konvergieren die letzten drei Ausdrücke für $j\to\infty$ gegen Null; da $\sum_{l=1}^{\infty} g_{1l}\,u_l$ existiert (vgl. (6) und (15)), gilt $u\in c_G$ und $\lim_G u=0$.

b) Aus a) folgt $\varepsilon \in c_C$ und $\lim_C \varepsilon = \lim_G u = 0$, da für $r,n=1,2,\ldots$ aus (4) und $0\leqq \varepsilon_r \leqq 1$ (vgl. (10) und (13))

$$\sum_{k=1}^{r} c_{nk}\,\varepsilon_k = \sum_{l=1}^{r} g_{nl}\,u_l - g_{n,r+1}\,\varepsilon_r \tag{16}$$

und somit

$$\sum_{k=1}^{\infty} c_{nk}\,\varepsilon_k = \sum_{l=1}^{\infty} g_{nl}\,u_l \quad \text{für } n=1,2,\ldots \tag{17}$$

folgt.

c) Es bleibt $\varepsilon \notin c_D$ zu zeigen. Da für $n = 1, 2, \ldots$ die Konvergenz der Reihen $\sum_{l=1}^{\infty} h_{nl} u_l$ wegen (6) und

$$\sum_{l=q_j+1}^{\infty} |h_{nl} u_l| = \sum_{\nu=j+1}^{\infty} |h_{nq_\nu} u_{q_\nu}| \leqq \sup_{l \in \mathbb{N}} |u_l| \sum_{\nu=j+1}^{\infty} \alpha_\nu < \infty \tag{18}$$

für $n = 1, 2, \ldots$ mit $n_{j-1} < n \leqq n_j$ $(j = 2, 3, \ldots)$ gesichert ist, genügt es nach (16) und (17) (angewendet auf D und H)

$$\sum_{l=1}^{\infty} h_{n_j l} u_l \to \begin{cases} 0 & \text{für } j = q(q+1) \to \infty \\ 1 & \text{für } j = q^2 \to \infty \end{cases} \tag{19}$$

zu zeigen. Für $j = 1, 2, \ldots$ gilt

$$\begin{aligned} \left| \sum_{l=1}^{\infty} h_{n_j l} u_l \right| &= \left| \sum_{\nu=1}^{\infty} h_{n_j q_\nu} u_{q_\nu} \right| \\ &\leqq \left| \sum_{\nu=1}^{j} h_{n_j q_\nu} u_{q_\nu} \right| + \sum_{\nu=j+1}^{\infty} |h_{n_j q_\nu} u_{q_\nu}|; \end{aligned} \tag{20}$$

der letzte Term strebt wegen (6), (15) und (18) für $j \to \infty$ gegen Null. Aus (11)–(15) folgt

$$\sum_{\nu=1}^{j} u_{q_\nu} \to \begin{cases} 0 & \text{für } j = q(q+1) \text{ und } q = 1, 2, \ldots \\ 1 & \text{für } j = q^2 \text{ und } q = 1, 2, \ldots. \end{cases} \tag{21}$$

Berücksichtigt man die Abschätzung (vgl. (9))

$$\begin{aligned} \left| \sum_{\nu=1}^{j} (h_{n_j q_\nu} - 1) u_{q_\nu} \right| &\leqq \sup_{l \in \mathbb{N}} |u_l| \sum_{\nu=1}^{j} |h_{n_j q_\nu} - 1| \\ &\leqq j \alpha_j \sup_{l \in \mathbb{N}} |u_l| \end{aligned}$$

und $\lim_{j \to \infty} j \alpha_j = 0$, so erhält man aus (21) und

$$\sum_{\nu=1}^{j} h_{n_j q_\nu} u_{q_\nu} = \sum_{\nu=1}^{j} (h_{n_j q_\nu} u_{q_\nu} - 1) + \sum_{\nu=1}^{j} u_{q_\nu}$$

die Beziehung

$$\sum_{\nu=1}^{j} h_{n_j q_\nu} u_{q_\nu} \to \begin{cases} 0 & \text{für } j = q(q+1) \to \infty \\ 1 & \text{für } j = q^2 \to \infty. \end{cases}$$

Damit folgt (19) aus (20) und somit $\varepsilon \notin c_D$.

Insgesamt ist das Lemma bewiesen.

Bemerkung 1. Ist (g_l) eine beliebige Nullfolge, so kann im Beweis des Lemmas nach geeigneter Wahl von α_1 die Indexfolge (q_j) derart gewählt werden, daß $|g_l|<\alpha_j$ für alle $l \geqq q_j$ und $j=1, 2, \ldots$ zusätzlich zu (8) gilt.

Beweis von Satz 3. Sei $x \in F_A \cap c_B$ und $\lim_A x \neq \lim_B x$, und seien die Matrixverfahren $A^*=(a^*_{nk})$ bzw. $B^*=(b^*_{nk})$ definiert durch

$$a^*_{nk} := a_{nk} - a_k \quad \text{bzw.} \quad b^*_{nk} := b_{nk} - b_k \qquad (n, k = 1, 2, \ldots).$$

Wegen der Verträglichkeit von A und B auf c gilt $a_k = b_k$ $(k=1, 2, \ldots)$, womit aus $x \in F_A = B_A \cap I_A$ die Existenz von $\sum_{k=1}^{\infty} a_k x_k = \sum_{k=1}^{\infty} b_k x_k$ und damit $x \in F_{A^*}$, $x \in c_{B^*}$ und $\lim_{A^*} x \neq \lim_{B^*} x$ folgt.

Ohne Beschränkung der Allgemeinheit können wir

$$\lim_{A^*} x = 0 \quad \text{und} \quad \lim_{B^*} x = 1 \tag{22}$$

annehmen. Werden die Matrixverfahren $C=(c_{nk})$ bzw. $D=(d_{nk})$ durch

$$c_{nk} := a^*_{nk} x_k \quad \text{bzw.} \quad d_{nk} := b^*_{nk} x_k \qquad (n, k = 1, 2, \ldots)$$

definiert, so erfüllen C und D die Voraussetzungen des Lemmas. Damit gibt es nach dem Lemma eine Folge mit $\varepsilon \in c_C$ und $\varepsilon \notin c_D$; wählt man die im Beweis des Lemmas konstruierte Folge ε und setzt $y := (x_k \varepsilon_k)$, so erhält man

$$y \in c_{A^*} \quad \text{und} \quad y \notin c_{B^*}. \tag{23}$$

Im folgenden zeigen wir $y \in F_A$ und $y \notin c_B$.

Da nach (2) insbesondere $x \in I_A$ gilt, ist die Folge (g_l) mit

$$g_l := \sum_{k=l}^{\infty} a_k x_k \qquad (l = 1, 2, \ldots)$$

eine Nullfolge. Damit kann nach Bemerkung 1 im Beweis des Lemmas die Indexfolge (q_j) derart konstruiert werden, daß

$$|g_l| < \alpha_j \qquad \text{für alle } l \geqq q_j \ (j = 1, 2, \ldots) \text{ gilt.} \tag{24}$$

Ist u durch (14) definiert, so folgt aus (15), (24) und (6) die absolute Konvergenz von $\sum_{l=1}^{\infty} g_l u_l$. Wie im Beweisteil c) des Lemmas zeigt man

$$\sum_{k=1}^{\infty} b_k y_k = \sum_{k=1}^{\infty} a_k y_k = \sum_{l=1}^{\infty} g_l u_l,$$

was zusammen mit (23) die Beziehungen $y \in I_A$ und $y \notin c_B$ impliziert. Um $y \in F_A$ zu erhalten, genügt es damit $y \in B_A$ zu zeigen (vgl. (2)); hierzu genügt es offensichtlich $y \in B_{A^*}$ zu zeigen.

Da C und D die Voraussetzungen des Lemmas erfüllen, gilt nach (16) mit den Bezeichnungen aus dem Beweis des Lemmas

$$\left|\sum_{k=1}^{r} a_{nk} y_k\right| = \left|\sum_{k=1}^{r} c_{nk} \varepsilon_k\right| = \left|\sum_{l=1}^{r} g_{nl} u_l - g_{n,r+1} \varepsilon_r\right|$$
$$\leqq \sup_{n\in\mathbb{N}} \sum_{l=1}^{\infty} |g_{nl} u_l| + \sup_{n,r\in\mathbb{N}} |g_{n,r+1} \varepsilon_r| \leqq N < \infty,$$

wobei N unabhängig von $n, r = 1, 2, \ldots$ ist; die letzte Abschätzung folgt aus (3), (10), (13) und der Abschätzung im Beweisteil a) des Lemmas. Damit ist $y \in B_{A^*}$ und insgesamt $y \in F_A$ und $y \notin c_B$ gezeigt.

Bemerkung 2. Wählt man in Satz 3 $x \in m \cap c_A \cap c_B$ bzw. $x \in |A| \cap c_A \cap c_B$, so kann mit obiger Beweiskonstruktion entsprechend $y \notin c_B$ und $y \in m \cap c_A$ bzw. $y \in |A| \cap c_A$ gezeigt werden, d.h. Satz 1 bzw. der entsprechende Satz (vgl. [5, Satz 1]) für das absolut-A-beschränkte Wirkfeld $|A| \cap c_A$ bewiesen werden.

Die Sätze 1 und 2 und Satz 1 in [5] gelten für koreguläre Matrixverfahren; im folgenden untersuchen wir, unter welchen Voraussetzungen die genannten Sätze für konulläre Verfahren gelten.

In [2, Theorem 3] (vgl. auch [3, Theorem 4]) wurde gezeigt, daß die Bedingung $\lim_A x = \sum_{k=1}^{\infty} a_k x_k$ *für alle* $x \in m \cap c_A$ notwendig und hinreichend für die Gültigkeit von Satz 1 für konulläre Verfahren ist. Wir zeigen nun, daß die entsprechende Bedingung notwendig und hinreichend für die Gültigkeit von Satz 2 für konulläre Verfahren ist; Bemerkung 3 gibt die entsprechende Antwort für [5, Satz 1].

Satz 4. *Sei A ein konulläres Matrixverfahren, dann sind folgende Bedingungen äquivalent:*

a) *Ist B mit A auf den konvergenten Folgen verträglich, dann folgt aus $F_A \subseteq c_B$ die Verträglichkeit von A und B auf F_A;*

b) $\lim_A x = \sum_{k=1}^{\infty} a_k x_k$ *für alle* $x \in F_A$.

Beweis. Es gelte a), und $B = (b_{nk})$ sei definiert durch $b_{nk} := a_k$ für alle $k \leqq n$ und $b_{nk} := 0$ für alle $k > n$ $(n = 1, 2, \ldots)$; wegen $c_B \supseteq l_A \supseteq F_A$ (vgl. (2)), $\chi(A) = 0$ und $\lim_A x = \sum_{k=1}^{\infty} a_k x_k = \lim_B x$ für alle $x \in c$ sind die Verfahren A und B auf F_A verträglich, was b) impliziert.

Sei umgekehrt A ein konulläres Matrixverfahren mit $\lim_A x = \sum_{k=1}^{\infty} a_k x_k$ für alle $x \in F_A$; weiter sei B mit A auf den konvergenten Folgen verträglich, und es gebe ein $x \in F_A \cap c_B$ mit $\lim_A x \neq \lim_B x$. Um a) zu zeigen, bleibt die Existenz einer Folge $y \in F_A$ mit $y \notin c_B$ nachzuweisen. Definiert man A^* und B^* wie im Beweis von Satz 3, so erhält man wie dort $x \in F_{A^*} \cap c_{B^*}$ und $\lim_{A^*} x \neq \lim_{B^*} x$. Da nach Voraussetzung $\lim_{A^*} x = 0$ gilt, kann ohne Beschränkung der Allgemeinheit $\lim_{B^*} x = 1$ angenommen werden. Der restliche Beweis verläuft wie der Beweis von Satz 3 ab der Stelle (22).

Bemerkung 3. Analog zum Beweis von Satz 4 kann der Satz 4 entsprechende Satz für das absolut-A-beschränkte Wirkfeld $|A| \cap c_A$ bewiesen werden (vgl. Bemerkung 2).

Analog zu den Untersuchungen von Chang u.a. in [2] können Verträglichkeits- und Perfektheitseigenschaften bezüglich dem absolut-A-beschränkten Wirkfeld $|A| \cap c_A$ bzw. bezüglich F_A untersucht werden; dadurch können weitere ersetzbare bzw. nicht-ersetzbare konulläre Matrixverfahren charakterisiert werden.

Literatur

1. Brudno, A. L.: Summation of bounded sequences by matrices. Mat. Sbornik, n. Ser. **16**(58), 191–247 (1945).
2. Chang, S. C., Macphail, M. S., Snyder, A. K., Wilansky, A.: Consistency and replaceability for conull matrices. Math. Z. **105**, 208–212 (1968).
3. Jurimäe, E.: Topological properties of conull matrices. Tartu Riikl. Ül. Toimetised Vih. **177**, 43–61 (1965).
4. Mazur, S., Orlicz, W.: On linear methods of summability. Studia Math. **14**, 129–160 (1955).
5. Volkov, I. I.: Über die Verträglichkeit zweier Summationsmethoden. Mat. Zametki **1**, 283–290 (1967).
6. Wilansky, A.: Distinguished subsets and summability invariants. J. Analyse Math. **12**, 327–350 (1964).
7. Zeller, K.: Allgemeine Eigenschaften von Limitierungsverfahren. Math. Z. **53**, 463–487 (1951).
8. Zeller, K.: Faktorfolgen bei Limitierungsverfahren. Math. Z. **56**, 134–151 (1952).

Dr. Johann Boos
Mathematisches Institut der Universität
D-7400 Tübingen, Brunnenstr. 27
Deutschland

(Eingegangen am 23. März 1972)

Math. Z. 128, 23 – 33 (1972)

An Area Method for Systems of Univalent Functions Whose Ranges do not Overlap

Duane W. DeTemple

Introduction

Let $\{w_0, w_1, \ldots, w_M\}$ be a given set of distinct points in the extended complex plane. A system of functions $\{F_0, F_1, \ldots, F_M\}$ will be said to belong to the class $\mathscr{M}(w_0, w_1, \ldots, w_M)$ if: (i) each $F_j(z)$ is meromorphic and univalent on the unit disc $E=\{|z|<1\}$; (ii) the ranges $D_j=F_j(E)$ are pairwise disjoint; (iii) $F_j(0)=w_j$. If the functions are required to be regular analytic, the class is denoted by $\mathscr{M}_\infty(w_0, \ldots, w_M)$. If the functions are required to be bounded (that is, each $|F_j(z)|<1$) the subclass is denoted by $\mathscr{M}_1(w_0, \ldots, w_M)$.

In this note, inequalities of the Grunsky-Nehari type are developed for these classes. Statements of the results are found in Section 1, with their proofs following in Section 2. In Section 3, some special cases of the inequalities are discussed, with particular attention given to the application of the inequalities to other related classes of univalent functions.

1. The Main Inequalities

Theorem 1. *Let* $\{F_0, \ldots, F_M\} \in \mathscr{M}_1(w_0, \ldots, w_M)$, *and let the coefficients* K^{jk}_{mn} *and* L^{jk}_{mn} $(m, n=0, 1, \ldots; j, k=0, 1, \ldots, M)$ *be defined by*

$$\log\frac{F_j(z)-F_j(\zeta)}{z-\zeta}=\sum_{m,n=0}^{\infty} K^{jj}_{mn} z^m \zeta^n$$

$$\log[F_j(z)-F_k(\zeta)]=\sum_{m,n=0}^{\infty} K^{jk}_{mn} z^m \zeta^n \quad (j \neq k) \tag{1}$$

$$-\log[1-F_j(z)\overline{F_k(\zeta)}]=\sum_{m,n=0}^{\infty} L^{jk}_{mn} z^m \bar{\zeta}^n.$$

Then if x_0^j $(j=0,1,\ldots,M)$ *are any* $M+1$ *real numbers and* x_n^j $(j=0,1,\ldots,M;$ $n=1,2,\ldots,N)$ *are any* $N(M+1)$ *complex numbers,*

$$\begin{aligned}
&2\sum_{j=0}^{M} x_0^j \operatorname{Re}\sum_{n=0}^{N}\sum_{k=0}^{M}(K_{0n}^{jk}x_n^k+L_{0n}^{jk}\overline{x_n^k})\\
&\qquad+\sum_{m=1}^{\infty} m\sum_{j=0}^{M}\left|\sum_{n=0}^{N}\sum_{k=0}^{M}(K_{mn}^{jk}x_n^k+L_{mn}^{jk}\overline{x_n^k})\right|^2 \qquad (2)\\
&\leqq \sum_{n=1}^{N}\frac{1}{n}\sum_{j=0}^{M}|x_n^j|^2.
\end{aligned}$$

Equality in (2) *is possible only if the complement of* $D_0\cup\cdots\cup D_M$ *with respect to the unit disc has zero area.*

Corollary 1. *Under the same hypotheses of Theorem 1,*

$$\operatorname{Re}\sum_{m,n=0}^{N}\sum_{j,k=0}^{M}(K_{mn}^{jk}x_m^j x_n^k+L_{mn}^{jk}x_m^j\overline{x_n^k})\leqq\sum_{n=0}^{N}\sum_{j=0}^{M}|x_n^j|^2. \qquad (3)$$

For equality to hold in (3) *it is necessary that*

$$\begin{aligned}
&\operatorname{Re}\sum_{n=0}^{N}\sum_{j,k=0}^{M}(K_{0n}^{jk}x_0^j x_n^k+L_{0n}^{jk}x_0^j\overline{x_n^k})=0\\
&\sum_{n=0}^{N}\sum_{k=0}^{M}(K_{mn}^{jk}x_n^k+L_{mn}^{jk}\overline{x_n^k})=\frac{\overline{x_m^j}}{m},\quad m=1,\ldots,N;\ j=0,\ldots,M \qquad (4)\\
&\qquad\qquad=0,\qquad m\geqq N+1;\ j=0,\ldots,M.
\end{aligned}$$

Remark 1. If each $x_0^j=0$, these inequalities reduce to a special case of some recent results of Gromova and Lebedev [6]. Thus our results are more general than theirs in some respects but less general in others. However, the availability of free real parameters has proved to be of great usefulness in the application of this type of inequality to coefficient problems. (See Schiffer and Tammi [14, 15]; DeTemple [3, 5].)

Corollary 2. *Let* $\{F_0,\ldots,F_M\}\in\mathcal{M}_\infty(w_0,\ldots,w_M)$, *and let the coefficients* K_{mn}^{jk} $(m,n=0,1,\ldots;\ j,k=0,1,\ldots,M)$ *be defined by*

$$\begin{aligned}
\log\frac{F_j(z)-F_j(\zeta)}{z-\zeta}&=\sum_{m,n=0}^{\infty}K_{mn}^{jj}z^m\zeta^n\\
\log[F_j(z)-F_k(\zeta)]&=\sum_{m,n=0}^{\infty}K_{mn}^{jk}z^m\zeta^n\qquad(j\neq k).
\end{aligned}\qquad(5)$$

Then if x_0^j $(j=0, 1, \ldots, M)$ *are any real numbers for which* $\sum_{j=0}^{M} x_0^j = 0$, *and* x_n^j $(j=0, 1, \ldots, M; n=1, 2, \ldots, N)$ *are any complex numbers,*

$$2\sum_{j=0}^{M} x_0^j \operatorname{Re} \sum_{n=0}^{N} \sum_{k=0}^{M} K_{0n}^{jk} x_n^k + \sum_{m=1}^{\infty} m \sum_{j=0}^{N} \left| \sum_{n=0}^{N} \sum_{k=0}^{M} K_{mn}^{jk} x_n^k \right|^2 \leqq \sum_{n=1}^{N} \frac{1}{n} \sum_{j=0}^{M} |x_n^j|^2. \quad (6)$$

Equality in (6) *is possible only if the complement of* $D_0 \cup \cdots \cup D_M$ *with respect to the extended plane has zero area.*

Corollary 3. *Under the hypotheses of Corollary 2,*

$$\operatorname{Re} \sum_{m,n=0}^{N} \sum_{j,k=0}^{M} K_{mn}^{jk} x_m^j x_n^k \leqq \sum_{n=1}^{N} \frac{1}{n} \sum_{j=0}^{M} |x_n^j|^2. \quad (7)$$

For equality to hold in (7) *it is necessary that*

$$\begin{aligned} &\operatorname{Re} \sum_{n=0}^{N} \sum_{j,k=0}^{M} K_{0n}^{jk} x_n^k = 0 \\ &\sum_{n=0}^{N} \sum_{k=0}^{M} K_{mn}^{jk} x_n^k = \frac{\overline{x_m^j}}{m}, \qquad m=1, 2, \ldots, N; j=0, \ldots, M \\ &\qquad = 0, \qquad m \geqq N+1; j=0, \ldots, M. \end{aligned} \quad (8)$$

Theorem 2. *Let* $\{F_0, F_1, \ldots, F_M\} \in \mathscr{M}(\infty, w_1, \ldots, w_M)$, *and let the coefficients* K_{mn}^{jk} $(m, n=0, 1, \ldots; j, k=0, 1, \ldots, M)$ *be defined by*

$$\begin{aligned} \log \frac{F_0(z)-F_0(\zeta)}{z^{-1}-\zeta^{-1}} &= \sum_{m,n=0}^{\infty} K_{mn}^{00} z^m \zeta^n \\ \log \frac{F_j(z)-F_j(\zeta)}{z-\zeta} &= \sum_{m,n=0}^{\infty} K_{mn}^{jj} z^m \zeta^n \qquad (j \neq 0) \\ \log[F_j(z)-F_k(\zeta)] &= \sum_{m,n=0}^{\infty} K_{mn}^{jk} z^m \zeta^n \qquad (j \neq k, j \neq 0, k \neq 0) \\ \log[z(F_0(z)-F_k(\zeta))] &= \sum_{m,n=0}^{\infty} K_{mn}^{0k} z^m \zeta^n \qquad (k \neq 0). \end{aligned} \quad (9)$$

Let x_0^j $(j=0, 1, \ldots, M)$ *be arbitrary real numbers satisfying* $\sum_{j=0}^{M} x_0^j = 0$, *and let* x_n^j $(j=0, 1, \ldots, M; n=1, 2, \ldots, N)$ *be arbitrary complex numbers. Then*

$$2\sum_{j=0}^{M} x_0^j \operatorname{Re} A_0^j + \sum_{m=1}^{\infty} m \sum_{j=0}^{M} |A_m^j|^2 \leqq \sum_{n=1}^{N} \frac{1}{n} \sum_{j=0}^{M} |x_n^j|^2, \quad (10)$$

where

$$\begin{aligned}
&A_0^0 = x_0^0 \log F_0'(0) \qquad \left(F_0'(0) = \lim_{z\to 0} \frac{1}{z\,F_0(z)}\right)\\
&A_m^0 = \sum_{\substack{n=0\\(n+k>0)}}^{N} \sum_{k=0}^{M} K_{mn}^{0k} x_n^k \qquad (m \geqq 1)\\
&A_m^j = \sum_{\substack{n=0\\(n+k>0)}}^{N} \sum_{k=0}^{M} K_{nm}^{kj} x_n^k \qquad (j \neq 0, m \geqq 0).
\end{aligned} \tag{11}$$

Equality in (10) *is possible only if the complement of* $D_0 \cup \cdots \cup D_M$ *with respect to the extended plane has zero area.*

Theorem 3. *Let* $\{F_0, F_1, \ldots, F_M\} \in \mathcal{M}(\infty, w_1, \ldots, w_M)$ *and furthermore suppose that* $F_0 \neq 0$. *Let the coefficients* K_{mn}^{jk} $(m, n = 0, 1, \ldots; j, k = 0, \ldots, M)$ *be defined by*

$$\begin{aligned}
\log \frac{F_0(z)^{-1} - F_0(\zeta)^{-1}}{z-\zeta} &= \sum_{m,n=0}^{\infty} K_{mn}^{00} z^m \zeta^n\\
\log \frac{F_j(z) - F_j(\zeta)}{z-\zeta} &= \sum_{m,n=0}^{\infty} K_{mn}^{jj} z^m \zeta^n \qquad (j \neq 0)\\
\log \left[1 - \frac{F_k(\zeta)}{F_0(z)}\right] &= \sum_{m,n=0}^{\infty} K_{mn}^{0k} z^m \zeta^n \qquad (k \neq 0)\\
\log [F_j(z) - F_k(z)] &= \sum_{m,n=0}^{\infty} K_{mn}^{jk} z^m \zeta^n \qquad (j \neq k, j \neq 0, k \neq 0).
\end{aligned} \tag{12}$$

Let x_0^j $(j = 0, 1, \ldots, M)$ *be arbitrary real numbers satisfying* $\sum_{j=0}^{M} x_0^j = 0$, *and let* x_n^j $(j = 0, 1, \ldots, M; n = 1, 2, \ldots, N)$ *be arbitrary complex numbers. Then*

$$2 \sum_{j=0}^{M} x_0^j \operatorname{Re} A_0^j + \sum_{m=1}^{\infty} m \sum_{j=0}^{M} |A_m^j|^2 \leqq \sum_{n=1}^{N} \frac{1}{n} \sum_{j=0}^{M} |x_n^j|^2, \tag{13}$$

where

$$\begin{aligned}
&A_m^0 = \sum_{n=0}^{N} \sum_{k=0}^{M} K_{mn}^{0k} x_n^k\\
&A_m^j = \sum_{n=0}^{N} \sum_{k=0}^{M} K_{nm}^{kj} x_n^k \qquad (j \neq 0).
\end{aligned} \tag{14}$$

Equality in (13) *is possible only if the complement of* $D_0 \cup \cdots \cup D_M$ *with respect to the extended plane has zero area.*

Remark 2. Lebedev [10, 11] has also given area inequalities for the unbounded classes $\mathcal{M}(w_0, \ldots, w_M)$. Again, though, our formulation involves some parameters and coefficient combinations not found in the Lebedev inequalities.

2. Proofs

The proofs rely on a version of the area method which is due to Schiffer. This procedure has been successfully utilized in the papers of Schiffer and Tammi [15], Hummel and Schiffer [8], Aharonov and Szargel [2], and DeTemple [4]. In certain respects, the present note is a generalization of [4], and so the details of the calculations are kept to a minimum.

Proof of Theorem 1. For an arbitrary system $\{F_0, \ldots, F_M\}$ in $\mathcal{M}_1(w_0, \ldots, w_M)$, the defining properties of this class assure us that the Taylor series expansions in (1) converge in the unit bicylinder $|z|<1 \times |\zeta|<1$. It should be observed that except for the cases $m=n=0$, $j \neq k$, we have the symmetry relation $K^{jk}_{mn}=K^{kj}_{nm}$; in the exceptional cases we still have equality of the real parts. We also have a Hermitian symmetry of the form $L^{jk}_{mn}=\overline{L^{kj}_{nm}}$ with no exceptions.

The rational functions $\Phi^j_n(t)$ $(n=1,2,\ldots; j=0,1,\ldots,M)$, with poles only at $t=1/w_j$, are introduced by means of the generating functions

$$\log \frac{1-t\,w_j}{1-t\,F_j(z)} = \sum_{n=1}^{\infty} \frac{1}{n}\, \Phi^j_n(t)\, z^n. \tag{15}$$

For $0<|z|<1$ fixed and ζ sufficiently near to the origin, we can develop the left hand sides of (1) in a power series of ζ. Comparing coefficients gives the following identities:

$$\begin{aligned}
&\sum_{m=0}^{\infty} K^{jj}_{m0}\, z^m = \log\left[(F_j(z)-w_j)/z\right] \\
&\sum_{m=0}^{\infty} K^{jk}_{m0}\, z^m = \log\left[F_j(z)-w_k\right] && (j \neq k) \\
&n\sum_{m=0}^{\infty} K^{jj}_{mn}\, z^m = z^{-n} - \Phi^j_n(1/F_j(z)) && (n \neq 0) \\
&n\sum_{m=0}^{\infty} K^{jk}_{mn}\, z^m = -\Phi^k_n(1/F_j(z)) && (j \neq k,\ n \neq 0) \\
&\sum_{m=0}^{\infty} L^{jk}_{m0}\, z^m = -\log\left[1-\overline{w}_k\, F_j(z)\right] \\
&n\sum_{m=0}^{\infty} L^{jk}_{mn}\, z^m = \overline{\Phi^k_n\left[\overline{F_j(z)}\right]} && (n \neq 0).
\end{aligned} \tag{16}$$

For $0<\rho<1$ given, let Δ_ρ be the $M+2$ connected sub-region of $|w|<1$ bounded by the unit circumference $|w|=1$ and the curves

$w=F_j(\rho\, e^{i\theta})$. With the given numbers x_n^j, define

$$q(w)=\sum_{j=0}^{M}\left[x_0^j \log\frac{w-w_j}{1-\overline{w}_j w}-\sum_{n=1}^{N}\frac{1}{n}\left\{x_n^j\,\Phi_n^j\left(\frac{1}{w}\right)-\overline{x_n^j}\,\overline{\Phi_n^j(\overline{w})}\right\}\right], \tag{17}$$

which is seen to be analytic in Δ_ρ. In addition it has a single-valued real part due to the realness of the parameters x_0^j, and so Green's formula may be employed to show

$$0\leqq \iint_{\Delta_\rho} |q'(w)|^2=\frac{1}{i}\int_{\partial\Delta_\rho} \operatorname{Re} q(w)\, dq(w).$$

Since $\operatorname{Re} q(w)=0$ on $|w|=1$, this inequality goes over to $I_0+\cdots+I_M\leqq 0$, where

$$I_j=\frac{1}{i}\int_0^{2\pi} \operatorname{Re} q[F_j(\rho\, e^{i\theta})]\,\frac{dq[F_j(\rho\, e^{i\theta})]}{d\theta}\, d\theta.$$

Inserting the identities (16) into the definition (17) we obtain the expansions

$$q[F_j(z)]=x_0^j \log z-\sum_{n=1}^{N}\frac{x_n^j}{n\, z^n}+\sum_{m=0}^{\infty} A_m^j\, z^m \tag{18}$$

where

$$A_m^j=\sum_{n=0}^{N}\sum_{k=0}^{M}(K_{mn}^{jk}\, x_n^k+L_{mn}^{jk}\, \overline{x_n^k}). \tag{19}$$

These expansions, valid in $0<|z|<1$, allow one to evaluate the integrals

$$I_j=2\pi(x_0^j)^2 \log\rho+2\pi\, x_0^j \operatorname{Re} A_0^j+\sum_{m=1}^{\infty} m\,|A_m^j|^2\,\rho^{2m}-\sum_{n=1}^{N}\frac{1}{n}|x_n^j|^2\,\rho^{-2n}.$$

The proof of inequality (2) is completed by summing over the indices and passing to the limit $\rho=1$.

Proof of Corollary 1. In terms of the A_m^j given in equation (19), inequality (3) is equivalent to showing $\mu+\operatorname{Re}\sum_{m=1}^{N}\sum_{j=0}^{M} A_m^j\, x_m^j\leqq \nu$, where $\mu=\sum_{j=0}^{M} x_0^j \operatorname{Re} A_0^j$ and $\nu=\sum_{n=1}^{N}\frac{1}{n}\sum_{j=0}^{M}|x_n^j|^2$. By successively using Schwarz's inequality, Theorem 1, and completion of the square, we have

$$\begin{aligned}\mu+\operatorname{Re}\sum_{m=1}^{N}\sum_{j=0}^{M} A_m^j\, x_m^j &\leqq \mu+\sum_{m=1}^{N}\sum_{j=0}^{M}\left|\frac{x_m^j}{\sqrt{m}}\cdot\sqrt{m}\, A_m^j\right|\\ &\leqq \mu+\left\{\nu\sum_{m=1}^{N}\sum_{j=0}^{M} m\,|A_m^j|^2\right\}^{\frac{1}{2}}\\ &\leqq \mu+\{\nu(\nu-2\mu)\}^{\frac{1}{2}}\\ &\leqq \nu.\end{aligned}$$

Proof of Corollary 2. First suppose that the D_j all lie in a large circle $|z|<B$. Theorem 1 can then be applied to the system

$$\{B^{-1}F_0, \ldots, B^{-1}F_M\} \in \mathcal{M}_1(B^{-1}w_0, \ldots, B^{-1}w_M).$$

The result then follows by sending $B\to\infty$. However, to ensure that the left hand side of the inequalities do not converge to $-\infty$, the coefficient of $-\log B$, namely $\left(\sum_{j=0}^{M} x_0^j\right)^2$, is set to zero.

Proof of Corollary 3. The method of showing Corollary 1 can be used.

Proof of Theorem 2. The steps used to proof Theorem 1 are followed, making the following changes. The $\Phi_n^0(t)$ are now taken to be the polynomials in t defined by the generating function

$$-\log\left[z F_0'(0)(F_0(z)-t)\right]=\sum_{n=1}^{\infty}\frac{1}{n}\Phi_n^0(t)\,z^n.$$

The generating functions (15) are retained for $j\neq 0$. The function $q(w)$ given by equation (17) is replaced by

$$q(w)=-\sum_{n=1}^{N}\frac{x_n^0}{n}\Phi_n^0(w)+\sum_{j=1}^{M}\left[x_0^j\log(w-w_j)-\sum_{n=1}^{N}\frac{x_n^j}{n}\Phi_n^j\left(\frac{1}{w}\right)\right].$$

Green's formula is then applied to the $M+1$ connected region Δ_ρ bounded by the curves $w=F_j(\rho\, e^{i\theta})$.

Proof of Theorem 3. Again (15) is retained when $j\neq 0$, but now $\Phi_n^0(t)$ are the polynomials given by the generating function

$$-\log\left[1-\frac{t}{F_0(z)}\right]=\sum_{n=1}^{\infty}\frac{1}{n}\Phi_n^0(t)\,z^n.$$

The remaining steps are analogous to those taken in the proof of Theorem 2; in particular $q(w)$ and Δ_ρ are taken as before.

3. Special Cases and Applications

The inequalities set forth in Section 1 contain many of the Grunsky type of inequalities for the standard classes of univalent mappings of the unit disc. In order to see easily which theorem is applicable, these classes are described below in terms of $\mathcal{M}_1(w_0, \ldots, w_M)$, $\mathcal{M}_\infty(w_0, \ldots, w_M)$, or $\mathcal{M}(\infty, w_1, \ldots, w_M)$. In addition we list some newer classes of functions which have lately come into consideration.

A function $f(z)$ belongs to:

$$\begin{array}{ll} \Sigma, & \text{if } f\in\mathcal{M}(\infty) \quad \text{and} \quad f'(0)=1; \\ S, & \text{if } f\in\mathcal{M}_\infty(0) \quad \text{and} \quad f'(0)=1; \\ S_1, & \text{if } f\in\mathcal{M}_1(0) \quad \text{and} \quad f'(0)>0; \\ C^*, & \text{if } \left\{\dfrac{1}{f}, f\right\}\in\mathcal{M}(\infty,0); \\ \Gamma^*, & \text{if } \left\{\dfrac{1}{f(z)}, -\overline{f(\bar{z})}\right\}\in\mathcal{M}(\infty,0); \\ D^*, & \text{if } \{f, -f\}\in\mathcal{M}_\infty(1,-1); \\ D^*(\beta), & \text{if } \{f, -f\}\in\mathcal{M}_1(\beta,-\beta), \quad (0<\beta<1). \end{array}$$

Remark 3. A function $f\in C^*$ is a univalent Bieberbach-Eilenberg function, characterized by $f(0)=0$ and $f(z)f(\zeta)\neq 1$ in E^2: $|z|<1\times|\zeta|<1$. Similarly the characterizing properties of a function f in Γ^* are $f(0)=0$ and $f(z)\overline{f(\zeta)}\neq -1$ in E^2. The class D^*, closely related to C^*, was first considered by Guelfer and more recently by DeTemple [4], who also defined the corresponding bounded class $D^*(\beta)$. Both D^* and $D^*(\beta)$ have the same characterizing property that $f(z)+f(\zeta)\neq 0$ in E^2.

The concept of a "pair" is also easily discussed within our framework. Introduced by Aharonov [1], (f, g) is a "pair" when the analytic functions f and g satisfy $f(0)=g(0)=0$ and $f(z)g(\zeta)\neq 1$ in E^2. For example, if $f\in C^*$, then (f, f) is a univalent pair. To complete the list, we note that (f, g) will belong to the class of univalent "pairs" if $(1/f, g)\in\mathcal{M}(\infty,0)$. It should also be observed that if $(f, g)\in\mathcal{M}_\infty(w_0, w_1)$, then $\left(\dfrac{f-w_0}{f-w_1}, \dfrac{g-w_1}{g-w_0}\right)$ is a univalent pair.

Many of the known if not classical inequalities are obtained when an appropriate theorem from Section 1 is specialized to one of the above classes. These include: the Grunsky inequalities for Σ and S; the Nehari inequalities [12] as generalized by Schiffer and Tammi [15] for S_1; the Hummel-Schiffer inequalities for C^* ([8], Eq. (45), p. 95); the inequalities of DeTemple [4] for D^* and $D^*(\beta)$; the Aharonov-Szargel inequalities for univalent "pairs" [2].

Next let us examine the inequalities of Theorems 1 and 2 and Corollary 3 for the case $M=1$, $N=0$, and $x_0^0=1=-x_0^1$.

Theorem 4. (a) *Let $f(z)$ and $g(z)$ be regular and univalent mappings of the unit disc $|z|<1$ into the unit disc. If the image domains have no point in common (i.e., $f(z)\neq g(\zeta)$ for any points z, ζ in the unit disc), then*

$$|f'(z)g'(\zeta)|\leqq\left(\frac{1-|f(z)|^2}{1-|z|^2}\right)\left(\frac{1-|g(\zeta)|^2}{1-|\zeta|^2}\right)\left|\frac{f(z)-g(\zeta)}{1-f(z)\overline{g(\zeta)}}\right|^2. \tag{20}$$

The inequality is sharp: For given z_0, ζ_0, equality will hold only for the functions mapping $|z|<1$ onto domains in $|w|<1$ separated by a circular arc orthogonal to $|w|=1$ and with respect to which $f(z_0)$ and $g(\zeta_0)$ are symmetric points.

(b) *Let $f(z)$ and $g(\zeta)$ be regular and univalent mappings of the unit disc onto non-overlapping domains. Then*

$$\frac{|f'(z)\, g'(\zeta)|}{|f(z)-g(\zeta)|^2} \leqq \frac{1}{(1-|z|^2)(1-|\zeta|^2)}. \tag{21}$$

The inequality is sharp: For given z_0, ζ_0, equality holds only for the functions which map the disc onto complementary half-planes, with $f(z_0)$ and $g(\zeta_0)$ symmetric points with respect to the dividing line.

(c) *Let $f(z)=\dfrac{\alpha}{z}+\cdots$ be a univalent and meromorphic mapping of $|z|<1$ whose range has no point in common with the range of the regular and univalent function $g(z)=w_1+\beta z+\cdots$. Then*

$$|\beta| \leqq |\alpha|. \tag{22}$$

The inequality is sharp, with equality holding only in the case $g(z)$ maps onto a disc and $f(z)$ maps onto the complement of that disc.

Proof. Since the arguments for each of the three parts are analogous, we will show only how part (b) is obtained from Corollary 3. First suppose $z=\zeta=0$. Then inequality (6) becomes (taking $F_0=f$, $F_1=g$, $x_0^0=1=-x_0^1$)

$$\log|f'(0)|+\log|g'(0)|-2\log|f(0)-g(0)| \leqq 0.$$

This is equivalent to (21) when $z=\zeta=0$. For equality to hold we employ the conditions (8) to see that $K_{m0}^{00}-K_{m0}^{01}=0$ and $K_{m0}^{10}-K_{m0}^{11}=0$ for $m\geqq 1$. Returning then to the generating functions (5) and taking real parts we find $\log\left|\dfrac{f(z)-f(0)}{z(f(z)-g(0))}\right|=0$ and $\log\left|\dfrac{g(z)-g(0)}{z(g(z)-f(0))}\right|=0$. Thus $\dfrac{f(z)-f(0)}{f(z)-g(0)}=\varepsilon z$, $|\varepsilon|=1$ and $\dfrac{g(z)-g(0)}{g(z)-f(0)}=\tau z$, $|\tau|=1$ and so it is easily seen the ranges of f and g are complementary half-planes, and $f(0)$ and $g(0)$ are symmetric points. For the case of general points z_0 and ζ_0, f and g are replaced by $f\left(\dfrac{z+z_0}{1+\bar{z}_0 z}\right)$ and $g\left(\dfrac{z+\zeta_0}{1+\bar{\zeta}_0 z}\right)$.

Remark 4. Parts (a) and (b) are due to Nehari ([12], pp. 277–281), who also shows in [13] how the requirement that f and g be univalent can be relaxed by a majorization argument. Part (c) is equivalent to a well known result of Bieberbach and Eilenberg.

For our final application, consider the family $S_1^{(p)}$ of bounded univalent and p-symmetric functions. That is, f is in $S_1^{(p)}$ if $f \in S_1$ and $f(\eta z) = \eta f(z)$, where $\eta = e^{2\pi i/p}$. Next suppose $F(z)$ is any regular and univalent function in $|z|<1$ for which $F(z) \neq \eta^j F(\zeta)$, $j=1, \ldots, p-1$, for all points $|z|<1$, $|\zeta|<1$. We may then apply Theorem 4(a) to the pair of functions $F^{p/2}$ and $-F^{p/2}$ to find (with $z=\zeta=0$)

$$|F'(0)| \leqq \frac{4}{p} |F(0)| \frac{1-|F(0)|^p}{1+|F(0)|^p}, \tag{23}$$

with equality holding only for the functions $\tilde{F}$ which map $|z|<1$ onto a sector symmetric to the ray passing through $\tilde{F}(0)$ and having a vertex angle of $2\pi/p$. Now let $|z_0|<1$ and consider $F = f \circ \tilde{F}$, where $f \in S_1^{(p)}$ and $\tilde{F}$ is that extremizing function of inequality (23) for which $\tilde{F}(0) = z_0$. Not only is F bounded and univalent, but $F(z) \neq \eta^j F(\zeta)$, $j=1, \ldots, p-1$, in $|z|, |\zeta|<1$. Indeed, were this not the case, then $f[\tilde{F}(z_0)] = \eta^{j_0} f[\tilde{F}(\zeta_0)] = f[\eta^{j_0} \tilde{F}(\zeta_0)]$, some j_0, z_0, ζ_0, and so by the univalence of f it would follow that $\tilde{F}(z_0) = \eta^{j_0} \tilde{F}(\zeta_0)$, a contradiction. Thus we may once again utilize the estimate (23). The outcome of this calculation gives us the following result.

Theorem 5. *Let f be a bounded, univalent and p-symmetric mapping of the unit disc (i.e. $f \in S_1^{(p)}$). Then*

$$\left| \frac{z f'(z)}{f(z)} \right| \leqq \frac{1-|f(z)|^p}{1+|f(z)|^p} \, \frac{1+|z|^p}{1-|z|^p}. \tag{24}$$

The inequality is sharp, with equality holding only for the functions mapping $|z|<1$ onto the disc $|w|<1$ minus p symmetrically placed slits extending radially toward $w=0$ from the periphery $|w|=1$.

Remark 5. Inequality (24) was obtained by Jakubowski [9] by the Loewner parametric method. It is the beginning point in obtaining a sequence of inequalities connecting pairs of the quantities $|z|$, $|f'(0)|$, $|f(z)|$, $|f'(z)|$, given the value of the remaining two quantities.

Additional Comment. Inequalities of a nature similar to those of Section 2 have recently appeared in a paper of Kühnau [9a]. He restricts himself to the case $M=1$ however.

References

1. Aharonov, D.: A generalization of a theorem of J.A. Jenkins. Math. Z. **110**, 218–222 (1969).
2. Aharonov, D., Szargel, J.: On pairs of functions. Technion's Preprint Series No. MT-65, Technion-Israel Institute of Technology, Haifa.

3. DeTemple, D. W.: On coefficient inequalities for bounded univalent functions. Ann. Acad. Sci. Fennicae Ser. A. I, no. 469, 1970.
4. DeTemple, D. W.: Grunsky-Nehari inequalities for a subclass of bounded univalent functions. Trans. Amer. Math. Soc. **159**, 317–328 (1971).
5. DeTemple, D. W.: Generalizations of the Grunsky-Nehari inequalities. Arch. Rat. Mech. Analysis **44**, 93–120 (1971).
6. Gromova, L. L., Lebedev, N. A.: Non-overlapping domains that lie in a disc (Russian. English summary). Vestnik Leningrad Univ. **24**, 7–12 (1969).
7. Guelfer, S. A.: On the class of regular functions which do not take on values w and $-w$. Rec. Math. (Mat. Sbornik), N.S. **19** (61), 33–46 (1946).
8. Hummel, J., Schiffer, M.: Coefficient inequalities for Bieberbach-Eilenberg functions. Arch. Rat. Mech. Analysis **32**, 87–99 (1969).
9. Jakubowski, Z. J.: Les fonctions univalents, p-symmetric et bornées dans la cercle unité. Ann. Polon. Math. **20**, 119–148 (1968).
9a. Kühnau, R.: Über vier Klassen schlichter Funktionen. Math. Nachr. **50**, 17–26 (1971).
10. Lebedev, N. A.: The area principle in the problem of non-overlapping regions. Doklady Akad. Nauk. SSSR **132**, 758–761 (1960) = Soviet Math. Doklady **1**, 640–643 (1960).
11. Lebedev, N. A.: An application of the area principle to non-overlapping domains [Russian]. Trudy Mat. Inst. Steklov **60**, 211–231 (1961).
12. Nehari, Z.: Some inequalities in the theory of functions. Trans. Amer. Math. Soc. **75**, 256–286 (1953).
13. Nehari, Z.: Some functions-theoretic aspects of linear second-order differential equations. J. Analyse Math. **18**, 259–276 (1967).
14. Schiffer, M., Tammi, O.: On bounded univalent functions which are close to identity. Ann. Acad. Sci. Fennicae Ser. A. I, no. 435, 1968.
15. Schiffer, M., Tammi, O.: On the coefficient problem for bounded univalent functions. Trans. Amer. Math. Soc. **140**, 461–474 (1969).

Dr. D. W. DeTemple
Department of Pure and Applied Mathematics
Washington State University
Pullman, Washington 99163
USA

(Received February 9, 1972)

Math. Z. 128, 34–42 (1972)

A Class of Unitary Block Designs

Michael J. Ganley

1. Introduction

A *unital* is a $2-(q^3+1, q+1, 1)$ design. Apart from the unitals associated with the unitary polarities of desarguesian planes of order q^2, there is one other known family, namely those of Ree type, discovered by Lüneburg [4]. Both of these families have doubly-transitive automorphism groups, and the two families are known to be distinct (see O'Nan [5]).

In [2], the following result was proved:

Theorem. *If $\mathfrak{S}$ is a finite commutative semifield admitting a non-trivial involutory automorphism, then the projective plane $\Pi(\mathfrak{S})$ admits a unitary polarity.*

The absolute points and non-absolute lines of a unitary polarity form a unital, and it is the purpose of this paper to study the unitals obtained when $\mathfrak{S}$ is the Dickson semifield of odd order q^2. In particular, we show, in section 4, that there exist infinitely many of these unitals which do not even possess a transitive automorphism group, and so must constitute at least part of a "new" family.

Our method of proof will be to demonstrate the existence of a particular configuration in these "new" Dickson unitals, which is known not to exist in the desarguesian case (O'Nan [5]). In section 3 most of the theory is presented, including a proof of the result of O'Nan, just mentioned, and section 4 is devoted exclusively to examples.

The work in this paper represents part of the author's Ph. D. Thesis, carried out at the University of London, under the supervision of Dr. F. C. Piper, and supported by a grant from the Science Research Council.

2. Preliminary Discussion

Definitions and basic results for finite projective planes may be found in Dembowski [1]. For a full discussion of finite semifields, the reader is referred to Knuth [3]. For convenience, we outline some of the fundamental notions.

A *semifield* (also known as a *division ring* or *distributive quasifield*) $\mathfrak{S}$ is a finite set with two binary operations $+, \cdot$, such that $\mathfrak{S}(+)$ is an abelian group, and $\mathfrak{S}^*(\cdot)$ is a loop, such that both distributive laws hold in $\mathfrak{S}$.

If Π is a projective plane coordinatized by the semifield $\mathfrak{S}$, then the points of $\Pi(\mathfrak{S})$ are elements of the following sets:

$\{(x, y)\colon x, y \in \mathfrak{S}\}$, $\{(m)\colon m \in \mathfrak{S}\}$, $\{(\infty)\colon \infty$ some symbol such that $\infty \notin \mathfrak{S}\}$.
Lines of $\Pi(\mathfrak{S})$ are the point sets:

$$[m, k] = (m) \cup \{(x, y)\colon\ mx + y = k, \text{ and } x, y \in \mathfrak{S}\} \text{ where } m, k \in \mathfrak{S},$$

$$[k] = (\infty) \cup \{(k, y)\colon\ y \in \mathfrak{S}\} \text{ where } k \in \mathfrak{S},$$

$$[\infty] = (\infty) \cup \{(m)\colon\ m \in \mathfrak{S}\}.$$

Incidence is set-theoretic inclusion.

The collineation groups of finite semifield planes are well-known, and may be described as follows:

Let

$$\mathcal{T} = \left\{ \tau(a, b) \ \textit{where}\ \tau(a, b)\colon \begin{array}{l} (x, y) \to (x + a, y + b) \\ (m) \to (m) \\ (\infty) \to (\infty) \end{array} \right\},$$

$$\mathcal{S} = \left\{ \sigma(c) \ \textit{where}\ \sigma(c)\colon \begin{array}{l} (x, y) \to (x, -cx + y) \\ (m) \to (m + c) \\ (\infty) \to (\infty) \end{array} \right\},$$

$$\mathcal{A} = \left\{ \gamma(A, B, C) \ \textit{where}\ \gamma(A, B, C)\colon \begin{array}{l} (x, y) \to (xC, yA) \\ (m) \to (mB) \\ (\infty) \to (\infty) \end{array} \right\}.$$

This last group $\mathcal{A}$ is known as the *autotopism group of* $\Pi(\mathfrak{S})$, and the mappings A, B and C are mappings of $\mathfrak{S}$ onto itself, such that A, B and C are additive, and $(mx)A = mB \cdot xC$ for all $m, x \in \mathfrak{S}$. Elements of $\mathcal{A}$ are known as *autotopisms*.

Result 1. *If $\mathcal{G}$ is the full collineation group of the finite semifield plane $\Pi(\mathfrak{S})$, then $\mathcal{G} = \mathcal{T}\mathcal{S}\mathcal{A}$.*

Now let us assume that $\mathfrak{S}$ is commutative with respect to multiplication, and admits a non-trivial involutory automorphism α. Consider the following mapping of $\Pi(\mathfrak{S})$:

$$\text{Let } \rho(\alpha)\colon \begin{array}{l} (x, y) \leftrightarrow [x^\alpha, -y^\alpha], \\ (m) \leftrightarrow [m^\alpha], \\ (\infty) \leftrightarrow [\infty]. \end{array}$$

Then the mapping $\rho(\alpha)$ is a polarity of $\Pi(\mathfrak{S})$.

Result 2. *If $\mathfrak{S}$ has odd characteristic, then the polarity $\rho(\alpha)$, described above, is a unitary polarity, whose absolute points are (∞) and all points of the form (k, y), where $k, y \in \mathfrak{S}$, and $y = -\frac{1}{2}k^\alpha \cdot k + y_1$ where $y_1 \in \mathfrak{S}$ and $y_1^\alpha = -y_1$.* The proof of this result may be found in [2].

Let $\mathfrak{F} = GF(q)$ where q is odd and not a prime. Let $\mathfrak{S} = \{(x, y): x, y \in \mathfrak{F}\}$ and define two binary operations $\oplus, \circ$, on $\mathfrak{S}$ by $(x, y) \oplus (u, v) = (x+u, y+v)$ and $(x, y) \circ (u, v) = (xu + \delta y^\sigma v^\sigma, yu + xv)$, where $\delta \in \mathfrak{F}$, and δ is a non-square in $\mathfrak{F}$, and $\sigma \in \mathrm{Aut} \cdot \mathfrak{F}$, $\sigma \neq I$. With these operations, $\mathfrak{S}$ becomes a (proper) commutative semifield known as a *Dickson semifield* and the mapping α: $(x, y) \to (x, -y)$ is an involutory automorphism of $\mathfrak{S}$. To distinguish elements of $\mathfrak{S}$ from points in the plane, we will write elements of $\mathfrak{S}$ in the form $x + \lambda y$, where $x, y \in \mathfrak{F}$. Then α: $x + \lambda y \to x - \lambda y$. According to Result 2, $\Pi(\mathfrak{S})$ admits a unitary polarity ρ, the absolute points of which are (∞) and all points of the form $(x_1 + \lambda x_2, -\frac{1}{2}(x_1^2 - \delta x_2^{2\sigma}) + \lambda y_2)$ where $x_1, x_2, y_2 \in \mathfrak{F}$.

Throughout the rest of this paper, we shall assume that $\mathfrak{S}$ is a Dickson semifield of order q^2. We denote the unital described above by $\mathscr{U}q, \sigma$, and the full automorphism group of $\mathscr{U}q, \sigma$ by $\mathrm{Aut} \cdot \mathscr{U}q, \sigma$.

3. Automorphisms of $\mathscr{U}q, \sigma$

Elements of $\mathrm{Aut} \cdot \mathscr{U}q, \sigma$ are of two types, namely those which extend to collineations of $\Pi(\mathfrak{S})$, and those which do not. For the unitals in the desarguesian plane, it is known that all automorphisms are of the first type (O'Nan [5]). If we let $\mathscr{C}(\mathscr{U}q, \sigma)$ denote the subgroup of $\mathrm{Aut} \cdot \mathscr{U}q, \sigma$ consisting of all the automorphisms of the first type, then it is trivial to prove the following lemma:

Lemma 1. *$\mathscr{C}(\mathscr{U}q, \sigma) = C_{\mathscr{G}}(\rho)$, i.e. the subgroup of all collineations of $\Pi(\mathfrak{S})$ which commute with ρ.*

Lemma 2. *$\mathscr{C}(\mathscr{U}q, \sigma) = \{\tau(a, b)\,\sigma(c)\,\gamma(A, B, C)$: $\tau \in \mathscr{T}, \sigma \in \mathscr{S}, \gamma \in \mathscr{A}$, such that $a^\alpha = c$, $ca = b + b^\alpha$, $A\alpha = \alpha A$ and $C\alpha = \alpha B\}$.*

Proof. Suppose θ is a collineation of $\Pi(\mathfrak{S})$, then by result 1, θ is of the form $\theta = \tau(a, b)\,\sigma(c)\,\gamma(A, B, C)$, and

$$\theta\colon\ (x, y) \to ((x+a)C,\ (-c(x+a) + y + b)A),$$

$$[m, k] \to [(m+c)B,\ (k + ma + b)A].$$

Also $\rho(\alpha)$: $(x, y) \leftrightarrow [x^\alpha, -y^\alpha]$, and so if $\rho\theta = \theta\rho$ we must have

$$((m+c)B\alpha,\ -(k+ma+b)A\alpha) = ((m^\alpha + a)C,\ (-c(m^\alpha + a) - k^\alpha + b)A).$$

Since this must be true for all $m, k \in \mathfrak{S}$, then setting $m = 0$ gives $cB\alpha = aC$, and so $mB\alpha = m^\alpha C$, so that $B\alpha = \alpha C$, and then $c^\alpha C = aC$, so $c^\alpha = a$. Also, setting $k=0$, we have $-(ma+b)A\alpha=(-c(m^\alpha+a)+b)A$, which gives $-kA\alpha=-k^\alpha A$, so that α and A commute. Then finally we have

$$-(ma+b)A\alpha=-(ma+b)^\alpha A=(-c(m^\alpha+a)+b)A,$$

so that

$$-(ma+b)^\alpha=-c(m^\alpha+a)+b=-a^\alpha(m^\alpha+a)+b=-(ma)^\alpha-a^\alpha\cdot a+b,$$

and so $b+b^\alpha=a^\alpha\cdot a=ca$, and the lemma is proved.

Lemma 3. *$\mathscr{C}(\mathscr{U}q,\sigma)$ is transitive on the affine points of $\mathscr{U}q,\sigma$, and hence either* Aut $\cdot\,\mathscr{U}q,\sigma$ *is doubly transitive on the points of the unital, or else* Aut $\cdot\,\mathscr{U}q,\sigma$ *fixes the point* (∞).

Proof. The collineations $\{\tau(a,b)\,\sigma(a^\alpha)\colon\ a,b\in\mathfrak{S},\ b+b^\alpha=a^\alpha\cdot a\}$ are transitive on the lines of $\mathscr{U}q,\sigma$ which contain the point (∞). On any such line, say $[k_1+\lambda k_2]$, the affine absolute points are

$$\{(k_1+\lambda k_2,\ -\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})+\lambda y)\ \text{ where } y\in\mathfrak{F}\}.$$

and clearly we are transitive on such points under the collineations $\{\tau(0,b)\colon\ b\in\mathfrak{S},\ b+b^\alpha=0\}$.

We now move onto the main result in this section. To do this we must consider the following configuration:

Theorem 1. *Suppose $\mathscr{U}q,\sigma$ contains a non-degenerate configuration of the type illustrated in Fig. 1, then* Aut $\cdot\,\mathscr{U}q,\sigma$ *fixes the point* (∞).

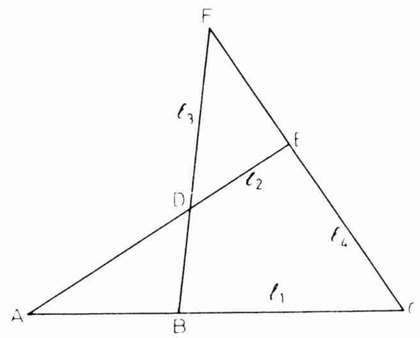


Fig. 1

Proof. Suppose there exists such a configuration in $\mathscr{U}q,\sigma$ and suppose that Aut $\cdot\,\mathscr{U}q,\sigma$ does not fix the point (∞), then by Lemma 3, Aut $\cdot\,\mathscr{U}q,\sigma$ is doubly transitive on the points of the unital, and so there is a configuration of the type illustrated which contains the points (∞) and $(0,0)$.

With this as the only restriction, we label the points and lines as follows:

$$\begin{array}{ll} A=(\infty), & D=(k_1+\lambda k_2,\ -\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})+\lambda s), \\ B=(0,\lambda r), & E=(k_1+\lambda k_2,\ -\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})+\lambda t), \\ C=(0,0), & F=(x_1+\lambda x_2,\ y_1+\lambda y_2), \\ l_1=[0], & l_3=[m_{21}+\lambda m_{22},\lambda r], \\ l_2=[k_1+\lambda k_2], & l_4=[m_{11}+\lambda m_{12},0] \end{array}$$

where $r \neq 0$, $s \neq t$, $k_1+\lambda k_2 \neq 0$, $(m_{11}+\lambda m_{12})(m_{21}+\lambda m_{22}) \neq 0$ and $m_{11}+\lambda m_{12} \neq m_{21}+\lambda m_{22}$. These conditions are necessary to ensure that the configuration is non-degenerate and is contained in $\mathscr{U}q,\sigma$.

We now make use of the fact that F is an absolute point, and that various incidences must hold, to obtain

$$x_1^2-\delta x_2^{2\sigma}=-2y_1, \tag{i}$$

$$(m_{11}+\lambda m_{12})(x_1+\lambda x_2)+(y_1+\lambda y_2)=0, \tag{ii}$$

$$(m_{21}+\lambda m_{22})(x_1+\lambda x_2)+(y_1+\lambda y_2)=\lambda r, \tag{iii}$$

$$(m_{11}+\lambda m_{12})(k_1+\lambda k_2)=\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})-\lambda t, \tag{iv}$$

$$(m_{21}+\lambda m_{22})(k_1+\lambda k_2)=\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})+\lambda(r-s). \tag{v}$$

Write $(m_{21}+\lambda m_{22})-(m_{11}+\lambda m_{12})=(a+\lambda b)$, then we have

$$\left.\begin{array}{l} a x_1+\delta b^\sigma x_2^\sigma=0 \\ b x_1+a x_2=r \\ a k_1+\delta b^\sigma k_2^\sigma=0 \\ b k_1+a k_2=(r-s+t). \end{array}\right\} \tag{vi}$$

We now consider two cases separately:

Case (i) $x_2=0$. Then $x_1 \neq 0$, since otherwise the point F would be incident with the line l_1. From (vi) $a=0$ and hence $b \neq 0$, which means that $k_2=0$ and $k_1 \neq 0$. From (iv) $m_{11}=\tfrac{1}{2}k_1$, and from (ii) $m_{11}x_1+y_1=0$. Using (i), $x_1^2=-2y_1=2m_{11}x_1$, and so $m_{11}=\tfrac{1}{2}x_1$. Hence $x_1=k_1$, and the point F is incident with l_2, and the configuration is degenerate.

Case (ii) $x_2 \neq 0$. In this case, (vi) implies that $a \neq 0$ and $k_2 \neq 0$.

We have $x_1=-\delta b^\sigma a^{-1} x_2^\sigma$, and hence

$$x_1^2-\delta x_2^{2\sigma}=\delta x_2^{2\sigma}(\delta b^{2\sigma}a^{-2}-1). \tag{vii}$$

From (ii)

$$m_{11}x_1+\delta m_{12}^\sigma x_2^\sigma=-y_1. \tag{viii}$$

From (iv)

$$m_{11}k_1+\delta m_{12}^{\sigma}k_2^{\sigma}=\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma}). \tag{ix}$$

From (vi) $k_1=-\delta b^{\sigma}a^{-1}k_2^{\sigma}$, and so (ix) becomes

$$\tfrac{1}{2}(k_1^2-\delta k_2^{2\sigma})=-m_{11}\delta b^{\sigma}a^{-1}k_2^{\sigma}+\delta m_{12}^{\sigma}k_2^{\sigma}=\delta k_2^{\sigma}(m_{12}^{\sigma}-m_{11}b^{\sigma}a^{-1}). \tag{x}$$

Then from (vi), (viii) and (x)

$$-y_1=-\delta m_{11}b^{\sigma}a^{-1}x_2^{\sigma}+\delta m_{12}^{\sigma}x_2^{\sigma}=\delta x_2^{\sigma}(m_{12}^{\sigma}-m_{11}b^{\sigma}a^{-1})$$

i.e.

$$-2y_1=x_2^{\sigma}k_2^{-\sigma}(k_1^2-\delta k_2^{2\sigma}). \tag{xi}$$

Finally from (vi)

$$\begin{aligned}k_1^2-\delta k_2^{2\sigma}&=\delta^2 b^{2\sigma}a^{-2}k_2^{2\sigma}-\delta k_2^{2\sigma}\\&=\delta k_2^{2\sigma}(\delta b^{2\sigma}a^{-2}-1).\end{aligned}$$

Then from (xi)

$$-2y_1=\delta x_2^{\sigma}k_2^{\sigma}(\delta b^{2\sigma}a^{-2}-1). \tag{xii}$$

From (i), (vii) and (xii)

$$\delta x_2^{2\sigma}(\delta b^{2\sigma}a^{-2}-1)=\delta x_2^{\sigma}k_2^{\sigma}(\delta b^{2\sigma}a^{-2}-1).$$

Since δ is a non-square in $\mathfrak{F}$, then $x_2=k_2$, and then from (vi) $x_1=k_1$ and again the point F is incident with l_2. This then completes the proof of the theorem. Note that we have also shown that there is no configuration, of the type in Fig. 1, which is contained in $\mathcal{U}q,\sigma$ and which has the point (∞) as one of its vertices. We also have the following result, due to O'Nan [5].

Corollary. *If q is odd, then the unital associated with the unitary polarity in the desarguesian plane of order q^2 does not contain a configuration of the type illustrated.*

Proof. By considering $\mathcal{U}q,\sigma$ when $\sigma=I$ we obtain the desarguesian unital. As remarked in the introduction, the full automorphism group of this unital is doubly transitive on the points, and so following the proof of Theorem 1 will prove the corollary.

4. Examples

Theorem 2. *Let p be a prime, $p\equiv 3\pmod 4$. Then there is an odd integer $r\leqq 2p-1$, such that whenever $q=p^{rk}$, where k is odd, and σ: $x\to x^p$, then $\mathrm{Aut}\cdot\mathcal{U}q,\sigma$ fixes the point (∞).*

Proof. Let $\mathfrak{C}$ be the configuration as labelled in Fig. 1. Let $\mathfrak{F}=GF(q)$ where $q\equiv 3\pmod 4$. Choose $\delta=-1$, a non-square in $\mathfrak{F}$, and label the points and lines of $\mathfrak{C}$ as follows:

$$\begin{aligned} &A=(0,\lambda a), && D=(b+\lambda 2,\ -\tfrac{1}{2}(b^2+4)+\lambda b),\\ &B=(0,\lambda b), && E=(1+\lambda,\ -1+\lambda),\\ &C=(0,0), && F=(\lambda 2,\ -2),\\ &l_1=[0], && l_2=[m_1+\lambda m_2, \lambda a],\\ &l_3=[\tfrac{1}{2}b-\lambda, \lambda b], && l_4=[-\lambda, 0] \end{aligned}$$

where $a, b \in \mathfrak{F}$ and $ab \neq 0$, $a \neq b$. It is straightforward to check that the points $A-F$ are absolute points and that the incidences required for l_1, l_3 and l_4 are satisfied. Hence we need only for A, D and E to be incident with l_2 for the configuration $\mathfrak{C}$ to exist in the unital $\mathcal{U}q, \sigma$.

Now $E \in l_2$ if and only if

$$m_1 - m_2^\sigma - 1 = 0 \tag{i}$$

and

$$m_2 + m_1 + 1 = a. \tag{ii}$$

$D \in l_2$ if and only if

$$m_1 b - 2m_2^\sigma - \tfrac{1}{2}(b^2+4) = 0, \tag{iii}$$

and

$$m_2 b + 2m_1 + b = a. \tag{iv}$$

Firstly note that $b \neq 2$. Solving (ii) and (iv) gives

$$\left.\begin{aligned} m_1 &= a(b-1)(b-2)^{-1}\\ m_2 &= (2-a-b)(b-2)^{-1}. \end{aligned}\right\} \tag{v}$$

Solving (i) and (ii)

$$\left.\begin{aligned} m_1 &= \tfrac{1}{2}b^2(b-2)^{-1}\\ m_2^\sigma &= (\tfrac{1}{2}b^2 - b + 2)(b-2)^{-1}. \end{aligned}\right\} \tag{vi}$$

From (v) and (vi) $a = \tfrac{1}{2}b^2(b-1)^{-1}$.

Thus

$$\begin{aligned} (\tfrac{1}{2}b^2 - b + 2)(b-2)^{-1} &= (2-a-b)^\sigma (b-2)^{-\sigma}\\ &= (2 - \tfrac{1}{2}b^2(b-1)^{-1} - b)^\sigma (b-2)^{-\sigma}. \end{aligned}$$

This equation reduces to

$$f(b) \equiv b^{2\sigma+2} + b^{2\sigma+1} - 2b^{2\sigma} - 3b^{\sigma+2} + 2b^2 = 0.$$

$f(b)=0$ has a double root $b=0$ and a single root $b=2$, both of which we do not require. If, however, $f(b)=0$ has another root in $GF(q)$, then $\mathcal{U}q, \sigma$ contains a non-degenerate configuration $\mathfrak{C}$, and $\mathcal{U}q, \sigma$ is a "new" unital.

Write $f(b) = b^2(b-2)\phi(b)$. Let $GF(q)$ have characteristic p, and let σ: $x \to x^p$. Then $\phi(b)$ is a polynomial of degree $2p-1$ over the ground field $GF(p)$. Write $\phi(b) = \phi_1(b)\phi_2(b)\ldots\phi_h(b)$, where $\phi_i(b)$ is an irreducible polynomial over $GF(p)$, $(i=1\ldots h)$, then the degree of $\phi(b)=$

$\sum_{i=1}^{h}$ degree $\phi_i(b)=2p-1$, and so at least one of the $\phi_i(b)$ has odd degree, r, say. Consequently if we choose $q=p^r\equiv 3 \pmod 4$, then $\phi(b)$ has a root in $GF(q)$, and also in any other field, which has $GF(p^r)$ as a subfield, and in which -1 is a non-square, i.e. in any finite field of order p^{rk} where k is odd. This remark completes the proof of the theorem.

As a specific example, consider the field $GF(3^{3k})$ where k is odd, and let σ: $x\to x^3$. The equation we require to solve becomes $b^8+b^7+b^6+2b^2=0$ i.e. $b^2(b+1)(b^2+b+2)(b^3+2b^2+1)=0$. Now b^3+2b^2+1 is irreducible over $GF(3)$ and hence whenever $q=3^{3k}$, where k is odd, and σ: $x\to x^3$ we obtain unitals which are isomorphic to neither the desarguesian nor Ree unitals having q^3+1 points.

As an obvious corollary to Theorem 2 we have

Corollary. *Let p be a prime, $p\equiv 3 \pmod 4$ and let s be any positive integer. Then there exists an odd integer $r\leqq 2p^s-1$, such that whenever $q=p^{rk}$, where k is odd and $s<rk$, and σ: $x\to x^{p^s}$, then* Aut $\cdot\,\mathcal{U}\,q,\sigma$ *fixes the point* (∞).

In the case when $q\equiv 1 \pmod 4$, we have some difficulty in choosing a suitable element δ which will make the working fairly straightforward. However, we conclude this section by giving two further examples, both found by trial and error methods, for the case $q\equiv 1 \pmod 4$.

Example 1. Let $\mathfrak{F}=GF(9)$, and write elements of $\mathfrak{F}$ in the form $a+b\sqrt{2}$, where $a,b\in GF(3)$. Choose $\delta=1+\sqrt{2}$, and note that σ: $a+b\sqrt{2}\to a-b\sqrt{2}$.

Let

$A=(0,\lambda)$,	$D=(2,1+\lambda)$,
$B=(0,\lambda(1+\sqrt{2}))$,	$E=(\sqrt{2}+\lambda(1+\sqrt{2}),2\sqrt{2}+\lambda 2\sqrt{2})$,
$C=(0,0)$,	$F=(2+\sqrt{2}+\lambda(1+\sqrt{2}),1+\lambda(2+2\sqrt{2}))$,
$l_1=[0]$,	$l_3=[1+\lambda 2\sqrt{2},\lambda(1+\sqrt{2})]$,
$l_2=[1,\lambda]$,	$l_4=[1+\sqrt{2}+\lambda 2,0]$.

Example 2. Let $\mathfrak{F}=GF(25)$, write elements in the form $a+b\sqrt{2}$, $a,b\in GF(5)$. Choose $\delta=\sqrt{2}$.

Let

$A=(0,\lambda)$,	$D=(2,3+\lambda)$,
$B=(0,\lambda(2+\sqrt{2}))$,	$E=(3+4\sqrt{2}+\lambda 4,2+\sqrt{2}+\lambda 2)$,
$C=(0,0)$,	$F=(3\sqrt{2}+\lambda\sqrt{2},1+\sqrt{2}+\lambda(4+\sqrt{2}))$,
$l_1=[0]$,	$l_3=[1+\lambda(3+3\sqrt{2}),\lambda(2+\sqrt{2})]$,
$l_2=[1,\lambda]$,	$l_4=[3+3\sqrt{2}+\lambda 2,0]$.

Theorem 3. *Let $q=9$ or 25, and let σ be the non-trivial automorphism of $GF(q)$, then $\mathscr{U}q,\sigma$ are "new" unitals.*

References

1. Dembowski, P.: Finite geometries. Berlin-Heidelberg-New York: Springer 1968.
2. Ganley, M.J.: Polarities in translation planes. Geometriae Dedicata **1**, 28–40 (1972).
3. Knuth, D.E.: Finite semifields and projective planes. J. Algebra **2**, 182–217 (1965).
4. Lüneburg, H.: Some remarks concerning the Ree groups of type (G_2). J. Algebra **3**, 256–259 (1966).
5. O'Nan, M.E.: Automorphisms of unitary block designs. J. Algebra **20**, 495–511 (1972).

Dr. M.J. Ganley
Department of Mathematics
University of York
York
England

(Received January 20, 1972)

Math. Z. 128, 43–74 (1972)

Variationsprobleme vom Dirichlet-Typ mit einer Ungleichung als Nebenbedingung

Friedrich Tomi

Einleitung

Es sei G ein Gebiet des $\mathbb{R}^n$ und

$$I(v)=\int f(x, v, \nabla v)\, dx$$

sei ein Variationsintegral, das auf einer Klasse $\mathscr{K}$ von in G erklärten, $\mathbb{R}^N$-wertigen Funktionen $v=v(x)$ definiert ist. Wir betrachten dann das Problem, das Minimum von I in einer Teilmenge $\mathscr{K}_0$ von $\mathscr{K}$ zu finden, wobei $\mathscr{K}_0$ aus allen Elementen von $\mathscr{K}$ besteht, die – außer einer Randbedingung – einer Ungleichung der Form

$$F(x, v(x)) \geqq 0 \quad \forall x \in G$$

genügen. Der Autor ist auf diese Fragestellung durch seine vorangegangenen Untersuchungen [19, 20] zum klassischen Plateauschen Problem bei Anwesenheit von Hindernissen geführt worden. Dieses Problem läßt sich hier als Spezialfall unterordnen und besteht anschaulich gesprochen darin, unter allen Flächen, welche eine gegebene Berandung haben und außerhalb gegebener Hindernisse liegen eine Fläche von kleinstem Inhalt zu finden. In den genannten Arbeiten [19, 20] wurde die spezielle Struktur des Plateauschen Problems entscheidend ausgenützt; insbesondere wurde die Tatsache verwendet, daß die Lösungen des Variationsproblems konform parametrisierte Flächen sind. Hier sollen hingegen allgemeinere Klassen von Variationsproblemen behandelt werden. So ist z.B. auch das Dirichletproblem in unseren Untersuchungen enthalten.

Entsprechend der Methodik der modernen Variationsrechnung formulieren wir unser Problem in einer Funktionenklasse, die so umfassend ist, daß der Existenzbeweis keine Mühe macht. Die interessante Frage ist die nach der Regularität der Lösungen, d.h. nach ihrer Stetigkeit und Differenzierbarkeit. Es sei gleich vermerkt, daß unsere Regularitätsresultate nur im Fall $n=2$ eine gewisse Vollständigkeit haben. Während wir nämlich für beliebiges n – unter geeigneten Voraussetzungen – nur die stetige Differenzierbarkeit bereits hölderstetiger Lösungen zeigen können, gelingt es im Fall $n=2$ die Voraussetzung der Hölderstetigkeit zu eliminieren und damit die Regularität derjenigen Lösungen nachzuweisen, die von der Existenztheorie geliefert werden. In der Tat können die Lösungen höherdimensionaler Variationsprobleme Singularitäten aufweisen [5].

Im letzten Paragraphen der Arbeit geben wir eine Anwendung der erzielten Resultate auf das Plateausche Problem. Gegenüber den Ergebnissen aus [19, 20] werden dabei folgende Fortschritte erzielt: Wir lassen eine allgemeinere Klasse von Hindernissen zu und behandeln außer der gewöhnlichen Plateauschen Randbedingung auch die sogenannte „halbfreie" Randbedingung ([3] Chap. VI). Der Leser, der speziell am Plateauschen Problem mit Hindernissen interessiert ist, sei auch auf die Arbeiten von Giusti [7], Lewy-Stampacchia [13], Miranda [14], und Nitsche [16] verwiesen. Die Resultate dieser Autoren unterscheiden sich von den unseren u.a. in dem Begriff von „Fläche", der zugrunde gelegt wird. Außerdem bringt der von uns gewählte klassische Zugang zum Plateauproblem eine Beschränkung auf zweidimensionale Flächen (in einem $\mathbb{R}^N$) mit sich, während es die Methoden von Giusti, Lewy-Stampacchia, und Miranda gestatten, $(N-1)$-dimensionale Hyperflächen in $\mathbb{R}^N$ zu betrachten.

Ein wesentlicher Gesichtspunkt, unter dem die vorliegende Arbeit gesehen werden muß, besteht darin, daß wir Variationsprobleme für $\mathbb{R}^N$-wertige Funktionen betrachten. Der skalare Fall ist bereits relativ eingehend untersucht; wir verweisen den Leser auf die in [20] angegebene Literatur.

Nach Einreichen des Manuskripts beim Herausgeber hat der Autor von einer Arbeit von Hildebrandt mit dem Titel "On the regularity of solutions of two-dimensional variational problems with obstructions" (Comm. Pure Appl. Math., erscheint demnächst) Kenntnis erhalten, in welcher den unseren ähnliche Ergebnisse bewiesen werden, jedoch mit einer ganz verschiedenen Methode.

§ 1. Ein Existenzsatz

Ein Existenzsatz für den Typ von Variationsproblem, den wir im Auge haben, kann mit den wohlbekannten Methoden der modernen Variationsrechnung ohne Schwierigkeit abgeleitet werden. Um dem Leser entgegenzukommen, wollen wir dennoch nicht auf eine explizite Formulierung eines solchen Existenzsatzes verzichten. Außerdem ist es für die von uns geplanten Anwendungen zweckmäßig, einen Existenzsatz zur Verfügung zu haben, der genau auf diese Anwendungen zugeschnitten ist.

Der unserem Problem adäquate Funktionenraum ist der Sobolevraum H_2^1. Die Definition und die grundlegenden Eigenschaften der

Räume H_p^m findet der Leser in [15], Chap. 3. Die Norm von H_p^m bezeichnen wir mit

$$\| \ \|_{m,p}.$$

Unter $\mathring{H}_p^m$ verstehen wir die Abschließung von C_0^∞, aufgefaßt als Teilmenge von H_p^m.

Für den Rest des Paragraphen sei G ein beschränktes Gebiet des $\mathbb{R}^n$, dessen Rand ∂G aus endlich vielen Zusammenhangskomponenten besteht, die sämtlich $(n-1)$-dimensionale reguläre Hyperflächen der Klasse C^1 sind. Außerdem möge G lokal stets auf einer Seite von ∂G liegen. Man weiß dann aus der Theorie der Sobolevräume ([15], 3.4), daß sich die auf $H_2^1(G) \cap C^0(\bar{G})$ im gewöhnlichen Sinne definierte Einschränkungsabbildung

$$u \mapsto u | \partial G$$

als stetige Abbildung von $H_2^1(G)$ in $L_2(\partial G)$ fortsetzen läßt. Für jedes $u \in H_2^1(G)$ sind daher die Randwerte $u|\partial G$ wohldefiniert und Elemente von $L_2(\partial G)$. Unter $L_2(\partial G)$ verstehen wir dabei den zum Oberflächenmaß von ∂G gehörenden L_2-Raum.

Es sei nun $\mathscr{R}$ eine Teilmenge von $L_2(\partial G)$ und F eine auf $G \times \mathbb{R}^N$ definierte, reellwertige Funktion. Wir setzen dann

$$\mathscr{K}(\mathscr{R}) := \{u \in H_2^1(G) \colon u|\partial G \in \mathscr{R}\}$$

und

$$\mathscr{K}(\mathscr{R}, F) := \{u \in \mathscr{K}(\mathscr{R}) \colon F(x, u(x)) \geqq 0 \text{ für fast alle } x \in G\}.$$

Aus wohlbekannten Sätzen von Morrey ergibt sich dann

Satz 1. *Voraussetzungen:*

(I) *Es gibt eine Zahl c mit*

$$\int_{\partial G} |v|^2 \, d\omega \leqq c$$

für alle $v \in \mathscr{R}$.

(II) *Es gilt $F \in C^0(G \times \mathbb{R}^N)$.*

(III) *Die Funktion $f = f(x, u, p)$ gehört zu $C^0(G \times \mathbb{R}^n \times \mathbb{R}^{nN})$, ist für jedes feste (x, u) stetig nach p differenzierbar und konvex in p. Ferner gilt die Abschätzung*

$$m|p|^2 \leqq f(x, u, p) \leqq M(1 + |p|^2), \tag{1}$$

wobei m und M feste Zahlen sind, $0 < m \leqq M < \infty$.

(IV) *$\mathscr{K}(\mathscr{R})$ ist abgeschlossen bezüglich schwacher Konvergenz in $H_2^1(G)$ und $\mathscr{K}(\mathscr{R}, F)$ ist nicht leer.*

Behauptung. Das Funktional

$$I(u)=\int_G f(x, u(x), \nabla u(x))\,dx$$

ist auf $H_2^1(G)$ wohldefiniert und es existiert ein $u_0 \in \mathscr{K}(\mathscr{R}, F)$ mit $I(u_0) \leqq I(u)$ für alle $u \in \mathscr{K}(\mathscr{R}, F)$.

Beweis. Es sei

$$d := \inf\{I(u) \colon u \in \mathscr{K}(\mathscr{R}, F)\}$$

und $u_k \in \mathscr{K}(\mathscr{R}, F)$ sei eine Folge mit $I(u_k) \to d$ bei $k \to \infty$. Aus (1) und einer Ungleichung vom Poincaré-Typ ([15], Theorem 3.6.4) schließen wir unter Benützung von Voraussetzung (I), daß es eine Zahl C gibt mit

$$\|u_k\|_{1,2} \leqq C \quad \forall k.$$

Wegen der schwachen Kompaktheit der Kugeln eines Hilbertraumes und wegen der Kompaktheit der Einbettung von $H_2^1(G)$ in $L_2(G)$ ([15], Theorem 3.4.4), können wir daher eine Teilfolge u_k' der Folge u_k auswählen, so daß u_k' schwach in $H_2^1(G)$ und fast überall gegen ein $u \in H_2^1(G)$ konvergiert. Wegen der Voraussetzungen (II) und (IV) muß u wieder zu $\mathscr{K}(\mathscr{R}, F)$ gehören. Die Behauptung des Satzes folgt nun aus einem Unterhalbstetigkeitssatz von Morrey ([15], Theorem 1.8.2).

§ 2. Ein Satz über lineare Variationsungleichungen

In diesem Paragraphen beweisen wir ein Resultat über lineare Variationsungleichungen, welches später bei der Untersuchung nichtlinearer Variationsungleichungen benötigt wird. Unser Satz ist in gewissem Sinne eine Verallgemeinerung eines Resultats von Lewy-Stampacchia ([12], Theorem 3.1).

Es sei $G \subset \mathbb{R}^n$ ein beschränktes Gebiet. Wir schreiben im folgenden H_2^1 und $\mathring{H}_2^1$ statt $H_2^1(G)$ und $\mathring{H}_2^1(G)$. Der Raum $\mathring{H}_2^1$ werde mit der Norm

$$\|v\|_1 := \Big(\int_G |\nabla v|^2\,dx\Big)^{\frac{1}{2}}$$

versehen. Den Dualraum von $\mathring{H}_2^1$ bezeichnen wir mit H^{-1}, die durch $\|\ \|_1$ in H^{-1} induzierte Norm mit $\|\ \|_{-1}$. Für die schwache Konvergenz verwenden wir das Symbol $\rightharpoonup$. Wir beginnen mit

Lemma 1. a) *Ist $u \in H_2^1 \cap L_\infty$, $v \in \mathring{H}_2^1 \cap L_\infty$, so gilt $u v \in \mathring{H}_2^1$.*

b) *Ist $u \in H_2^1 \cap L_\infty$, $v \in \mathring{H}_2^1 \cap L_\infty$ und ist $u_k \in H_2^1 \cap L_\infty$ eine Folge mit*

$$\|u_k\|_{0,\infty} \leqq C \quad \forall k$$

und

$$u_k \rightharpoonup u \quad \text{in } H_2^1,$$

so gilt

$$u_k v \rightharpoonup u v \quad \text{in } \mathring{H}_2^1.$$

c) *Es sei h eine auf* $\mathbb{R}$ *erklärte, gleichmäßig lipschitzstetige Funktion. Dann ist* $h(u)\in H_2^1$ *für jedes* $u\in H_2^1$. *Sind* $u, v\in H_2^1$ *mit* $u-v\in \mathring{H}_2^1$, *so gilt auch* $h(u)-h(v)\in \mathring{H}_2^1$.

d) *Es sei h wie in* c). *Dann ist die Abbildung*

$$u\mapsto h(u)$$

schwach stetig von H_2^1 *in sich.*

Beweis. Die Aussagen a) und c) beweist man leicht durch Approximation. Die Beweise von b) und d) verlaufen sehr ähnlich, so daß wir uns darauf beschränken können, die Behauptung b) zu beweisen. Wir führen einen Widerspruchsbeweis und nehmen an, daß es ein $\Phi\in H^{-1}$, ein $\varepsilon>0$, und eine Teilfolge u_k' der Folge u_k gibt mit

$$|\Phi(u_k' v-u v)|\geqq\varepsilon \quad \forall\, k. \tag{2}$$

Da die Folge der $u_k' v$ in der $\|\ \|_1$-Norm beschränkt ist, existiert ein $w\in \mathring{H}_2^1$ und eine Teilfolge u_k'' der Folge u_k' mit

$$u_k'' v \rightharpoonup w \quad \text{in } \mathring{H}_2^1, \tag{3}$$

woraus auch

$$u_k'' v \rightharpoonup w \quad \text{in } H_2^1$$

folgt. Für jedes $G'\Subset G$ ist aber $H_2^1(G')$ kompakt in $L_2(G')$ eingebettet, so daß man

$$u_k \to u \quad \text{in } L_2(G')$$

und

$$u_k'' v \to w \quad \text{in } L_2(G')$$

erhält. Hieraus schließt man sofort $w=u\,v$, was jedoch einen Widerspruch zu (2) und (3) darstellt.

Wir betrachten im folgenden positive, stetige Funktionale auf $\mathring{H}_2^1$, d.h. stetige Funktionale, welche auf nichtnegativen Funktionen einen nichtnegativen Wert annehmen. Es gilt das folgende

Lemma 2. *Es sei* $\Phi\in H^{-1}$ *ein positives Funktional und* $u\in H_2^1\cap L_\infty$. *Für alle* $v\in \mathring{H}_2^1\cap L_\infty$ *gilt dann die Abschätzung*

$$|\Phi(u\,v)|\leqq\|u\|_{0,\infty}\,\|\Phi\|_{-1}\,\|v\|_1, \tag{4}$$

und das durch die Zuordnung

$$v\mapsto \Phi(u\,v) \tag{5}$$

auf dem Raum $\mathring{H}_2^1\cap L_\infty$ *definierte Funktional läßt sich auf eindeutige Weise zu einem auf ganz* $\mathring{H}_2^1$ *erklärten, stetigen Funktional fortsetzen.*

Beweis. Nach Lemma 1 a) ist die linke Seite von (4) sinnvoll. Es genügt offenbar, (4) zu beweisen. Dazu bemerken wir zunächst, daß nach Lemma 1 c) mit einer Funktion w auch $|w|$ zu $\mathring{H}_2^1$ gehört und außerdem

$$\| |w| \|_1 = \|w\|_1 \tag{6}$$

gilt. Aus der Positivität von Φ entnehmen wir

$$\begin{aligned} |\Phi(u\,v)| &\leqq \Phi(|u\,v|) \leqq \Phi(\|u\|_\infty |v|) \\ &\leqq \|u\|_\infty \|\Phi\|_{-1} \| |v| \|_1 . \end{aligned}$$

Zusammen mit (6) ergibt sich daraus (4).

Lemma 2 rechtfertigt die folgende

Definition. Es sei $u \in H_2^1 \cap L_\infty$ und $\Phi \in H^{-1}$ sei positiv. Dann bezeichnen wir mit $u\Phi$ das durch die Zuordnung (5) eindeutig bestimmte Element von H^{-1}.

Lemma 3. *Es sei $\Phi \in H^{-1}$ ein positives Funktional und es seien u, u_k $(k=1, 2, \dots)$ Elemente von $H_2^1 \cap L_\infty$ mit*

$$u_k \rightharpoonup u \quad \text{in } H_2^1$$

und

$$\sup_k \|u_k\|_{0,\infty} < \infty .$$

Dann gilt

$$u_k \Phi \rightharpoonup u\Phi \quad \text{in } H^{-1}.$$

Beweis. Wegen (4) genügt es, die Limesrelation

$$(u_k \Phi)(v) \rightharpoonup (u\Phi)(v) \quad (k \to \infty)$$

für jedes v aus dem in $\mathring{H}_2^1$ dichten Teilraum $\mathring{H}_2^1 \cap L_\infty$ zu zeigen. Nach Lemma 1 b) hat man aber

$$(u_k \Phi)(v) = \Phi(u_k v) \to \Phi(u\,v) = (u\Phi)(v) \quad (k \to \infty).$$

Der Beweis des nächsten Lemmas kann dem Leser überlassen werden.

Lemma 4. *Es sei V ein Vektorraum über $\mathbb{R}$ und K eine Teilmenge von V. Ferner sei Φ eine Linearform und B eine positiv definite Bilinearform auf V. Dann gibt es höchstens ein $u \in K$, welches die Ungleichung*

$$B(u, v-u) + \Phi(v-u) \geqq 0$$

für alle $v \in K$ erfüllt.

Nach diesen Vorbereitungen können wir nun zur Formulierung und zum Beweis des angekündigten Satzes übergehen. Im folgenden ist über doppelt auftretende Indizes stets zu summieren. Wir betrachten die

bilineare Form
$$B(u,v):=\int a_{ij}(x)\frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^j}\,dx,$$

wobei wir annehmen, daß die Koeffizienten a_{ij} zu $L_\infty(G)$ gehören und für alle $x\in G$ und $y\in\mathbb{R}^n$ den Ungleichungen
$$|y|^2\leqq a_{ij}(x)\,y^i\,y^j\leqq C\,|y|^2 \tag{7}$$
mit einer Zahl C genügen.

Wir verwenden den Begriff des Maßes im Sinne von [1] und erinnern den Leser an folgende Tatsache: Ist Φ ein reelles Maß auf dem Gebiet G, so ist durch
$$|\Phi|(v):=\sup\{\Phi(u)\colon u\in C_0^0(G), |u|\leqq v\}\qquad (v\in C_0^0(G), v\geqq 0)$$
ein positives Maß $|\Phi|$ definiert, welches wir den Absolutbetrag von Φ nennen ([1], Chap. III, § 1.5).

Satz 2. *Es sei $u\in H_2^1$ mit $u\geqq 0$ und Φ sei ein Maß auf G, dessen Absolutbetrag ein stetiges Funktional auf $\mathring{H}_2^1$ ist. Die Funktion u genüge der Variationsungleichung*
$$B(u,v-u)+\Phi(v-u)\geqq 0 \tag{8}$$
für alle $v\in H_2^1$ mit $v-u\in\mathring{H}_2^1$ und $v\geqq 0$.

Dann ist durch die Zuordnung
$$w\mapsto -B(u,w)\qquad (w\in C_0^\infty)$$
ein Maß Φ_0 definiert mit
$$|\Phi_0|\leqq|\Phi|.$$

Beweis. Wir setzen
$$\Phi^+:=\tfrac{1}{2}(|\Phi|+\Phi),\qquad \Phi^-:=\tfrac{1}{2}(|\Phi|-\Phi)$$
so daß
$$\Phi=\Phi^+-\Phi^-$$
gilt mit positiven Maßen Φ^+ und Φ^-, welche außerdem zu H^{-1} gehören. Wir verwenden nun eine Konstruktion von Lewy-Stampacchia ([12], Beweis von Theorem 3.1). Es sei $g\in C^1(\mathbb{R})$ eine reellwertige Funktion mit den Eigenschaften
$$\begin{aligned} &g(t)=0 \quad \text{für } t\leqq -1\\ &g(t)=1 \quad \text{für } t\geqq 0\\ &0\leqq g\leqq 1,\quad g'\geqq 0.\end{aligned} \tag{9}$$

Wir definieren einen nichtlinearen Operator

$$T\colon \mathring{H}_2^1 \to H^{-1}$$

durch

$$T(w) = B(w, \cdot) + g(u+w)\,\Phi^+ .$$

Wir wollen nun mit Hilfe eines wohlbekannten Theorems über monotone Operatoren ([2], Théorème 1) zeigen, daß T surjektiv ist. Dazu genügt es, folgendes nachzuweisen:

(I) T ist schwach stetig,

(II) T ist koerziv,

(III) T ist monoton.

Die Eigenschaft (I) folgt aus Lemma 1 a) und Lemma 3, (II) ergibt sich aus der Abschätzung

$$\big(T(w)\big)(w) \geqq (\|w\|_1)^2 - \|\Phi^+\|_{-1}\,\|w\|_1, \tag{10}$$

welche man aus (4) und (7) erhält. Für (III) ist es offenbar ausreichend, wenn wir die Ungleichung

$$\big((g(v_1) - g(v_2))\,\Phi^+\big)(v_1 - v_2) \geqq 0 \tag{11}$$

für alle $v_1, v_2 \in H_2^1$ mit $v_1 - v_2 \in \mathring{H}_2^1$ zeigen. Für beliebiges $w \in H_2^1$ und $m = 1, 2, \ldots$ setzen wir dazu

$$w^{(m)} := \begin{cases} m & \text{falls } w > m \\ w & \text{falls } -m \leqq w \leqq m \\ -m & \text{sonst.} \end{cases}$$

Nach Lemma 1 c) gilt dann $v_i^{(m)} \in H_2^1$, $v_1^{(m)} - v_2^{(m)} \in \mathring{H}_2^1$ und wegen der speziellen Gestalt der Funktion g hat man

$$g(v_i) = g(v_i^{(m)}).$$

Aus der Monotonie von g und der Positivität von Φ^+ folgt daher

$$\big((g(v_1) - g(v_2))\,\Phi^+\big)(v_1 - v_2) = \lim_{m \to \infty} \Phi^+\big((g(v_1^{(m)}) - g(v_2^{(m)}))(v_1^{(m)} - v_2^{(m)})\big) \geqq 0,$$

so daß (11) gezeigt ist. Definieren wir

$$g_k(t) := g(k\,t) \qquad (k = 1, 2, \ldots),$$

so haben die Funktionen g_k ebenfalls die Eigenschaften (9). Nach dem bereits Bewiesenen existieren daher Elemente $w_k \in \mathring{H}_2^1$ mit

$$B(w_k, \cdot) + g_k(u + w_k)\,\Phi^+ = \Phi^- - B(u, \cdot).$$

Setzen wir $v_k = u + w_k$, so gilt offenbar

$$B(v_k, \cdot) + g_k(v_k)\, \Phi^+ = \Phi^- \tag{12}$$

und

$$v_k - u \in \mathring{H}_2^1. \tag{13}$$

Wir wollen zeigen, daß die Folge v_k schwach gegen u konvergiert. Zunächst zeigen wir jedoch die Ungleichung

$$v_k \geqq -k^{-1}. \tag{14}$$

Wegen $u \geqq 0$ und (13) kann man aus Lemma 1c) schließen, daß die Funktion

$$\bar{v}_k := \min\{v_k + k^{-1}, 0\}$$

zu $\mathring{H}_2^1$ gehört und daher in (12) als Testfunktion zulässig ist. Wegen der speziellen Gestalt von g_k verschwinden aber die Funktionen $g_k(v_k)\, \bar{v}_k^{(m)}$ und man erhält

$$\big(g_k(v_k)\, \Phi^+\big)(\bar{v}_k) = \lim_{m \to \infty} \Phi^+ \big(g_k(v_k)\, \bar{v}_k^{(m)}\big) = 0.$$

Andererseits ist $\Phi^-(\bar{v}_k) \leqq 0$, so daß sich insgesamt aus (12) die Ungleichung

$$B(\bar{v}_k, \bar{v}_k) = B(v_k, \bar{v}_k) \leqq 0$$

ergibt, was offenbar $\bar{v}_k = 0$ und somit (14) zur Folge hat.

Aus (14) ist unmittelbar klar, daß jeder schwache Häufungspunkt der Folge v_k eine nichtnegative Funktion ist. Es soll darüber hinaus gezeigt werden, daß jeder solche Häufungspunkt eine Lösung der Variationsungleichung (8) ist. Dazu zeigen wir zunächst, daß die Ungleichung

$$B(v, v - v_k) + \Phi(v - v_k) \geqq 0 \tag{15}$$

für alle $v \in H_2^1$ mit $v - u \in \mathring{H}_2^1$ und $v \geqq 0$ richtig ist. Unter Beachtung von (9), (11), und (12) ergibt sich

$$\begin{aligned} B(v, v - v_k) + \Phi(v - v_k) &= B(v, v - v_k) + \big(g_k(v)\, \Phi^+\big)(v - v_k) - \Phi^-(v - v_k) \\ &= B(v - v_k, v - v_k) + \big(\big(g_k(v) - g_k(v_k)\big)\, \Phi^+\big)(v - v_k) \geqq 0. \end{aligned}$$

Ist nun v_0 ein schwacher Häufungspunkt der Folge v_k, so ergibt sich aus (15), daß auch

$$B(v, v - v_0) + \Phi(v - v_0) \geqq 0 \tag{16}$$

für alle $v \in H_2^1$ mit $v - u \in \mathring{H}_2^1$ und $v \geqq 0$ gelten muß. Mit einem solchen v ist aber auch $v_0 + t(v - v_0)$ für alle $t \in [0, 1]$ als Testfunktion in (16) zulässig, so daß man

$$t\, B\big(v_0 + t(v - v_0), v - v_0\big) + t\, \Phi(v - v_0) \geqq 0$$

erhält. Nach Division dieser Ungleichung durch t ergibt sich bei $t \to 0$ daß, wie angekündigt, v_0 eine Lösung von (8) ist. Nach Lemma 4 können wir dann aber schließen, daß $v_0 = u$ gelten muß, d.h. die Folge v_k besitzt allenfalls u als schwachen Häufungspunkt. Andererseits entnimmt man aus (10), daß die Folge v_k in H_2^1 beschränkt ist. Nach einem üblichen Kompaktheitsschluß muß sie daher schwach gegen u konvergieren. Aus (12) folgt dann

$$\begin{aligned} |-B(u,w)| &= |\lim_{k\to\infty} B(v_k,w)| \\ &\leqq \limsup_{k\to\infty} |g_k(v_k)\,\Phi^+(w) - \Phi^-(w)| \leqq |\Phi|(|w|) \end{aligned}$$

für alle $w \in C_0^1$. Hieraus sieht man, daß die Linearform $-B(u,\cdot)$ ein Maß Φ_0 definiert mit $|\Phi_0| \leqq |\Phi|$, womit Satz 2 bewiesen ist.

§ 3. Die Differenzierbarkeit der Lösungen einer Klasse von nichtlinearen Variationsungleichungen

In diesem Paragraphen untersuchen wir Variationsungleichungen, wie sie sich für Lösungen gewisser Variationsprobleme von dem in § 1 betrachteten Typ ergeben. Wie bereits in der Einleitung bemerkt, können wir i.a. eine Aussage nur für hölderstetige Lösungen machen. In § 4 werden wir allerdings zeigen, daß im Fall $n=2$ die Lösungen tatsächlich hölderstetig sind.

Da wir uns nur für die Differenzierbarkeit im Innern des Definitionsgebietes der Lösung interessieren, können wir unsere Betrachtung auf ein Teilgebiet einschränken, in dessen Abschluß die Lösung dann noch hölderstetig ist. Es sei also $u = (u^1, \dots, u^N)$ eine solche Lösung, die für ein $\beta > 0$ zum Raum $C^\beta(\bar G) \cap H_2^1(G)$ gehört. Die Klasse der zulässigen Vergleichsfunktionen sei dann

$$\mathscr{K}_u(F) := \{v \in C^0(\bar G) \cap H_2^1(G) \colon v\,|\,\partial G = u\,|\,\partial G,\ F(x, v(x)) \geqq 0 \quad \forall\, x \in G\}.$$

$\mathscr{K}_u(F)$ wird als Teilmenge des linearen Raumes $C^0(\bar G) \cap H_2^1(G)$ aufgefaßt, der auf natürliche Weise zu einem Banachraum gemacht wird. Ist $v \in \mathscr{K}_u(F)$, so definieren wir den „Tangentialraum" von $\mathscr{K}_u(F)$ in v als

$$T_v \mathscr{K}_u(F) := \left\{ w = \left.\frac{dv[t]}{dt}\right|_{t=0} : v[\cdot] \text{ ist eine stetig differenzierbare Abbildung von } [0,1] \text{ in } \mathscr{K}_u(F) \text{ mit } v[0] = v \right\}.$$

Wenn nun $u \in \mathscr{K}_u(F)$ einem Variationsintegral I seinen minimalen Wert in der Menge $\mathscr{K}_u(F)$ erteilt, so muß offenbar die Variationsungleichung

$$d I_u(w) \geqq 0 \tag{17}$$

für alle $w \in T_u \mathscr{K}_u(F)$ gelten. Hierbei bezeichnet dI_u das Differential von I an der Stelle u, welches unter geeigneten Voraussetzungen an den Integranden von I existiert. Wir können nicht die denkbar allgemeinste Form der Ungleichung (17) behandeln, wie sie sich aus einem Variationsintegral ergeben kann, das lediglich den Bedingungen des Existenzsatzes unterworfen ist. Vielmehr müssen wir annehmen, daß (17) die folgende Gestalt hat:

$$\int \left(A_{kl}(x,u)\, a^{ij}(x,u) \frac{\partial u^k}{\partial x^i} \frac{\partial w^l}{\partial x^j} + f_k(x,u,\nabla u)\, w^k \right) dx \geqq 0 \quad \forall\, w \in T_u \mathscr{K}_u(F). \tag{18}$$

Über die Koeffizienten von (18) wollen wir folgendes voraussetzen:

$$\begin{aligned} &A_{kl} \in C^2(G \times \mathbb{R}^N), \qquad a^{ij} \in C^1(G \times \mathbb{R}^N),\\ &(A_{kl}) \text{ und } (a^{ij}) \text{ sind symmetrisch und positiv definit},\\ &f(x,u(x),\nabla u(x)) \text{ ist meßbar und es gilt}\\ &|f(x,u,\nabla u)| \leqq a + b\,|\nabla u|^2 \qquad \text{mit } a,b \in \mathbb{R}. \end{aligned} \tag{19}$$

Das Problem der Differenzierbarkeit der Lösungen von (18) kann, wie wir zeigen werden, auf das entsprechende Problem für Lösungen gewisser Differentialgleichungen zurückgeführt werden. Wir wollen uns daher zunächst dem Regularitätsproblem solcher Differentialgleichungssysteme zuwenden. Es ist dabei von entscheidender Bedeutung, von gewissen Koeffizienten dieser Systeme nicht mehr als Meßbarkeit vorauszusetzen. Es gilt der

Satz 3. *Es sei $u \in C^0 \cap H_2^1(G)$ eine schwache Lösung des elliptischen Systems*

$$\frac{\partial}{\partial x^j}\left(a_{ij}(x,u) \frac{\partial u^k}{\partial x^i} \right) = a_k(x) + b_k(x)\,|\nabla u|^2 \qquad (k = 1, \ldots, N). \tag{20}$$

Dabei gelte

$$a_{ij} \in C^1(G \times \mathbb{R}^N),$$

$$a_k,\ b_k \in L_\infty(G),$$

$$a_{ij}(x,v) \text{ positiv definit } \forall\,(x,v) \in G \times \mathbb{R}^N.$$

Dann ist $u \in C^{1+\alpha}(G)$ für alle $\alpha \in [0,1)$.

Dieser Satz kann mit Hilfe der a-priori-Abschätzungen bewiesen werden, die Ladyzhenskaya-Uraltseva ([11], Chap. 8) für C^2-Lösungen von Systemen des Typs (20) abgeleitet haben. Für den Fall $n = 2$ wurde ein einfacherer Spezialfall von Satz 3 in [9] und [18] bewiesen. In einem Anhang am Schluß der Arbeit wollen wir einen Beweis von Satz 3 skizzieren, welcher von der sehr schwierigen Theorie von Ladyzhenskaya und Uraltseva im wesentlichen unabhängig ist, und, wie wir glauben, einen relativ einfachen Zugang zu Satz 3 darstellt. Es sei jedoch nicht verschwiegen, daß die Theorie von Ladyzhenskaya-Uraltseva ein etwas stärkeres Resultat als Satz 3 liefern würde; für unsere Zwecke ist dies aber ohne Bedeutung.

Zur Abkürzung setzen wir

$$B_r^n(x) := \{y \in \mathbb{R}^n : |y-x| < r\},$$
$$B_r^n := B_r^n(0)$$

und beginnen dann mit einem Lemma, das für $n=2$ von Morrey bewiesen wurde ([15], Lemma 5.4.1). Die Übertragung auf beliebiges n ist jedoch trivial, und kann daher dem Leser überlassen werden.

Lemma 5. *Es sei G ein beschränktes Gebiet des $\mathbb{R}^n$, $q \in L_1(G)$, und es gebe positive Zahlen β und C, so daß für alle $x \in G$ und $r > 0$ die Abschätzung*

$$\|q \mid G \cap B_r^n(x)\|_{0,1} \leqq C r^{n-2+2\beta}$$

erfüllt ist.

Dann existiert eine Konstante C_1, so daß für alle $v \in C_0^\infty(G)$ die Ungleichung

$$\left|\int q\, v\, dx\right| \leqq C_1 \left(\int |\nabla v|^2\, dx\right)^{\frac{1}{2}}$$

gilt.

Das folgende Lemma, welches die Einführung eines speziellen Koordinatensystems beschreibt, können wir ebenfalls ohne Beweis anführen, da es im wesentlichen aus der Riemannschen Geometrie bekannt ist ([6], S. 42).

Lemma 6. *Es sei (A_{kl}) eine C^2-Abbildung von $B_r^n \times B_r^N$ in die Menge der symmetrischen, nichtsingulären $N \times N$-Matrizen. Ferner sei F eine reelle Funktion der Klasse $C^3(B_r^n \times B_r^N)$ mit*

$$F(0,0) = 0$$

und

$$\nabla_y F(0,y)|_{y=0} \neq 0.$$

Dann gibt es ein s, $0 < s < r$, und eine C^2-Abbildung g von $B_s^n \times B_s^N$ in $\mathbb{R}^N$ mit den folgenden Eigenschaften:

(I) *Für jedes $x \in B_s^n$ ist $g(x,\cdot)$ ein Diffeomorphismus von B_s^N auf eine Umgebung der $0 \in \mathbb{R}^N$.*

(II) *Es gilt $g^N(x,y) = F(x,y)$.*

(III) *Setzt man $h(x,\cdot) := g(x,\cdot)^{-1}$, so gelten für $m = 1, \dots, N-1$ die Gleichungen*

$$A_{kl}(x, h(x,y)) \frac{\partial h^k}{\partial y^m}(x,y) \frac{\partial h^l}{\partial y^N}(x,y) = 0. \tag{21}$$

Wir sind jetzt in der Lage, das Hauptergebnis dieses Paragraphen zu formulieren und zu beweisen.

Satz 4. *Es sei* $u \in C^{\beta}(\bar{G}) \cap H_2^1(G)$ *mit* $0<\beta<1$ *und es gelte*

$$F(x, u(x)) \geqq 0 \qquad \forall x \in G$$

mit einer Funktion $F \in C^3(G \times \mathbb{R}^N)$, *welche der Bedingung*

$$\nabla_y F(x, y) \neq 0 \qquad \text{falls } F(x, y)=0$$

genügt. Ist dann u eine Lösung der Variationsungleichung (18), *wobei alle Bedingungen* (19) *erfüllt sind, so gilt* $u \in C^{1+\alpha}(G)$ *für alle* $\alpha \in [0, 1)$.

Beweis. (I) Es sei $x_0 \in G$ mit $F(x_0, u(x_0))>0$. Dann existiert ein $\varepsilon>0$, so daß auch

$$F(x, u(x)) \geqq \varepsilon \qquad \forall x \in B_\varepsilon^n(x_0)$$

gilt. Hieraus sieht man, daß für beliebiges $w \in C_0^0 \cap H_2^1(G)$ mit Träger $(w) \subset B_\varepsilon^n(x_0)$ durch

$$u[t]=u+t\,w \qquad (0 \leqq t<T)$$

eine stetig differenzierbare Kurve in $\mathscr{K}_u(F)$ gegeben ist, falls nur T hinreichend klein gewählt ist. Aus (18) folgt daher die Beziehung

$$\int_{B_\varepsilon^n(x_0)} \left(A_{kl}(x, u)\, a^{ij}(x, u) \frac{\partial u^k}{\partial x^i} \frac{\partial w^l}{\partial x^j}+f_k(x, u, \nabla u)\, w^k\right) dx=0 \tag{22}$$

für alle $w \in C_0^0 \cap H_2^1(B_\varepsilon^n(x_0))$. Mit der Substitution

$$w^l(x)=C_m^l(x, u(x))\, v^m(x),$$

wobei (C_m^l) die zu (A_{kl}) inverse Matrix und v ein beliebiges Element aus $C_0^0 \cap H_2^1(B^n(x_0))$ ist, erhält man aus (22) die Gleichung

$$\int \left(a^{ij}(x, u) \frac{\partial u^k}{\partial x^i} \frac{\partial v^k}{\partial x^j}+\bar{f}_k(x, u, \nabla u)\, v^k\right) dx=0 . \tag{23}$$

Hierbei ist $\bar{f}$ eine Vektorfunktion mit den selben Eigenschaften, wie sie f besitzt. Aus (23) kann nun vermöge Satz 3 die behauptete Regularität von u in $B_\varepsilon^n(x_0)$ gefolgert werden.

(II) Es sei nun $x_0 \in G$ mit

$$F(x_0, u(x_0))=0 .$$

Wir können ohne Einschränkung $x_0=0$ annehmen. Es seien dann g und h die in Lemma 6 beschriebenen, zu F und (A_{kl}) konstruierten Abbildungen. Die Funktion

$$U(x):=g(x, u(x))$$

ist dann für $x \in B_r^n$ wohldefiniert und Element von $H_2^1(B_r^n)$, falls nur r hinreichend klein gewählt ist. Ferner gilt

$$U^N(x) = F(x, u(x)) \geqq 0.$$

Es sei nun $V \in C^0 \cap H_2^1(B_r^n)$ mit

$$V = U \quad \text{auf } \partial B_r^n$$

und

$$V^N \geqq 0 \quad \text{in } B_r^n.$$

Für $t \in [0, T]$, $0 < T \leqq 1$, setzen wir dann

$$u[t](x) := \begin{cases} h(x, U(x) + t(V(x) - U(x))) & \text{falls } x \in B_r^n \\ u(x) & \text{sonst.} \end{cases}$$

Falls T hinreichend klein gewählt ist, ist $u[t]$ wohldefiniert und es gilt

$$F(x, u[t](x)) = U^N(x) + t(V^N(x) - U^N(x)) \geqq 0.$$

Somit ist $u[\cdot]$ eine stetig differenzierbare Kurve in $\mathscr{K}_u(F)$ mit Anfangspunkt u. Gemäß unserer Voraussetzung gilt dann für

$$w := \left.\frac{du[t]}{dt}\right|_{t=0}$$

die Relation (18). Für $x \in B_r^n$ ist nun

$$\begin{aligned} u(x) &= h(x, U(x)), \\ \frac{\partial u}{\partial x^j} &= \frac{\partial h}{\partial U^i}(x, U)\frac{\partial U^i}{\partial x^j} + \frac{\partial h}{\partial x^j}(x, U), \\ w(x) &= \frac{\partial h}{\partial U^i}(x, U)(V^i - U^i), \end{aligned} \tag{24}$$

$$\frac{\partial w}{\partial x^j} = \frac{\partial h}{\partial U^i}\frac{\partial (V^i - U^i)}{\partial x^j} + \left(\frac{\partial^2 h}{\partial U^i \partial U^k}\frac{\partial U^k}{\partial x^j} + \frac{\partial^2 h}{\partial U^i \partial x^j}\right)(V^i - U^i).$$

Einsetzen dieser Gleichungen in (18) führt auf einen Ausdruck der Gestalt

$$\begin{aligned} \int_{B_r^n} \Big(& \bar{A}_{kl}(x, U)\, \bar{a}^{ij}(x, U) \frac{\partial U^k}{\partial x^i}\frac{\partial (V^l - U^l)}{\partial x^j} \\ & + c_k(x, U, \nabla U)(V^k - U^k)\Big)\, dx \geqq 0 \end{aligned} \tag{25}$$

mit

$$\bar{A}_{kl}(x, U) = A_{mp}(x, h(x, U)) \frac{\partial h^m}{\partial U^k}(x, U) \frac{\partial h^p}{\partial U^l}(x, U),$$

$$\bar{a}^{ij}(x, U) = a^{ij}(x, h(x, U)),$$

und Funktionen c_k, welche von höchstens quadratischem Wachstum in ∇U sind.

Wir werden nun zeigen, daß sich aus (25) für U eine „Dirichletwachstumsbedingung" der Gestalt

$$\int_{B_s^n(x)} |\nabla U|^2 \, dy \leqq \text{const } s^{n-2+2\beta}, \qquad \left(x \in B_r^n, s \leqq \tfrac{1}{2}(r - |x|)\right) \tag{26}$$

ableiten läßt. Dazu sei $x_0 \in B_r^n$, $0 < s \leqq \frac{1}{2}(r - |x_0|)$, und η sei eine reellwertige Funktion der Klasse C^1 mit

$$\eta(y) = 1 \quad \text{für } y \in B_s^n(x_0),$$

$$\eta(y) = 0 \quad \text{für } y \notin B_{2s}^n(x_0),$$

$$0 \leqq \eta \leqq 1, \quad |\nabla \eta| \leqq s^{-1}.$$

Die Vektorfunktion

$$V := U + \eta^2 (U(x_0) - U)$$

ist dann in (25) als Testfunktion zugelassen, da V^N offenbar nichtnegativ ist. Man erhält daher aus (25) die Ungleichung

$$\begin{aligned} &\int \eta^2 \bar{A}_{kl} \bar{a}^{ij} \frac{\partial U^k}{\partial x^i} \frac{\partial U^l}{\partial x^j} \, dx \\ &\leqq \int \left(2\eta \, \bar{A}_{kl} \bar{a}^{ij} \frac{\partial U^k}{\partial x^i} (U^l(x_0) - U) \frac{\partial \eta}{\partial x^j} + c_k \eta^2 (U^k(x_0) - U^k) \right) dx. \end{aligned} \tag{27}$$

Durch eine Argumentation, welche analog zum Beweis von Lemma 1.2 in [11], Chap. 4, verläuft, gewinnt man aus (27) leicht die gewünschte Ungleichung (26).

Wenn man in (25) die Funktion

$$V = (U^1, \ldots, U^{N-1}, v)$$

einsetzt, wobei $v \in C^0(\overline{B_r^n}) \cap H_2^1(B_r^n)$ den Bedingungen

$$v \geqq 0, \qquad v = U^N \text{ auf } \partial B_r^n$$

genügt, so erhält man

$$\int_{B_r^n} \left(\bar{A}_{NN} \bar{a}^{ij} \frac{\partial U^N}{\partial x^i} \frac{\partial (v - U^N)}{\partial x^j} + c_N (v - U^N) \right) dx \geqq 0. \tag{28}$$

Definieren wir nun ein Maß Φ durch

$$\Phi(w) = \int c_N w\, dx,$$

so kann man, da c_N höchstens von quadratischem Wachstum in ∇U ist, aus (26) und Lemma 5 schließen, daß Φ ein stetiges Funktional auf $H_2^1(B_\rho^n)$ für alle $\rho < r$ ist. Aus Satz 2 und dem Theorem von Lebesgue-Nikodym ([1], Chap. V, § 5.5) erhalten wir dann die Existenz einer meßbaren Funktion $\bar{c}_N$ mit

$$|\bar{c}_N| \leqq |c_N|, \tag{29}$$

so daß

$$\int \left(\bar{A}_{NN} \bar{a}^{ij} \frac{\partial U^N}{\partial x^i} \frac{\partial v}{\partial x^j} + \bar{c}_N v \right) dx = 0$$

für alle $v \in C_0^\infty(B_r^n)$ gilt. Aus der Variationsungleichung (25) haben wir somit die Gleichung

$$\int_{B_r^n} \left(\bar{A}_{kl} \bar{a}^{ij} \frac{\partial U^k}{\partial x^i} \frac{\partial V^l}{\partial x^j} + \bar{c}_k V^k \right) dx = 0 \qquad \forall V \in C_0^\infty$$

gewonnen, wobei

$$\bar{c} = (c_1, \ldots, c_{N-1}, \bar{c}_N)$$

gesetzt ist. Wegen (29) gilt eine Abschätzung

$$|\bar{c}| \leqq A + B |\nabla U|^2, \qquad (A, B \in \mathbb{R}),$$

und wir können wie im ersten Teil des Beweises schließen, daß U zu $C^{1+\alpha}(B_r^n)$ gehört für alle $\alpha \in [0, 1)$. Für u ergibt sich dann die gleiche Regularität, womit Satz 4 bewiesen ist.

§ 4. Die Hölderstetigkeit im Fall $n = 2$

In diesem Paragraphen zeigen wir die Hölderstetigkeit der Lösungen zweidimensionaler Variationsprobleme von dem in § 1 betrachteten Typ. Damit wird im Fall $n = 2$ die Lücke geschlossen, die noch zwischen der Existenztheorie (Satz 1) und der Regularitätstheorie (Satz 4) klafft. Die Methode unseres Beweises läuft grob gesprochen darauf hinaus, ein Resultat von Morrey über das Dirichletproblem für Funktionen mit Werten in einem Riemannschen Raum ([15], Theorem 9.4.2) auf den Fall des Riemannschen Raumes mit Rand zu übertragen.

Wir benützen im folgenden die Abkürzung

$$D(u) = \int |\nabla u|^2\, dx$$

und schreiben Punkte $z \in \mathbb{R}^N$ in der Form

$$z = (\hat{z}, z^N) \quad \text{mit} \quad \hat{z} = (z^1, \ldots, z^{N-1}).$$

Wir benötigen die folgende

Definition. Es sei G ein Gebiet des $\mathbb{R}^2$ und ρ, R seien positive, λ und L nichtnegative Zahlen. Dann gehört eine reellwertige, auf $G \times \mathbb{R}^N$ definierte Funktion genau dann zur Klasse $\mathrm{Lip}(G \times \mathbb{R}^N; \rho, R; \lambda, L)$ wenn folgendes gilt

(I) $F \in C^0(G \times \mathbb{R}^N)$.

(II) Zu jedem $(x_0, z_0) \in G \times \mathbb{R}^N$ mit $F(x_0, z_0) = 0$ existiert eine Kongruenztransformation T des $\mathbb{R}^N$ und eine in $B_\rho^2(x_0) \times B_R^{N-1}(\hat{z}_0)$ erklärte, reellwertige Funktion g mit

$$|g(x, y) - g(x', y')| \leqq \lambda\,|x - x'| + L\,|y - y'| \tag{30}$$

für alle $x, x' \in B_\rho^2(x_0)$ und $y, y' \in B_R^{N-1}(\hat{z}_0)$ und der weiteren Eigenschaft, daß für alle $x \in G \cap B_\rho^2(x_0)$ die Beziehungen

$$\{z \in B_R^N(z_0) \colon F(x, z) > 0\} = T\{z \in B_R^N \colon z^N > g(x, \hat{z})\},$$

$$\{z \in B_R^N(z_0) \colon F(x, z) = 0\} = T\{z \in B_R^N \colon z^N = g(x, \hat{z})\}$$

gelten.

Das folgende Lemma ist ein Analogon zu einem Lemma von Morrey ([15], Lemma 9.4.8).

Lemma 7. *Die Funktion*

$$v \colon \partial B_r^2 \to \mathbb{R}^N$$

sei als Funktion der Bogenlänge von ∂B_r^2 absolutstetig. Ferner gelte

$$\int_0^{2\pi} \left| \frac{d}{d\theta} v(r e^{i\theta}) \right|^2 d\theta < \infty$$

und

$$F(x, v(x)) \geqq 0,$$

wobei F zu $\mathrm{Lip}(B_s^2 \times \mathbb{R}^N; \rho, R; \lambda, L)$ mit $s > r$ gehört. Außerdem sei

$$r < \min\left\{\frac{1}{2}\rho, \frac{R}{8\lambda}\right\}$$

und

$$\operatorname{osc} v \leqq \varepsilon := \frac{R}{8(1+L)}. \tag{31}$$

Dann existiert $h_0 \in H_2^1(B_r^2)$ mit

$$\begin{aligned} F(x, h_0(x)) &\geqq 0 \qquad (x \in B_r^2), \\ h_0(x) &= v(x) \qquad (x \in \partial B_r^2), \end{aligned} \tag{32}$$

und

$$D(h_0) \leqq (1 + 4L^2) \int_0^{2\pi} \left| \frac{d}{d\theta} v(r e^{i\theta}) \right|^2 d\theta + 4\pi\lambda^2 r^2. \tag{33}$$

Beweis. Es sei $x_1 \in \partial B_r^2$ und $z_1 := v(x_1)$. Wir unterscheiden zwei Fälle.

1. Fall. Für alle $x \in B_r^2(x_0)$ und alle $z \in B_\varepsilon^N(z_1)$ gilt $F(x,z) \geqq 0$. Wählen wir dann h_0 als die harmonische Funktion mit den Randwerten von v, so gilt nach dem Maximumprinzip für subharmonische Funktionen

$$\sup |h_0 - z_1| \leqq \varepsilon,$$

so daß gemäß unserer Annahme die Ungleichung (32) erfüllt ist. Für eine harmonische Funktion h gilt jedoch bekanntlich ([15], Theorem 1.10.2)

$$D(h) \leqq \int_0^{2\pi} \left| \frac{d}{d\theta} h(r e^{i\theta}) \right|^2 d\theta, \tag{34}$$

so daß umsomehr (33) richtig ist.

2. Fall. Es gibt ein $(x,z) \in B_r^2 \times B_\varepsilon^N(z_1)$ mit $F(x,z) < 0$.

Wegen $F(x_1, z_1) \geqq 0$ muß es dann ein $(x_2, z_2) \in \overline{B_r^2} \times \overline{B_\varepsilon^N(z_1)}$ geben mit $F(x_2, z_2) = 0$. Gemäß unserer Voraussetzung über F können wir ohne Einschränkung annehmen, daß F in $B_r^2 \times B_R^N(z_2)$ die Form

$$F(x,z) = z^N - g(x, \hat{z}) \tag{35}$$

hat, wobei g der Ungleichung (30) genügt.

Es sei nun h die in B_r^2 harmonische Funktion mit den Randwerten von v, und g_1 sei die in B_r^2 harmonische Funktion mit den Randwerten von $g(\cdot, \hat{v})$. Wie in [20] definieren wir dann

$$h_0 := (\hat{h}, g(\cdot, \hat{h}) + h^N - g_1).$$

Das Maximumprinzip für harmonische Funktionen liefert

$$h_0^N(x) - g(x, \hat{h}_0(x)) = h^N(x) - g_1(x) \geqq 0.$$

Wegen (35) ist die Ungleichung (32) daher gezeigt, sobald wir noch

$$\sup |h_0 - z_2| < R \tag{36}$$

nachgewiesen haben. Dazu führen wir die C^0-Norm

$$|w|_0 := \sup |w| \quad .$$

ein und erhalten mit Hilfe von (30) und (31) die folgenden elementaren Abschätzungen:

$$|h_0 - z_2|_0 \leqq |\hat{h} - \hat{z}_2|_0 + |h_0^N - g(x_2, \hat{z}_2)|_0, \tag{37}$$

$$|\hat{h} - \hat{z}_2|_0 \leqq |v - z_2|_0 \leqq |v - z_1|_0 + |z_2 - z_1| \leqq 2\varepsilon, \tag{38}$$

$$\begin{aligned} |h_0^N - g(x_2, \hat{z}_2)|_0 &\leqq |h^N - g_1|_0 + |g(\cdot, \hat{h}) - g(x_2, \hat{z}_2)|_0 \\ &\leqq |v^N - g(\cdot, \hat{v})|_0 + 2\lambda r + L|\hat{h} - \hat{z}_2|_0 \\ &\leqq |v^N - z_2^N|_0 + |g(x_2, \hat{z}_2) - g(\cdot, v)|_0 + 2\lambda r + 2L\varepsilon \\ &\leqq 2\varepsilon + 4\lambda r + 4L\varepsilon. \end{aligned} \tag{39}$$

Einsetzen von (38) und (39) in (37) ergibt

$$|h_0 - z_2|_0 \leqq 4(1+L)\,\varepsilon + 4\lambda\, r,$$

so daß (36) aus den Voraussetzungen des Lemmas folgt. Es bleibt die Ungleichung (33) zu zeigen. Durch elementare Rechnung erhalten wir

$$\begin{aligned} D(h_0) &= D(h) + D(g(\cdot, h) - g_1) = D(h) + \int (\nabla(g(\cdot, h) - g_1))(\nabla g(\cdot, h))\,dx \\ &\leqq D(h) + D(g(\cdot, \hat{h}))^{\frac{1}{2}} (D(g(\cdot, \hat{h}))^{\frac{1}{2}} + D(g_1)^{\frac{1}{2}}). \end{aligned}$$

Nach dem Dirichletschen Prinzip ist

$$D(g_1) \leqq D(g(\cdot, \hat{h})),$$

so daß sich schließlich die folgende Abschätzung ergibt:

$$\begin{aligned} D(h_0) &\leqq D(h) + 2D(g(\cdot, \hat{h})) \\ &\leqq D(h) + 4 \int_{B_r^2} (\lambda^2 + L^2 |\nabla \hat{h}|^2)\,dx \\ &\leqq (1 + 4L^2)\,D(h) + 4\pi\,\lambda^2\,r^2. \end{aligned}$$

Die Behauptung des Lemmas folgt nun aus (34).

Mit Hilfe des vorhergehenden Lemmas können wir nun unser Resultat über die Hölderstetigkeit der Minimallösungen beweisen. Der Beweis stimmt, was die Stetigkeit im Innern des Gebiets betrifft, im wesentlichen mit dem Beweis eines entsprechenden Theorems von Morrey ([15], Theorem 9.4.2) überein, wobei unser Lemma 7 an die Stelle von Morreys Lemma 9.4.8 tritt. Der Beweis der Stetigkeit in Randpunkten erfordert allerdings einige Änderungen der Morreyschen Methode. Aus beweistechnischen Gründen müssen wir die Grundzüge des gesamten Beweises darstellen.

Satz 5. *Es sei $G \subset \mathbb{R}^2$ ein beschränktes Gebiet, dessen Rand aus endlich vielen disjunkten Jordankurven der Klasse C^1 besteht. Es sei $F \in C^0(G \times \mathbb{R}^N)$ und es gebe positive Zahlen ρ, R und nichtnegative Zahlen λ und L, so daß F zur Klasse* $\mathrm{Lip}(G \times \mathbb{R}^N; \rho, R; \lambda, L)$ *gehört. Ferner sei I ein Variationsintegral der Gestalt*

$$I(v) = \int_G f(x, v(x), \nabla v(x))\,dx,$$

dessen Integrand f zu $C^0(G \times \mathbb{R}^N \times \mathbb{R}^{2N})$ gehört und den Abschätzungen

$$m|p|^2 \leqq f(x, v, p) \leqq M_1 + M_2 |p|^2 \quad \forall (x, v, p) \in G \times \mathbb{R}^N \times \mathbb{R}^{2N} \tag{40}$$

mit positiven Zahlen m, M_1, und M_2 genügt.

Ist dann $u \in H_2^1(G)$ *mit*

$$F(x, u(x)) \geqq 0 \quad \textit{für fast alle} \quad x \in G$$

und gilt

$$I(u) \leqq I(v)$$

für alle $v \in H_2^1(G)$ *mit* $u - v \in \mathring{H}_2^1(G)$ *und*

$$F(x, v(x)) \geqq 0 \quad \textit{für fast alle} \quad x \in G,$$

so existiert ein $\alpha > 0$, *so daß* u *zu* $C^\alpha(G)$ *gehört. Ist darüber hinaus* $u|S$ *stetig, wobei* S *ein relativ offener Teil von* ∂G *ist, so gilt* $u \in C^0(G \cup S)$.

Beweis. Für $x \in G$ und $r > 0$ setzen wir

$$\varphi(x, r) := D(u | B_r^2(x) \cap G).$$

Wir werden zeigen, daß es ein $\alpha > 0$ und zu jedem Kompaktum $K \subset G$ positive Zahlen C und s gibt mit

$$\varphi(x, r) \leqq C r^{2\alpha} \quad \forall x \in K, \ r \leqq s. \tag{41}$$

Wir beginnen mit der Bemerkung, daß φ bei festem x als Funktion von r in dem Intervall $[0, \text{Abstand}(x, \partial G)]$ absolutstetig ist und für fast alle r aus diesem Intervall

$$r \frac{d}{dr} \varphi(x, r) \geqq \int_0^{2\pi} \left| \frac{\partial}{\partial \theta} u(x + r e^{i\theta}) \right|^2 d\theta$$

gilt. Für fast alle r ist daher $u | \partial B_r^2$ absolutstetig und genügt der Ungleichung

$$\operatorname{osc} u | \partial B_r^2(x) \leqq \left(\pi \int_0^{2\pi} \left| \frac{\partial}{\partial \theta} u(r e^{i\theta}) \right|^2 d\theta \right)^{\frac{1}{2}}.$$

$$\leqq \left(\pi r \frac{\partial}{\partial r} \varphi(x, r) \right)^{\frac{1}{2}}.$$

Ist daher $r \leqq s$, wobei

$$s := \min \left\{ \text{Abstand}(K, \partial G), \frac{\rho}{2}, \frac{R}{8\lambda} \right\},$$

und gilt

$$r \varphi'(x, r) \leqq \frac{R^2}{64\pi} (1 + L)^{-2},$$

so ergeben sich aus (40) und Lemma 7 die Ungleichungen

$$\begin{aligned} m\,\varphi(x,r) &\leqq I(u|B_r^2(x)) \leqq I(h_0) \leqq \pi M_1 r^2 + M_2 D(h_0) \\ &\leqq \pi(M_1+4M_2\lambda^2)\,r^2 + M_2(1+4L^2)\int_0^{2\pi}\left|\frac{\partial}{\partial\theta}u(x+r\,e^{i\theta})\right|^2 d\theta \\ &\leqq M_2(1+4L^2)\,r\,\varphi'(x,r)+\pi(M_1+4M_2\lambda^2)\,r^2. \end{aligned}$$

Ist hingegen

$$r\,\varphi'(x,r) > \frac{R^2}{64\pi}(1+L)^{-2},$$

so folgt trivial

$$\varphi(x,r) \leqq 64\pi R^{-2}(1+L)^2 D(u)\,r\,\varphi'(x,r),$$

so daß sich insgesamt für fast alle $r \leqq s$ die Abschätzung

$$\varphi(x,r) \leqq k^{-1} r\,\varphi'(x,r) + l\,r^2$$

ergibt mit

$$k^{-1} := \max\{m^{-1} M_2(1+4L^2),\ 64\pi R^{-2}(1+L)^2 D(u)\}$$

und

$$l := \pi m^{-1}(M_1+4M_2\lambda^2).$$

Hieraus erhält man durch Integration

$$\varphi(x,r) \leqq \left(\varphi(x,s)+\frac{k\,l}{2-k}s^2\right)\left(\frac{r}{s}\right)^k, \tag{42}$$

womit (41) gezeigt ist. Nach Morrey folgt aus (41) die Hölderstetigkeit von u in G mit dem Exponenten $\alpha = \frac{1}{2}k$.

Wir wenden uns nun dem Beweis der Stetigkeit von u in einem Randpunkt x_0 zu, unter der Annahme, daß die Randwerte von u in einer Umgebung von x_0 stetig sind. Es bedeutet keine Einschränkung, $x_0 = 0$ vorauszusetzen. Wegen unserer Annahmen über ∂G können wir außerdem voraussetzen, daß

$$B_1^2 \cap G = \{x = (x^1, x^2) \in B_1^2 : x^2 > 0\}$$

gilt. Man kann nun leicht eine Schar von injektiven Abbildungen

$$T_s\colon \overline{B_1^2} \to \overline{B_1^2 \cap G} \qquad (0 < s < 1)$$

konstruieren, welche die folgenden Eigenschaften besitzt:

(I) $T_s \in C^1(\overline{B_1^2})$.

(II) $T_s(0) = \dfrac{i}{2}$, $T_s(i) = 0$.

(III) $T_s(\{e^{i\theta}\colon 0 \leqq \theta \leqq \pi\}) = \left[-\dfrac{s}{2}, \dfrac{s}{2}\right]$. (43)

(IV) Es gibt eine Zahl $C \geqq 1$, so daß für die Funktionalmatrizen ∇T_s die Abschätzungen

$$C^{-1} s |y| \leqq |\nabla T_s y| \leqq C s |y| \qquad \forall\, y \in \mathbb{R}^2,\ 0 < s < 1$$

gelten.

Wir definieren

$$G_s := T_s(B_1^2),$$
$$u_s := u \circ T_s,$$
$$F_s(y, v) := F(T_s y, v).$$

Offenbar gilt dann

$$u_s \in H_2^1 \cap C^0(B_1^2),$$
$$F_s \in \operatorname{Lip}(B_1^2 \times \mathbb{R}^N;\ C^{-1} s^{-1} \rho, R;\ C s \lambda, L),$$
$$F_s(y, u_s(y)) \geqq 0 \quad \forall\, y \in B_1^2.$$

Weiterhin existiert eine stetige Funktion $f_s = f_s(y, v, p)$, welche den Abschätzungen

$$m\, C^{-4} |p|^2 \leqq f_s(y, v, p) \leqq C^2 M_1 s^2 + C^4 M_2 |p|^2$$

genügt und die Eigenschaft hat, daß die Gleichung

$$I(v) = I_s(v \circ T_s)$$

für alle $v \in H_s^1(G_s)$ gilt, wobei

$$I_s(v) := \int_{B_1^2} f_s(y, v(y), \nabla v(y))\, dy$$

gesetzt wurde. Es ist dann offensichtlich, daß die Ungleichung

$$I_s(u_s) \leqq I_s(v)$$

für alle $v \in H_2^1(B_1^2)$ mit

$$v | \partial B_1^2 = u | \partial B_1^2, \qquad F_s(y, v(y)) \geqq 0$$

erfüllt ist. Setzen wir daher

$$\varphi_s(r) := D(u_s | B_r^2),$$

so ergibt sich für

$$s \leqq \min\{(2C)^{-1} \rho, (8 C \lambda)^{-1} R\}$$

die zu (42) analoge Ungleichung

$$\varphi_s(r) \leqq (\varphi_s(1) + l(s))\, r^{k(s)} \qquad (0 \leqq r \leqq 1) \tag{44}$$

mit

$$k(s)^{-1} = \max\{m^{-1} C^8 M_2(1+4L^2), 64\pi R^{-2}(1+L)^2 D(u_s)\}$$

und

$$l(s) = \pi m^{-1} C^6 (M_1 + 4C^4 M_2 \lambda^2) s^2.$$

Nach Morrey ([15], Beweis von Theorem 4.3.2) folgt aus (44) die Abschätzung

$$\int_0^\pi |u_s(e^{i\theta}) - u_s(0)|\, d\theta \leqq (2\pi)^{\frac{1}{2}} \big(k(s)^{-1} + 1\big)\big(D(u_s) + l(s)\big)^{\frac{1}{2}}. \tag{45}$$

Nun gilt offenbar

$$D(u_s) \to 0, \qquad l(s) \to 0 \quad (s \to 0),$$

wohingegen $k(s)^{-1}$ bei $s \to 0$, beschränkt bleibt. Unter Beachtung von (43) erhalten wir daher aus (45) als endgültiges Resultat

$$\begin{aligned}\left|u\left(\frac{s}{2}i\right) - u(0)\right| &= |u_s(0) - u_s(i)| \\ &\leqq \frac{1}{\pi} \int_0^\pi |u_s(e^{i\theta}) - u_s(0)|\, d\theta + \operatorname{osc} u \left|\left[-\frac{s}{2}, \frac{s}{2}\right]\right. \to 0 \quad (s \to 0).\end{aligned}$$

Wir haben somit die Stetigkeit von u bei Annäherung an einen Punkt $x_0 \in \partial G$ in Richtung der Normalen von ∂G gezeigt, unter der Annahme, daß $u|\partial G$ in einer Umgebung von x_0 stetig ist. Ist nun S_0 eine kompakte Teilmenge von ∂G, so daß $u|\partial G$ noch in einer ∂G-Umgebung von S_0 stetig ist, so sieht man aus den erbrachten Abschätzungen ohne weiteres, daß diese normale Stetigkeit gleichmäßig für alle $x_0 \in S_0$ ist. Damit ist aber die Stetigkeit von u in $G \cup S_0$ bewiesen.

§ 5. Anwendungen auf das Plateauproblem

Wir setzen voraus, daß der Leser mit dem klassischen zweidimensionalen Plateauproblem [3, 15] vertraut ist. Wir wollen hier eine Verallgemeinerung dieses Problems behandeln, nämlich das „Plateauproblem bei Anwesenheit von Hindernissen". Was genau darunter zu verstehen ist, wird im folgenden erklärt.

Wir setzen zur Abkürzung

$$B := B_1^2,$$

$$\partial^+ B := \{x = (x^1, x^2) \in \partial B \colon x^2 > 0\},$$

$$\partial^- B := \{x \in \partial B \colon x^2 < 0\}.$$

Sodann definieren wir Klassen monotoner Parametertransformationen von ∂B und $\partial^+ B$, nämlich

$$M_1 := \{\mu\colon \partial B \to \partial B\colon \mu \text{ ist stetig};$$
$$\mu(e^{2\pi ik/3}) = e^{2\pi ik/3} \quad \text{für } k=1,2,3;$$
$$\arg \mu(e^{it}) \leqq \arg \mu(e^{is}) \text{ für } 0 \leqq t \leqq s < 2\pi\}$$

und

$$M_2 := \{\mu\colon \overline{\partial^+ B} \to \overline{\partial^+ B}\colon \mu \text{ ist stetig und surjektiv};$$
$$\mu(i) = i;\ \arg \mu(e^{it}) \leqq \arg \mu(e^{is}) \text{ für } 0 \leqq t \leqq s \leqq \pi\}.$$

Es sei nun Γ eine Jordankurve in $\mathbb{R}^N$, d.h.

$$\Gamma\colon \partial B \to \mathbb{R}^N$$

sei eine stetige, injektive Abbildung. Dann setzen wir

$$\mathscr{K}_1(\Gamma) := \{u \in H_2^1(B)\colon \exists\, \mu \in M_1\colon u \mid \partial B = \Gamma \circ \mu\}.$$

Die Elemente von $\mathscr{K}_1(\Gamma)$ können als Flächen interpretiert werden, welche von der Jordankurve Γ berandet sind. Wir werden nun eine weitere Klasse von Flächen definieren, deren Rand teils ein Jordanscher Kurvenbogen ist, teils auf einer vorgegebenen Punktmenge liegt (vgl. [3], Chap. VI). Genauer, es sei $A \subset \mathbb{R}^N$ eine abgeschlossene Menge und

$$\Lambda\colon \overline{\partial^+ B} \to \mathbb{R}^N$$

sei eine stetige injektive Abbildung mit

$$\Lambda(1) \in A, \qquad \Lambda(-1) \in A,$$
$$\Lambda(e^{it}) \notin A \qquad \forall\, t \in (0, \pi).$$

Wir definieren dann

$$\mathscr{K}_2(\Lambda, A) := \{u \in H_2^1(B)\colon u(x) \in A \text{ für fast alle}$$
$$x \in \partial^- B,\ \exists\, \mu \in M_2\colon u \mid \partial^+ B = \Lambda \circ \mu\}.$$

Es gilt dann das folgende

Lemma 8. *Unter den genannten Voraussetzungen sind $\mathscr{K}_1(\Gamma)$ und $\mathscr{K}_2(\Lambda, A)$ als Teilmengen von $H_2^1(B)$ schwach abgeschlossen.*

Beweis. Die Aussage des Lemmas bezüglich $\mathscr{K}_1(\Gamma)$ folgt aus Lemma 2 in [20], so daß die schwache Abgeschlossenheit von $\mathscr{K}_2(\Lambda, A)$ zu beweisen bleibt. Es sei also $u_k \in \mathscr{K}_2(\Lambda, A)$ $(k=1,2,\ldots)$ mit

$$u_k \rightharpoonup u \quad \text{in } H_2^1(B). \tag{46}$$

Es gilt dann ([15], Theorem 3.4.5)

$$u_k|\partial B \to u|\partial B \quad \text{in } L_2(\partial B) \tag{47}$$

und folglich

$$u_k' \to u \quad \text{fast überall auf } \partial B$$

für eine Teilfolge u_k'. Hieraus folgt $u(x) \in A$ für fast alle $x \in \partial^- B$, womit gezeigt ist, daß u auf $\partial^- B$ die geforderte Randbedingung erfüllt. Nun untersuchen wir die Randwerte von u auf $\partial^+ B$. Es gilt die Darstellung

$$u_k|\partial^+ B = \Lambda \circ \mu_k, \qquad \mu_k \in M_2.$$

Aus (46) folgt die Existenz einer Konstanten C mit

$$D(u_k) \leqq C \quad \forall k.$$

Hieraus kann man mit Hilfe des Lemmas von Courant-Lebesgue ([3], Lemma 3.1) und unter Benützung der Monotonie der μ_k schließen, daß die Abbildungen μ_k auf jedem Kompaktum von $\partial^+ B$ gleichgradig stetig sind. Es läßt sich daher eine Folge natürlicher Zahlen $k_l \to \infty$ und eine stetige Abbildung

$$\mu\colon \partial^+ B \to \partial^+ B$$

finden mit

$$\mu_{k_l} \to \mu \qquad (l \to \infty)$$

gleichmäßig in jedem Kompaktum von $\partial^+ B$. Daher ist insbesondere

$$\mu(i) = i \tag{48}$$

und μ ist monoton in dem oben definierten Sinn. Wegen seiner Monotonie läßt sich μ zu einer auf ganz $\overline{\partial^+ B}$ erklärten, stetigen Abbildung fortsetzen, welche wiederum mit μ bezeichnet werde. Wegen (47) ist offenbar

$$u|\partial^+ B = \Lambda \circ \mu$$

und wir sind fertig, sobald wir gezeigt haben, daß

$$\mu(1) = 1, \qquad \mu(-1) = -1$$

gilt. Dazu betrachten wir die auf ganz $\mathbb{R}^N$ definierte Funktion

$$d(x) := \text{Abstand}\,(x, A).$$

Die Funktion d ist auf $\mathbb{R}^N$ gleichmäßig lipschitzstetig, weshalb $d \circ u$ zu $H_2^1(B)$ gehört. Offenbar ist

$$d \circ u|\partial^- B = 0,$$

und man kann daher mit Hilfe des Lemmas von Courant-Lebesgue Folgen $x_k \in \partial^+ B$, $y_k \in \partial^+ B$ finden mit

$$x_k \to 1, \quad y_k \to -1 \quad (k \to \infty)$$

und

$$d \circ u(x_k) \to 0, \quad d \circ u(y_k) \to 0 \quad (k \to \infty).$$

Es gilt aber

$$u(x_k) \to \Lambda \circ \mu(1), \quad u(y_k) \to \Lambda \circ \mu(-1) \quad (k \to \infty),$$

so daß man

$$\Lambda \circ \mu(1) \in A, \quad \Lambda \circ \mu(-1) \in A$$

erhält. Wir hatten jedoch vorausgesetzt, daß $\Lambda(1)$ und $\Lambda(-1)$ die einzigen Punkte von $\Lambda(\overline{\partial^+ B})$ sind, welche zu A gehören. Wegen der Monotonie von μ können nur die folgenden Fälle auftreten:

$$\mu(1) = \mu(-1) = 1,$$

oder

$$\mu(1) = \mu(-1) = -1,$$

oder

$$\mu(1) = 1, \quad \mu(-1) = -1.$$

Die ersten beiden Möglichkeiten stehen jedoch im Widerspruch zu (48) und der Monotonie von μ. Das Lemma ist somit bewiesen.

Wir wollen nun Flächen betrachten, welche über ein Hindernis H gespannt sind. Dabei setzen wir voraus, daß H eine kompakte Teilmenge des $\mathbb{R}^N$ von der Gestalt

$$H = \{x \in \mathbb{R}^N : F(x) \leqq 0\}$$

ist, wobei F zu $C^3(\mathbb{R}^N)$ gehört und der Bedingung

$$\nabla F(x) \neq 0 \quad \text{falls } F(x) = 0$$

genügt. Es seien nun

$$\Gamma : \partial B \to \mathbb{R}^N$$

eine Jordankurve und

$$\Lambda : \overline{\partial^+ B} \to \mathbb{R}^N$$

ein Jordanbogen, so daß

$$\Gamma(x) \notin \mathring{H} \quad \forall x \in \partial B,$$

$$\Lambda(x) \notin H \quad x \in \partial^+ B,$$

$$\Lambda(1) \in \partial H, \quad \Lambda(-1) \in \partial H.$$

Wir setzen dann

$$\mathscr{H} := \{u \in H_2^1(B) \colon u(x) \notin \mathring{H} \text{ für fast alle } x\},$$

$$\mathscr{L}_1(\Gamma, H) := \mathscr{K}_1(\Gamma) \cap \mathscr{H},$$

$$\mathscr{L}_2(\Lambda, H) := \mathscr{K}_2(\Lambda, \partial H) \cap \mathscr{H}.$$

Es gilt dann der folgende

Satz 6. *Es seien alle genannten Voraussetzungen über Γ, Λ, und H erfüllt, ferner seien $\mathscr{L}_1(\Gamma, H)$ und $\mathscr{L}_2(\Lambda, H)$ nicht leer. Dann existieren $u_1 \in \mathscr{L}_1(\Gamma, H)$ und $u_2 \in \mathscr{L}_2(\Lambda, H)$ mit*

$$D(u_1) = \inf\{D(u) \colon u \in \mathscr{L}_1(\Gamma, H)\} \tag{49}$$

und

$$D(u_2) = \inf\{D(u) \colon u \in \mathscr{L}_2(\Lambda, H)\} \tag{50}$$

und es gilt

$$u_1 \in C^0(\bar{B}) \cap C^{1+\alpha}(B),$$

$$u_2 \in C^0(B \cup \partial^+ B \cup \partial^- B) \cap C^{1+\alpha}(B)$$

für alle $\alpha \in [0, 1)$.

Beweis. Die Existenz von u_1 bzw. u_2 mit (49) bzw. (50) folgt aus Satz 1 und Lemma 8. Aus Satz 5 ergibt sich $u_1, u_2 \in C^\beta(B)$ für ein $\beta \in (0, 1)$, sowie $u_1 \in C^0(\bar{B})$ und $u_2 \in C^0(B \cup \partial^+ B)$. Satz 4 liefert dann $u_1, u_2 \in C^{1+\alpha}(B)$ für alle $\alpha \in (0, 1)$. Es bleibt noch $u_2 \in C^0(B \cup \partial^- B)$ zu zeigen. Dies beweist man leicht durch Kombination von Lemma 7 mit einer Methode von Hildebrandt [10]. Die Einzelheiten seien dem Leser überlassen.

Anhang. Ein Beweis von Satz 3

Wir betrachten das System

$$\frac{\partial}{\partial x^i} a_{ij}(x, u) \frac{\partial u^k}{\partial x^j} = a^k + b^k |\nabla u|^2 \qquad (k = 1, \ldots, N) \tag{51}$$

und werden zeigen, daß schwache Lösungen von (51) der Klasse $C^0 \cap H_2^1$ zu $C^{1+\alpha}$ gehören für alle $\alpha \in [0, 1)$. Aus dem Beweis wird unmittelbar ersichtlich sein, daß es genügen würde, an Stelle der Stetigkeit hinreichende Kleinheit der Oszillation zu fordern. Die Schranke für die Oszillation berechnet sich dabei aus gewissen Schranken für die Koeffizienten von (51). Über diese Koeffizienten setzen wir das Folgende voraus: Die Matrix (a_{ij}) gehört zu $C^1(\mathbb{R}^n \times \mathbb{R}^N)$ und die Vektorfunktionen $a = (a^1, \ldots, a^N)$ und $b = (b^1, \ldots, b^N)$ sind meßbar. Ferner gibt es eine positive Zahl M, und nichtnegative Zahlen A, B, C, D, so daß die folgenden Abschätzungen gelten:

$$M^{-1}|y|^2 \leqq a_{ij} y^i y^j \leqq M|y|^2 \qquad \forall y \in \mathbb{R}^n$$

$$|a| \leqq A, \qquad |b| \leqq B \tag{52}$$

$$\left|\frac{\partial a_{ij}}{\partial x^k}\right| \leqq C, \qquad \frac{\partial a_{ij}}{\partial u^k} y^i z^j_k \leqq D|y||z| \quad \forall y, z.$$

Das Prinzip des Beweises ist dasselbe wie bei Lemma 2 in [9], die Einzelheiten sind allerdings komplizierter. Wir werden zunächst zeigen, daß das schwache Dirichletproblem zu (51) in der Kugel B_r^n eine Lösung $v \in H_2^1 \cap C^{1+\alpha}$ besitzt, falls nur r und die Oszillation der Randdaten hinreichend klein sind. Wenn wir dann als Randdaten $u|\partial B_r^n$ vorschreiben, wobei u eine stetige Lösung von (51) ist, und ein Eindeutigkeitsargument von Ladyshenskaya-Uraltseva benutzen, erhalten wir $u=v$, und der Regularitätssatz ist bewiesen. Der Existenzsatz für das Dirichletproblem beruht auf einer Reihe von a-priori-Abschätzungen für Lösungen von (51), welche im wesentlichen bekannt sind oder zumindest auf bekannte Resultate zurückgeführt werden können. Wir begnügen uns daher damit, an Stelle von vollständigen Beweisen einige Hinweise zu geben.

Die Räume C^k und $C^{k+\alpha}$ ($k \in \mathbb{N}$, $0<\alpha<1$) seien auf die übliche Weise normiert. Ihre Normen bezeichnen wir mit $|\ |_k$ und $|\ |_{k+\alpha}$.

Wir beginnen mit einer einfachen Maximumabschätzung.

Lemma 9. *Es sei* $u \in C^2(B_1^n) \cap C^0(\overline{B_1^n})$ *eine Lösung des Systems* (51) *mit*

$$M(B+D)|u|_0 \leqq 1. \tag{53}$$

Dann kann man eine Zahl $K=K(B, C, D, M, n)$ *angeben, so daß die Abschätzung*

$$||u|^2|_0 \leqq ||u|^2|\partial B_1^n|_0 + AK \tag{54}$$

gilt.

Beweis. Man rechnet leicht nach, daß (53) die Ungleichung

$$a_{ij}\frac{\partial^2(|u|^2)}{\partial x^i \partial x^j} + \frac{\partial a_{ij}}{\partial x^j}\frac{\partial(|u|^2)}{\partial x^i} \geqq -2A|u|_0$$

zur Folge hat. Die Anwendung eines wohlbekannten Maximumprinzips ([4], S. 330) liefert (54).

Lemma 10. *Es sei* $u \in H_2^1(B_1^n)$ *eine schwache Lösung von* (51) *und es gebe ein* $\varepsilon \in (0, 1)$ *mit*

$$2MB\|u\|_{0,\infty} \leqq 1-\varepsilon. \tag{55}$$

Dann gilt für alle $v \in H_2^1(B_1^n)$ *mit* $u - v \in \mathring{H}_2^1(B_1^n)$ *die Ungleichung*

$$\int |\nabla u|^2 \, dx \leqq 2M^2 \varepsilon^{-2} \left(M^2 \int |\nabla v|^2 \, dx + 2\omega_n A^2\right). \tag{56}$$

Hierbei ist ω_n *das Volumen von* B_1^n.

Beweis. Nach dem Dirichletschen Prinzip genügt es offenbar, (56) für $v = h$ zu zeigen, wobei h die harmonische Funktion mit den Randwerten von u ist. In die schwache Form von (51) setze man $u - h$ als Testfunktion ein und beachte die Ungleichung

$$\|u - h\|_{0,\infty} \leqq 2 \, \|u\|_{0,\infty} .$$

Mit Hilfe von (55) ergibt sich dann die Behauptung.

Die für den Existenzsatz erforderlichen a-priori-Abschätzungen in der $C^{1+\alpha}$-Norm können aus einer früheren Arbeit des Autors entnommen werden [17]. Wie eine Durchsicht des Beweises von Hilfssatz 2 aus [17] zeigt, kann dieser Hilfssatz in der folgenden etwas schärferen Form ausgesprochen werden.

Lemma 11. *Es gibt eine Zahl* $\delta_0 = \delta_0(B, D, M, n) > 0$ *mit der folgenden Eigenschaft: Sind* $u, v \in C^2(\overline{B_1^n})$ *mit*

$$u \,|\, \partial B_1^n = v \,|\, \partial B_1^n$$

und ist u eine Lösung von (51) *mit*

$$|u|_0 \leqq (1 - \delta)\, \delta_0$$

für ein $\delta > 0$, *so gelten für alle* $r, \alpha \in (0, 1)$ *Abschätzungen der Form*

$$|u| \, B_r^n|_{1+\alpha} \leqq C_1(\delta, r, \alpha, A, B, C, D, M, n), \tag{57}$$

$$|u|_{1+\alpha} \leqq C_2(\delta, |v|_2, \alpha, A, B, C, D, M, n). \tag{58}$$

Nun sind alle Hilfsmittel bereitgestellt, um den angekündigten Existenzsatz zu beweisen.

Satz 6. *Es sei* $h \in H_2^1(B_1^n)$ *und es gebe positive Zahlen* ε *und* δ, *so daß die Ungleichung*

$$\begin{aligned} &\| \, |h|^2 \|_{0,\infty} + AK(B, C, D, M, n) \\ &\quad < \left(\min\{(1-\delta)\,\delta_0(B, D, M, n), (1-\varepsilon)(2MB)^{-1}, M^{-1}(B+D)^{-1}\}\right)^2 \end{aligned} \tag{59}$$

erfüllt ist, wobei K *und* δ_0 *die Konstanten aus Lemma* 9 *und Lemma* 11 *sind.*

Dann existiert eine schwache Lösung $u \in H_2^1(B_1^n)$ *des Systems* (51), *welche der Randbedingung*

$$u - h \in \mathring{H}_2^1(B_1^n)$$

und der Ungleichung

$$(\|u\|_{0,\infty})^2 \leqq (\|h\|_{0,\infty})^2 + AK(B, C, D, M, n) \tag{60}$$

genügt und darüber hinaus zu $C^{1+\alpha}$ *gehört für alle* $\alpha \in [0, 1)$.

Beweis. Der Beweis verläuft nach dem selben Prinzip wie der Beweis von Satz 4 in [8]. Wir nehmen zunächst zusätzlich an, daß die Funktion h und die Koeffizienten des Systems (51) im Abschluß ihrer jeweiligen Definitionsgebiete unendlich oft differenzierbar sind. Wir wählen dann ein $\beta \in (0, 1)$ und bemerken, daß für jedes $v \in C^{1+\beta}(\overline{B_1^n})$ das lineare Randwertproblem

$$\frac{\partial}{\partial x^j} a_{ij}(x, v) \frac{\partial u}{\partial x^i} = a(x) + b(x) |\nabla v|^2,$$

$$u = h \quad \text{auf } \partial B_1^n \tag{61}$$

eine eindeutig bestimmte Lösung $u = Tv \in C^{2+\beta}(\overline{B_1^n})$ besitzt. Die Abbildung T ist dabei eine beschränkte Abbildung von $C^{1+\beta}(\overline{B_1^n})$ in $C^{2+\beta}(\overline{B_1^n})$ und eine vollstetige Abbildung von $C^{1+\beta}(\overline{B_1^n})$ in sich ([11], Chap. 3). Die $C^{2+\beta}$-Lösungen von (51), welche der Randbedingung (61) genügen, sind offenbar genau die Fixpunkte von T. Wir definieren nun das Gebiet $\Omega \subset C^{1+\beta}(\overline{B_1^n})$ als die Menge derjenigen $v \in C^{1+\beta}(\overline{B_1^n})$, welche den Ungleichungen

$$|v|_0 \leqq (M(B+D))^{-1},$$

$$|v|_{1+\beta} \leqq C_2(\delta, |h|_2, \beta, A, B, C, D, M, n) + 1$$

genügen, wobei C_2 die Konstante aus (58) ist. Aus Lemma 9, der Bedingung (59), und aus Lemma 11 folgt dann, daß der Operator $\mathrm{Id} - sT$ für $s \in [0, 1]$ keine Nullstellen auf $\partial\Omega$ besitzen kann. Der Abbildungsgrad von $\mathrm{Id} - T|\Omega$ hat daher den Wert 1, woraus unmittelbar folgt, daß T einen Fixpunkt $u \in \Omega$ hat. Die Ungleichung (60) erhält man aus Lemma 9. Der Satz ist damit unter den obigen zusätzlichen Differenzierbarkeitsannahmen bewiesen. Der allgemeine Fall ergibt sich mit Hilfe der Lemmata 9, 10, und 11 durch Approximation.

Die folgende Eindeutigkeitsaussage wird mit einer Methode von Ladyzhenskaya-Uraltseva bewiesen ([11], Chap. 4, Theorem 2.1).

Lemma 12. *Es seien* $u_1, u_2 \in H_2^1(B_1^n)$ *schwache Lösungen von* (51) *mit* $u_1 - u_2 \in \mathring{H}_2^1(B_1^n)$ *und es gelte*

$$2MB\, \|u_k\|_{0,\infty} \leqq 1$$

und

$$2M^2(\sqrt{2}B + D)^2 (8M^4(\|u_1\|_{0,\infty}^2 + \|u_2\|_{0,\infty}^2) + AM(\|u_1\|_{0,\infty} + \|u_2\|_{0,\infty})) < 1.$$

Dann ist $u_1 = u_2$.

Beweis von Satz 3. Es sei $u \in H_2^1 \cap C^0(G)$ eine schwache Lösung von (51) und x_0 sei ein beliebiger Punkt von G. Für $r>0$ setzen wir

$$u_r(x) := u(x_0 + r\,x) - u(x_0).$$

Wenn r hinreichend klein ist, so ist u_r in B_1^n definiert und schwache Lösung eines Systems

$$\frac{\partial}{\partial x^j}\left(a_{ij}^r(x, u_r)\frac{\partial u_r}{\partial x^i}\right) = a_r + b_r\,|\nabla u_r|^2, \tag{62}$$

dessen Koeffizienten den Ungleichungen (52) genügen, wobei allerdings die Konstante A durch $r^2 A$ ersetzt werden darf. Nach Satz 6 existiert daher eine schwache Lösung $v_r \in H_2^1 \cap C^{1+\alpha}(B_1^n)$ des Systems (62) mit den Randwerten von u_r, falls nur r genügend klein ist. Da v_r einer (60) entsprechenden Abschätzung genügt, wird — nach gegebenenfalls nochmaliger Verkleinerung von r — Lemma 12 anwendbar und man erhält $u_r = v_r$, womit Satz 3 bewiesen ist.

Literatur

1. Bourbaki, N.: Eléments de Mathématique, Livre VI, Intégration (seconde edition). Paris: Hermann 1965.
2. Browder, F. E.: Problèmes non-linéaires. Montréal: Les Presses de l'Université 1966.
3. Courant, R.: Dirichlets principle, conformal mapping, and minimal surfaces. New York: Interscience 1950.
4. Courant, R., Hilbert, D.: Methods of mathematical physics, vol. II. New York-London: Interscience 1962.
5. DeGiorgi, E.: Un esempio di estremali discontinue per un problema variazionale di tipo ellitico. Boll. Un. Mat. Ital., Ser. IV, Anno I, 135–137 (1968).
6. Eisenhart, L. P.: Riemannian geometry. Princeton: University Press 1949.
7. Giusti, E.: Superfici minime cartesiane con ostacoli discontinui. Arch. Rat. Mech. Analysis **40**, 251–267 (1971).
8. Heinz, E.: Über die Existenz einer Fläche konstanter mittlerer Krümmung bei vorgegebener Berandung. Math. Ann. **127**, 258–287 (1954).
9. Heinz, E., Tomi, F.: Zu einem Satz von Hildebrandt über das Randverhalten von Minimalflächen. Math. Z. **111**, 372–386 (1969).
10. Hildebrandt, S.: Ein einfacher Beweis für die Regularität der Lösungen gewisser zweidimensionaler Variationsprobleme unter freien Randbedingungen. Erscheint demnächst in Math. Ann.
11. Ladyzhenskaya, O. A., Uraltseva, N. N.: Linear and quasilinear elliptic equations (Übersetzung aus dem Russischen). New York-London: Academic Press 1968.
12. Lewy, H., Stampacchia, G.: On the regularity of the solution of a variational inequality. Commun. Pure Appl. Math. **22**, 153–188 (1968).
13. Lewy, H., Stampacchia, G.: On existence and smoothness of solutions of some noncoercive variational inequalities. Arch. Rat. Mech. Analysis **41**, 241–253 (1971).
14. Miranda, M.: Frontiere minimali con ostacoli. Erscheint demnächst in Ann. Univ. Ferrara Sez. VII.
15. Morrey, C. B.: Multiple integrals in the calculus of variations. Berlin-Heidelberg-New York: Springer 1966.

16. Nitsche, J. C. C.: Variational problems with inequalities as boundary conditions or how to fashion a cheap hat for Giacomettis brother. Arch. Rat. Mech. Analysis **35**, 82 – 113 (1969).
17. Tomi, F.: Über semilineare elliptische Differentialgleichungen zweiter Ordnung. Math. Z. **111**, 350 – 366 (1969).
18. Tomi, F.: Ein einfacher Beweis eines Regularitätssatzes für schwache Lösungen gewisser elliptischer Systeme. Math. Z. **112**, 214 – 218 (1969).
19. Tomi, F.: Ein teilweise freies Randwertproblem für Flächen vorgeschriebener mittlerer Krümmung. Math. Z. **115**, 104 – 112 (1970).
20. Tomi, F.: Minimal surfaces and surfaces of prescribed mean curvature spanned over obstacles. Math. Ann. **190**, 248 – 264 (1971).

Dr. F. Tomi
Mathematisches Institut der Universität
D-4400 Münster i. Westf., Roxelerstraße 64
Deutschland

(Eingegangen am 18. Februar 1972)

Math. Z. 128, 75–83 (1972)

Infinite Matrices and Almost-Convergence

J. Peter Duran

§ 1. Introduction

1.1. In this paper we prove two theorems concerning the relationship between almost-convergence and infinite matrices. Theorem 1 gives necessary and sufficient conditions for an infinite matrix to transform each bounded sequence into an almost-convergent sequence. The infinite matrices which transform each almost-convergent sequence into an almost-convergent sequence are characterized by Theorem 2. As will become apparent, the two theorems are closely related. As an application of our results we characterize the multipliers of certain spaces of almost-convergent sequences.

The author wishes to thank Professor R.A. Raimi for several helpful discussions.

§ 2. Background

2.1. We denote by ℓ^∞ the Banach space of all bounded sequences $x=(x_0, x_1, \ldots)$ of real numbers. The shift operator, S, on ℓ^∞ is defined by $(Sx)_n = x_{n+1}$. A Banach limit φ is an element of $(\ell^\infty)^*$ satisfying $\varphi \geqq 0$, $\varphi(u)=1$ where $u_n=1, n=0, 1, \ldots$, and $\varphi(Sx-x)=0$ for all $x \in \ell^\infty$. It is well-known that the set of all Banach limits is a non-empty, convex, compact subset of $(\ell^\infty)^*$.

2.2. **Definition.** $x \in \ell^\infty$ is *almost-convergent* to $a \in \mathbb{R}$ if $\varphi(x)=a$ for each Banach limit φ. This is written $F\text{-}\lim x = a$. ($\mathbb{R}$ is the set of real numbers.)

We denote by F the space of all almost-convergent sequences and by F_0 the subspace of F consisting of all sequences almost-convergent to 0. F is a closed, non-separable subspace of ℓ^∞, invariant under S and $F = F_0 \oplus \mathbb{R}\, u$ [4].

2.3. Let $S_n = \dfrac{1}{n+1} \sum_{i=0}^{n} S^i$ where S^i is the composition of the shift operator with itself i times. We shall need the following known characterization of F:

The following conditions on an element $x \in \ell^\infty$ *are equivalent:*

2.3.1. $F\text{-}\lim x = a$;

2.3.2. $S_n x$ *converges to* $a u$ *in the norm of* ℓ^∞;

2.3.3. $x \in a u +$ *closed span of* $\{S y - y \mid y \in \ell^\infty\}$. *Condition* 2 *can be restated as:*

2.3.4. $\lim\limits_n \dfrac{x_k + \cdots + x_{k+n}}{n+1} = a$ *uniformly in* k.

Lorentz [4] first proved the equivalence of 1 and 2, while the equivalence of 1 and 3 can be found in [8].

2.4. Let c denote the subspace of ℓ^∞ consisting of all convergent sequences and c_0 the subspace of all sequences converging to 0. A Schur matrix A is a matrix satisfying $A(\ell^\infty) \subseteq c$. Although we shall not need the full force of Schur's theorem, we shall need the following easy corollaries of it (see [6, pp. 17–22] for a good discussion of Schur matrices):

2.4.1. *If* A *is a Schur matrix, then* $\lim\limits_n a_{m,n} = \alpha_n$ *exists for each* n *and if* $x \in \ell^\infty$, $\lim\limits_m (A x)_m = \sum\limits_{n=0}^{\infty} \alpha_n x_n$.

2.4.2. $A(\ell^\infty) \subseteq c_0$ *if and only if* $\lim\limits_m \sum\limits_{n=0}^{\infty} |a_{m,n}| = 0$.

§ 3. Almost Schur Matrices

3.1. **Definition.** A matrix is called *almost Schur* provided that $A(\ell^\infty) \subseteq F$.

It is the purpose of this section to characterize the almost Schur matrices. In their paper [2], Eizen and Laush purport to give such a characterization. Their conditions differ from ours in that no absolute value signs appear in 3.3.3 below. The proof that their conditions imply that A is almost Schur is insufficient. Moreover, if our (stronger) conditions were substituted for theirs, their proof of the converse would then become inadequate.

3.2. We shall need the following generalization of a lemma of Schaefer [9]:

3.2.1. *Let* $\{A^{(m)}\}$ *be a sequence of infinite matrices. Then for each* $x \in \ell^\infty$, $\lim\limits_q (A^{(m)} x)_q = 0$ *uniformly in* m *if and only if* $\lim\limits_q \sum\limits_{n=0}^{\infty} |a_{q,n}^{(m)}| = 0$ *uniformly in* m.

Proof. Suppose that

$$\lim_q \sum_{n=0}^{\infty} |a_{q,n}^{(m)}| = 0$$

uniformly in m. Let $x\in\ell^\infty$. Then

$$\lim_q |(A^{(m)}x)_q| = \lim_q \left|\sum_{n=0}^{\infty} a_{q,n}^{(m)} x_n\right| \leqq \|x\| \cdot \lim_q \sum_{n=0}^{\infty} |a_{q,n}^{(m)}| = 0$$

uniformly in m.

For the converse, assume that for each $x\in\ell^\infty$, $\lim_q (A^{(m)}x)_q = 0$ uniformly in m. Then by 2.4.2,

$$\lim_q \sum_{n=0}^{\infty} |a_{q,n}^{(m)}| = 0$$

for all m. If this limit were not uniform in m, there would be an $\alpha>0$ and increasing sequences (q_l) and (m_l) of natural numbers such that for each l,

$$\sum_{n=0}^{\infty} |a_{q_l,n}^{(m_l)}| \geqq \alpha .$$

Thus by 2.4.2, applied to the matrix $b_{l,n} = a_{q_l,n}^{(m_l)}$, there would be an $x\in\ell^\infty$ such that not

$$\lim_l \sum_{n=0}^{\infty} a_{q_l,n}^{(m_l)} x_n = 0 .$$

For this x, we would not have

$$\lim \sum_{n=0}^{\infty} a_{q,n}^{(m)} x = 0$$

uniformly in m.

Remark. 3.2.1 is clearly a direct generalization of and corollary to 2.4.2.

3.3. We are now in a position to characterize the almost Schur matrices.

Theorem 1. *A is almost Schur* (i.e. $A(\ell^\infty)\subseteq F$) *if and only if:*

3.3.1. $\sup_m \sum_{n=0}^{\infty} |a_{m,n}| \equiv \|A\| < +\infty$;

3.3.2. *F-$\lim a_{m,n} = \alpha_n$ exists for all n: and*

3.3.3. $\lim_q \sum_{n=0}^{\infty} \left| \frac{1}{q+1} \sum_{i=0}^{\infty} a_{m+i,n} - \alpha_n \right| = 0$ *uniformly in m.*

Moreover, if A satisfies 3.3.1–3.3.3, *then for each $x\in\ell^\infty$, F-$\lim Ax = \sum_{n=0}^{\infty} \alpha_n x_n$.*

Proof. Suppose first that A is almost Schur. By Theorem 1.3.2 of [6], $\|A\| < +\infty$. Fix n and set $e^{(n)}{}_m = \delta_{m,n}$, the Kronecker delta. Let $\alpha_n = F$-$\lim A e^{(n)}$. Then since $(A e^{(n)})_m = a_{m,n}$, we have $\alpha_n = F$-$\lim a_{m,n}$.

Now define a sequence $\{B^{(m)}\}$ of infinite matrices as follows:

$$b_{q,n}^{(m)} = \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n}.$$

Then each $B^{(m)}$ is a Schur matrix. Indeed, if $x \in \ell^\infty$, $(B^{(m)}x)_q = (S_q A x)_m$ which converges uniformly in m as $q \to \infty$.

Since for each m and n, $\lim_q b_{q,n}^{(m)} = \alpha_n$, 2.4.1 implies that $\lim_q (S_q A x)_m = \lim_q (B^{(m)} x)_q = \sum \alpha_n x_n$. Further this limit must be uniform in m since $(S_q A x)_m$ converges uniformly in m. Thus $F\text{-}\lim Ax = \sum \alpha_n x_n$ for each x.

Now consider the sequence $\{C^{(m)}\}$ of matrices defined by

$$C_{q,n}^{(m)} = \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} - \alpha_n.$$

Since $(C^{(m)} x)_q = (S_q A x)_m - \sum \alpha_n x_n$, it follows that for each $x \in \ell^\infty$, we have $\lim_q (C^{(m)} x)_q = 0$ uniformly in m. Hence by the lemma, we have

$$\lim_q \sum_{n=0}^{\infty} \left| \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} - \alpha_n \right| = 0$$

uniformly in m.

For the converse, we assume that A satisfies 3.3.1–3.3.3. Then if $x \in \ell^\infty$,

$$|(S_q A x)_m - \sum \alpha_n x_n| \leqq \|x\| \sum_{n=0}^{\infty} \left| \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} - \alpha_n \right| \to 0$$

uniformly in m. Hence $F\text{-}\lim Ax = \sum \alpha_n x_n$ and A is almost Schur.

3.4. **Corollary.** *A matrix A transforms each bounded sequence into a sequence which is almost-convergent to 0 if and only if*

$$\lim_q \sum_{n=0}^{\infty} \left| \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} \right| = 0$$

uniformly in m and $\|A\| < \infty$.

Proof. The proof is immediate once we observe that for each n, $\alpha_n = F\text{-}\lim_m a_{m,n} = F\text{-}\lim A e^{(n)} = 0$.

3.5. There is one more corollary we would like to draw from Theorem 1. However, before we can do so, we need some facts due to King [3].

Definition. A matrix A is *almost conservative* if:

3.5.1. $\|A\| < +\infty$;

3.5.2. $F\text{-}\lim_m a_{m,n} = \alpha_n$ exists for each n; and

3.5.3. $F\text{-}\lim_m \sum_{n=0}^{\infty} a_{m,n} = \alpha$ exists.

A is almost regular if $\alpha_n = 0$ for each n in 3.5.2 and $\alpha = 1$ in 3.5.3.

3.6. The following facts are Theorems 3.1 and 3.2 respectively in [3]:

3.6.1. *A is almost conservative if and only if $A(c) \subseteq F$.*

3.6.2. *A is almost regular if and only if $A(c) \subseteq F$ and $\lim_m x_m = F\text{-}\lim x$ for each $x \in c$.*

These two theorems are generalizations of the Kojima-Schur Theorem and the Silverman-Toeplitz Theorem respectively. In the next section we shall generalize them still further.

3.7. In analogy to a well-known theorem (cf. [6, Corollary 1, p. 19]) we have the following:

Corollary (to Theorem 1). *No matrix is both almost Schur and almost regular.*

Proof. If A is almost regular, then $\alpha_n = 0$ for all n. Hence we would have the contradictory conditions that

$$\lim_q \sum_{n=0}^{\infty} \left| \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} \right| = 0$$

uniformly in m and

$$\lim_q \sum_{n=0}^{\infty} \frac{1}{q+1} \sum_{i=0}^{q} a_{m+i,n} = 1$$

uniformly in m.

The above corollary is also stated in [2].

3.8. We give an example of an almost Schur matrix which is not a Schur matrix. Let $\{a_n\}$ be a sequence of non-negative real numbers with the properties:

1) $a_n > 0$ for all n;
2) $\sum a_n < +\infty$; and
3) $\sum (-1)^n a_n = 0$.

Define a matrix A by: $a_{2m,n} = (-1)^n a_n$, $a_{2m+1,n} = (-1)^{n+1} a_n$. Then if $x \in \ell^\infty$, $(Ax)_{2m} = -(Ax)_{2m+1}$. Hence $F\text{-}\lim Ax = 0$ for each $x \in \ell^\infty$ and A is almost Schur. However, if $x = (1, -1, 1, \ldots)$ then $(Ax)_{2m} = \sum a_n > 0$ and $(Ax)_{2m+1} = -\sum a_n < 0$, so $\lim Ax$ does not exist and A is not a Schur matrix.

§4. Matrices Leaving F Invariant

4.1. The next problem with which we shall be concerned is that of characterizing those matrices which transform almost-convergent sequences into almost-convergent sequences. A first step in this direction was taken by Schaefer [9]. He gave necessary and sufficient conditions for an almost regular matrix to transform each almost-convergent sequence into an almost-convergent sequence with the same F-limit. His theorem is an easy corollary of 4.4 below. As will become apparent, the solution of the general problem rests on the characterization of almost Schur matrices given in Theorem 1.

4.2. **Theorem 2.** *An infinite matrix A transforms each almost-convergent sequence into an almost-convergent sequence if and only if:*

4.2.1. $\|A\| < +\infty$;

4.2.2. $F\text{-}\lim_m \sum_{n=0}^{\infty} a_{m,n} = \alpha$ *exists; and*

4.2.3. *the matrix $A(S-I)$ is an almost Schur matrix (where S is the shift, and I is the identity).*

Remark. Since

$$A(S-I)_{m,n} = \begin{cases} a_{m,n-1} - a_{m,n} & \text{if } n = 1, 2, 3, \ldots \\ -a_{m,0} & \text{if } n = 0 \end{cases}$$

Theorem 1 allows us to restate condition 4.2.3 in the following equivalent form:

4.2.4. (a) $F\text{-}\lim_m a_{m,n} = \alpha_n$ exists for all n, and

(b) $\lim_q \sum_{n=1}^{\infty} \frac{1}{q+1} \left| \sum_{i=0}^{q} \{a_{m+i,n-1} - \alpha_{n-1} + \alpha_n - a_{m+i,n}\} \right| = 0$

uniformly in m.

Note also that conditions 4.2.1, 4.2.2 and 4.2.4 (a), state simply that A is almost conservative. Thus our conditions amount to the statement that A is almost conservative and 4.2.4 (b) holds.

Proof (of Theorem 2). Suppose first that 4.2.1, 4.2.2, and 4.2.3 are satisfied. If $x \in \ell^\infty$, then $A(Sx - x) \in F$ by 4.2.3. Since A is a bounded linear operator on ℓ^∞ by 4.2.1, it follows from 2.3.3 that $A(F_0) \subseteq F$. If $y \in F$, then $y = x + au$ where $x \in F_0$. Hence $Ay = Ax + aAu$. Since $(Au)_m = \sum_{n=0}^{\infty} a_{m,n}$, 4.2.2 shows that $Au \in F$. Since $Ax \in F$, we conclude that $Ay \in F$. Thus $A(F) \subseteq F$.

Assume, conversely, that $A(F) \subseteq F$. Since $c_0 \subseteq F$ and $F \subseteq \ell^\infty$, Theorem 1.3.2 of [6] shows that $\|A\| < \infty$. Since $u \in F$, there is a number α such that $\alpha = F\text{-}\lim_m Au = F\text{-}\lim_m \sum_{n=0}^{\infty} a_{m,n}$. Lastly, we note that by 2.3.3 $Sx - x \in F$ for each $x \in \ell^\infty$. Hence $A(Sx - x) \in F$ for each $x \in \ell^\infty$, i.e. $A(S-I)$ is almost Schur.

4.3. Using methods entirely analogous to those used in the proof of Theorem 2, we obtain the following theorem whose proof we omit:

Theorem 3. *A transforms F_0 into F_0 if and only if*

4.3.1. $\|A\| \neq +\infty$; *and*

4.3.2. *The matrix $A(S-I)$ transforms each bounded sequence into a sequence which is almost-convergent to* 0.

Remark. As in 4.2, condition 4.3.2 can be replaced by

4.3.3. (a) $F\text{-}\lim a_{m,n}=0$ for all n;

(b) $\lim_q \sum_{n=1}^{\infty} \frac{1}{q+1}\left|\sum_{i=0}^{q}\{a_{m+i,n-1}-a_{m+i,n}\}\right|=0$ uniformly in m.

4.4. We turn now to the question of consistency—i.e. we wish to describe those matrices A which map F into F in such a way that $F\text{-}\lim x = F\text{-}\lim A x$ for all $x\in F$. It is clear that a necessary condition for this to occur is that A be almost regular. This is also sufficient.

Theorem 4. *Suppose that A transforms F into F. Then $F\text{-}\lim x = F\text{-}\lim A x$ for each $x\in F$ if and only if A is almost regular.*

Proof. Suppose that $A(F)\subseteq F$ and A is almost regular. Then A satisfies conditions 4.2.1, 4.2.2, and 4.2.4. By the definition of almost regular, we have $\alpha=1$ in 4.2.2 and $\alpha_n=0$ for each n in 4.2.4(a). The latter fact implies, via 3.4, that $A(S-I)$ maps ℓ^∞ into F_0. Linearity and continuity of A now yield that $A(F_0)\subseteq F_0$. Moreover, we have

$$1=F\text{-}\lim_m \sum_n a_{m,n} = F\text{-}\lim A\,u.$$

Now if $x\in F$ and $F\text{-}\lim x=a$, then $x=y+a\,u$ where $y\in F_0$. Hence $F\text{-}\lim A\,x=F\text{-}\lim A\,y+a\cdot F\text{-}\lim u=a$.

The converse, as noted previously, is trivial.

4.5. As an immediate consequence of the foregoing, we have:

Corollary. *$A(F)\subseteq F$ and $F\text{-}\lim A\,x=F\text{-}\lim x$ for each $x\in F$ if and only if conditions* 4.2.1, 4.2.2, *and* 4.2.4 *are satisfied with $\alpha=1$ and $\alpha_n=0$ for each n.*

4.6. We conclude this section by giving an example of a matrix A which transforms F into F in a consistent fashion but which does not satisfy $A(c)\subseteq c$, i.e. which is not regular.

Define

$$a_{m,n}=\begin{cases}\dfrac{1}{m+1}[1+(-1)^m], & 0\leqq n\leqq m\\ 0, & n>m.\end{cases}$$

That the matrix A so defined has the asserted properties is a straightforward verification left to the reader.

§ 5. An Application-Multipliers

5.1. **Definition.** Let E be a closed subspace of ℓ^∞. the *multipliers* of E are defined to be the set

$$\mathcal{M}_E = \{x \in \ell^\infty \mid x\, y \in E \ \ \forall\, y \in E\}.$$

(Multiplication is coordinatewise.)

It is clear that $\mathcal{M}_E$ is a closed subalgebra of ℓ^∞ contained in E. Also, if $u \in E$, then $\mathcal{M}_E \subseteq E$. Thus $\mathcal{M}_F \subseteq F$. However, it is not clear that $\mathcal{M}_{F_0} \subseteq F_0$ and we shall see that this is in fact not the case. We remark that Lloyd [5] has shown that even when $u \in E$, $\mathcal{M}_E$ need not be a maximal closed subalgebra of ℓ^∞ contained in E. In fact, $\mathcal{M}_F$ fails to satisfy this property.

It is the purpose of this section to use Theorems 3 and 2 to characterize $\mathcal{M}_{F_0}$ and $\mathcal{M}_F$ respectively. The space $\mathcal{M}_F$ has already been characterized by Chou [1]. Our characterization of $\mathcal{M}_{F_0}$, however, seems to be new.

5.2. For $y \in \ell^\infty$, let $|y|$ denote the sequences defined by $|y|_n = |y_n|$. We have:

Theorem 5. *Let* $x \in \ell^\infty$. $x \in \mathcal{M}_{F_0}$ *if and only if* $|S\,x - x| \in F_0$.

Proof. Let M be the matrix such that $M\,y = x\,y$ for all $y \in \ell^\infty$, i.e. $M_{m,n} = x_m \delta_{m,n}$. Then clearly $x \in \mathcal{M}_{F_0}$ if and only if F_0 is invariant under M. Since M satisfies 4.3.1 and 4.3.3 (a) in a trivial way, $x \in \mathcal{M}_{F_0}$ precisely when M satisfies 4.3.3 (b). Now for M this latter condition becomes

$$\lim_q \frac{1}{q+1} \sum_{n=1}^{q} |x_{m+n-1} - x_{m+n}| = 0$$

uniformly in n. This is easily seen to be equivalent to

$$F\text{-}\lim |S\,x - x| = 0.$$

We remark that Petersen and Zame ([7, Theorem 3]) have shown that if $|S\,x - x| \in F_0$, then $x \in \mathcal{M}_{F_0}$. They did not, however, consider the reverse inclusion and their methods differ greatly from ours.

5.3. We give an example of a sequence $x \notin F_0$ but such that $x \in \mathcal{M}_{F_0}$. Let $x = (1, 0, 1, 1, 0, 0, 1, \ldots)$ where the blocks of 0's and 1's grow equally. The verification that x has the asserted properties is easy and is left to the reader.

5.4. The following theorem follows from Theorem 2 exactly as Theorem 5 follows from Theorem 3. (It is also a consequence of Theorem 5.) We therefore omit the proof.

Theorem 6. *Let* $x \in \ell^\infty$. $x \in \mathcal{M}_F$ *if and only if* $x \in F$ *and* $|S\,x - x| \in F_0$.

References

1. Chou, C.: The multipliers for the space of almost convergent sequences. To appear in Illinois J. Math.
2. Eizen, C., Laush, G.: Infinite matrices and almost convergence. Math. Japonicae **14**, 137–143 (1969).
3. King, J. P.: Almost summable sequences. Proc. Amer. Math. Soc. **17**, 1219–1225 (1966).
4. Lorentz, G. G.: A contribution to the theory of divergent sequences. Acta Math. **80**, 167–190 (1948).
5. Lloyd, S. P.: Subalgebras in a subspace of $C(X)$. Illinois J. Math. **14**, 259–267 (1970).
6. Petersen, G. M.: Regular matrix transformations. London: McGraw-Hill 1966.
7. Petersen, G. M., Zame, A.: Summability properties for the distribution of sequences. Monatsh. Math. **73**, 147–158 (1969).
8. Raimi, R. A.: On Banach's generalized limits. Duke Math. J. **26**, 17–28 (1959).
9. Schaefer, P.: Matrix transformations of almost convergent sequences. Math. Z. **112**, 321–325 (1969).

Dr. J. P. Duran
Department of Mathematics
The University of Puerto Rico
Mayagüez, Puerto Rico 00708

(Received March 21, 1972)

Math. Z. 128, 85–92 (1972)

Über die Faltung schlichter Funktionen

Stephan Ruscheweyh

I. Einführung

Mit S, C, S^*, K seien die Familien der im Einheitskreis $EK := \{z \mid |z| < 1\}$ holomorphen und in der üblichen Weise normierten, schlichten, fast-konvexen[1], sternförmigen bzw. konvexen Funktionen bezeichnet. Bekanntlich ist $S \supset C \supset S^* \supset K$. Wenn

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \qquad g(z) = \sum_{k=0}^{\infty} b_k z^k$$

in EK holomorph sind, so sei

$$h(z) = (f * g)(z) := \sum_{k=0}^{\infty} a_k b_k z^k$$

das Hadamardprodukt von f und g.

Pólya und Schönberg [5] stellten die folgende Vermutung auf:

$$f \in K, \; g \in K \Rightarrow f * g \in K, \tag{1}$$

$$f \in S^*, \; g \in K \Rightarrow f * g \in S^*. \tag{2}$$

Die Vermutungen (1) und (2) sind offenbar äquivalent. Ein Beweis konnte bisher nicht gefunden werden, lediglich einige Teilmengen von K und S^*, für die (1) und (2) richtig sind, wurden konstruiert (s. z. B. Mairhuber *et al.* [4], Ruscheweyh [8]). Suffridge [13] konnte jedoch eine schwächere Aussage beweisen:

Lemma 1.

$$f \in K, \; g \in K \Rightarrow f * g \in C. \tag{3}$$

Die Vermutung von Pólya und Schoenberg legt die Betrachtung der Funktionenfamilie

$$M := \{f \in S \mid f * g \in S \text{ für alle } g \in K\} \tag{4}$$

nahe. Während nach Lemma 1 $K \subset M$ gilt, würde (2) $S^* \subset M$ bedeuten. Man kann leicht sehen (Satz 1), daß in M die Bieberbachsche Vermutung richtig ist, jedoch ist $M \neq S$ wie ein – aus anderem Anlaß konstruiertes – Beispiel von Epstein und Schoenberg [1] zeigt.

[1] Im englischen Sprachgebrauch: close-to-convex.

Das wichtigste Ergebnis dieser Arbeit ist der

Satz 4. *Sei $f \in S$. Es gebe ein $\alpha \in \mathbb{R}$ und Zahlen $\beta, \gamma \in \mathbb{C}$ mit $|\beta| \leqq 1$, $|\gamma| \leqq 1$, so daß in EK*

$$\operatorname{Re} e^{i\alpha}(1-\beta z)(1-\gamma z) f'(z) > 0 \tag{5}$$

gilt. Dann ist $f * g \in C$ für jedes $g \in K$. Insbesondere ist $f \in M$.

Dieser Satz umfaßt, neben der Aussage von Lemma 1, auch bekannte Ergebnisse von Rahmanov [6] und Sheil-Small [11]. Der Beweis basiert auf einer Dreipunkteformel für die Funktionen aus K (Satz 3), die eine große Anzahl bekannter Abschätzungen für diese Funktionen als Spezialfall enthält.

Herrn Dr. K.J. Wirths bin ich für wertvolle Diskussionen zu Dank verpflichtet.

Unterstützung bei der Abfassung dieser Arbeit wurde vom Sonderforschungsbereich 40 (Theoretische Mathematik) gewährt.

II. Eigenschaften der Familie M

Wir beweisen zunächst, daß in M die Bieberbachsche Vermutung richtig ist.

Satz 1. *Sei $f(z) = \sum_{k=1}^{\infty} a_k z^k \in M$. Dann ist $|a_k| \leqq k$, $k \in \mathbb{N}$.*

Beweis. Man verifiziert leicht die Gültigkeit der folgenden Aussagen

$$z + \sigma z^k \in K \Leftrightarrow |\sigma| \leqq k^{-2}, \tag{6}$$

$$z + \sigma z^k \in S \Leftrightarrow |\sigma| \leqq k^{-1}, \tag{7}$$

für $k = 2, 3, \ldots$. Wegen (4) und (6) ist

$$f * \left(z + \frac{z^k}{k^2}\right) = z + \frac{a_k}{k^2} z^k \in S,$$

und mit (7) folgt die Behauptung.

Sehr wichtig für viele Untersuchungen ist die im nächsten Satz ausgesprochende Charakterisierung der Familie M.

Satz 2. *Sei $f(z)$ in EK holomorph, $f(0)=0$, $f'(0)=1$. Dann gilt*

$$f \in M \Leftrightarrow \frac{1}{z}(f * p) \neq 0 \quad \textit{für alle } z \in EK \textit{ und alle } p \in S^*. \tag{8}$$

Zum Beweis benötigen wir ein Ergebnis von Suffridge [13]:

Lemma 2. *Sei $g \in K$, $|\sigma| \leqq 1$, $\sigma \neq 1$. Dann ist $\dfrac{g(z) - g(\sigma z)}{1-\sigma} \in S^*$.*

Beweis (von Satz 2). 1) Seien $f \in M$ und $p \in S^*$. Dann ist

$$g(z) := \int_0^z \frac{p(x)}{x} \, dx \in K,$$

also $f * g \in S$. Hieraus folgt $(f * g)' = \frac{1}{z}(f * p) \neq 0$ in EK.

2) Seien $g \in K$, $h = f * g$. Seien z_1, z_2 beliebig in EK gewählt mit $|z_1| \leqq |z_2|$, $z_1 \neq z_2$, und sei $\sigma = z_1/z_2$. Nach Lemma 2 ist $p = \frac{g(z) - g(\sigma z)}{1 - \sigma} \in S^*$, und eine leichte Rechnung zeigt $\frac{1}{z}(f * p) = \frac{h(z) - h(\sigma z)}{(1 - \sigma) z}$. Setzt man $z = z_2 \neq 0$, so folgt aus der Voraussetzung $h(z_2) - h(z_1) \neq 0$, also $h \in S$. Da $g_0(z) = \frac{z}{1 - z} \in K$ und $f * g_0 = f$ ist, haben wir mit dem Obigen auch $f \in S$ bewiesen. Damit ist alles gezeigt.

Als nächstes geben wir ein notwendiges und hinreichendes Kriterium dafür an, daß gewisse Funktionenfamilien in M liegen. Dazu bezeichnen wir für jedes feste $h \in K$:

$$A_h := \left\{ f \in S \,|\, \operatorname{Re}\left[e^{i\alpha} \frac{f'(z)}{h'(z)} \right] > 0 \text{ für ein } \alpha \in \mathbb{R},\ z \in EK \right\}. \tag{9}$$

Offenbar ist $A_h \subset C$ (siehe z.B. [3]), und es gilt:

Lemma 3. *$A_h \subset M$ ist gleichbedeutend mit*

$$\operatorname{Re} \frac{\frac{z h'}{1 - \alpha z} * g}{z h' * g} > \frac{1}{2} \quad \textit{für alle } |\alpha| \leqq 1,\ z \in EK, \textit{ und alle } g \in K. \tag{10}$$

Bemerkung 1. Da $\bigcup_{h \in K} A_h = C$ ist, beinhaltet die Richtigkeit von (10) für alle $h \in K$ die Relation $C \subset M$. Dafür sprechen auch verschiedene andere Ergebnisse in [9], die die Variabilitätsbereiche von $\frac{f(z)}{z}$ und $\frac{z f'(z)}{f(z)}$ in C und M vergleichen.

Beweis. Nach Satz 2 ist die Aussage $A_h \subset M$ gleichbedeutend mit

$$\frac{1}{z}(f * p) \neq 0 \quad \text{für alle } p \in S^*,\ f \in A_h. \tag{11}$$

Sei P die Familie der in EK holomorphen Funktionen $F(z)$, für die dort $\operatorname{Re} F(z) \geqq 0$ und $F(0) = 1$ gilt. Eine Funktion $f \in S$ ist genau dann in A_h, wenn sie eine Ableitung der Form

$$f'(z) = e^{-i\alpha} h'(z)[\cos\alpha\, F(z) + i \sin\alpha] \tag{12}$$

mit $F \in P$ und $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ besitzt. Sei nun $p \in S^*$, also $g(z) = \int\limits_0^z \frac{p(x)}{x}\, dx \in K$. Dann ist

$$\frac{1}{z}(f * p) = \frac{1}{z}(z f' * g)$$

$$= \frac{e^{-i\alpha}}{z}(z h' F * g)\cos\alpha + i\,\frac{e^{-i\alpha}}{z}(z h' * g)\sin\alpha,$$

und (11) ist äquivalent mit

$$\frac{z h' F * g}{z h' * g} \neq -i\,\frac{\sin\alpha}{\cos\alpha} \qquad \text{für alle } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],\ F \in P,\ g \in K. \tag{13}$$

Da die Funktion auf der linken Seite von (13) in $z=0$ den Wert 1 hat, folgt aus dem Satz über die Gebietstreue, daß (13) äquivalent zu

$$\mathrm{Re}\,\frac{z h' F * g}{z h' * g} > 0 \qquad \text{für alle } F \in P,\ g \in K, \tag{14}$$

ist. Nach der Formel von Herglotz [2] stimmt P überein mit der Menge der Funktionen

$$F(z) = \frac{1}{2\pi} \int\limits_0^{2\pi} \frac{1 + z e^{-i\varphi}}{1 - z e^{-i\varphi}}\, d\mu(\varphi),$$

wobei $\mu(\varphi)$ eine monoton wachsende Funktion mit $\mu(2\pi) - \mu(0) = 2\pi$ ist. Setzt man für festes $g \in K$, $z \in EK$

$$k(\varphi) = \mathrm{Re}\,\frac{z h' \dfrac{1 + z e^{-i\varphi}}{1 - z e^{-i\varphi}} * g}{z h' * g},$$

so ist nach (14)

$$\frac{1}{2\pi} \int\limits_0^{2\pi} k(\varphi)\, d\mu(\varphi) > 0 \qquad \text{für alle zulässigen } \mu(\varphi).$$

Das ist aber genau dann der Fall, wenn $k(\varphi) > 0$ für alle $\varphi \in [0, 2\pi]$ ist. Beachtet man nun noch

$$\frac{1}{2} k(\varphi) = \mathrm{Re}\,\frac{\dfrac{z h'}{1 - \alpha z} * g}{z h' * g} - \frac{1}{2}, \qquad \alpha = e^{-i\varphi},$$

so wird die Äquivalenz von (10) und (11) deutlich.

Bemerkung 2. Sei für festes $h\in K$: $A_h\subset M$. Wenn dann $h*g\in K$ für alle $g\in K$ ist (das ist z.B. der Fall, wenn Vermutung (1) richtig ist), so kann man aus (12) und (14) sofort ablesen:

$$f\in A_h,\ g\in K \Rightarrow f*g\in A_{h*g}\subset C. \tag{15}$$

III. Eine Ungleichung für Funktionen aus K

Wir wollen (11) für einen Spezialfall beweisen. Dazu benötigen wir die folgende Dreipunkteformel für die Funktionen aus K:

Satz 3. *Sei $g\in K$, $z_j\in EK$, $j=1,2,3$. Dann gilt:*

$$\operatorname{Re}\left[\frac{z_1}{z_1-z_3}\,\frac{z_2-z_3}{z_2-z_1}\,\frac{g(z_2)-g(z_1)}{g(z_2)-g(z_3)}\right]\geqq\frac{1}{2}\operatorname{Re}\frac{z_1+z_3}{z_1-z_3}. \tag{16}$$

Beweis. Nach Sheil-Small [10] ist für $h\in K$, $x, y\in EK$

$$\operatorname{Re}\left[\frac{x}{h(x)}\,\frac{h(x)-h(y)}{x-y}\right]\geqq\frac{1}{2}. \tag{17}$$

Sei nun g die im Satz genannte Funktion, von der wir ohne Beschränkung der Allgemeinheit annehmen wollen, daß sie in $|z|\leqq 1$ holomorph ist. Sei

$$h(\eta):=\frac{g\left(\dfrac{\eta+z_1}{1+\overline{z_1}\,\eta}\right)-g(z_1)}{g'(z_1)(1-|z_1|^2)},$$

also $h(\eta)\in K$, und $x=\dfrac{z_2-z_1}{1-\overline{z_1}\,z_2}$, $y=\dfrac{z_3-z_1}{1-\overline{z_1}\,z_3}$. Dann folgt aus (17), angewandt auf $h(\eta)$:

$$\operatorname{Re}\left[\frac{1-z_3\overline{z_1}}{1-|z_1|^2}\,\frac{z_2-z_1}{z_2-z_3}\,\frac{g(z_2)-g(z_3)}{g(z_2)-g(z_1)}\right]\geqq\frac{1}{2},$$

also

$$\operatorname{Re}\left[\frac{1}{1-z_3\overline{z_1}}\,\frac{z_2-z_3}{z_2-z_1}\,\frac{g(z_2)-g(z_1)}{g(z_2)-g(z_3)}\right]\geqq 0.$$

Aus Stetigkeitsgründen ist diese Beziehung auch noch für $|z_1|\leqq 1$, $|z_3|\leqq 1$, $z_1\neq z_3$ gültig. Für $|z_1|=1$ folgt

$$\operatorname{Re}\left[\frac{z_1}{z_1-z_3}\,\frac{z_2-z_3}{z_2-z_1}\,\frac{g(z_2)-g(z_1)}{g(z_2)-g(z_3)}\right]\geqq 0. \tag{18}$$

Wie man unmittelbar sieht, ist für festes $z_2\in EK$ die Funktion

$$H(z_1, z_3):=\frac{z_1}{z_1-z_3}\,\frac{z_2-z_3}{z_2-z_1}\,\frac{g(z_2)-g(z_1)}{g(z_2)-g(z_3)}-\frac{1}{2}\,\frac{z_1+z_3}{z_1-z_3}$$

im Dizylinder $|z_1| \leqq 1$, $|z_3| \leqq 1$ holomorph. Auf dem Rand gilt jedoch

$$\operatorname{Re} H(z_1, z_3) \geqq -\frac{1}{2} \operatorname{Re} \frac{z_1 + z_3}{z_1 - z_3} = 0.$$

Aus dem Minimumprinzip für harmonische Funktionen folgt nun (16).

Bemerkung 3. Eine leichte Umrechnung zeigt, daß (16) äquivalent zu der Formel

$$\operatorname{Re} \frac{\dfrac{z}{(1-\alpha z)(1-\beta z)(1-\gamma z)} * g}{\dfrac{z}{(1-\beta z)(1-\gamma z)} * g} \geqq \frac{1}{2} \tag{19}$$

für $z \in EK$, $|\alpha| \leqq 1$, $|\beta| \leqq 1$, $|\gamma| \leqq 1$, $g \in K$ ist. (19) enthält eine Reihe bekannter Abschätzungen der Funktionen aus K. Für $\gamma = 0$ entsteht wieder (17). Mit $\alpha = \beta = 1$ wird (19) zu

$$\operatorname{Re} \frac{z g'(z)}{g(z) - g(\gamma z)} \geqq \frac{1}{2} \operatorname{Re} \frac{1+\gamma}{1-\gamma},$$

was ebenfalls von Sheil-Small [10] bewiesen wurde. Die Fälle $\alpha = \beta = 1$, $\gamma = 0$ und $\alpha = 1$, $\beta = \gamma = 0$ ergeben die bekannten Abschätzungen

$$\operatorname{Re} \frac{z g'(z)}{g(z)} \geqq \frac{1}{2} \quad \text{und} \quad \operatorname{Re} \frac{g(z)}{z} \geqq \frac{1}{2},$$

die zuerst von Strohhäcker [12] bewiesen wurden. Setzt man schließlich $\alpha = \beta = \gamma = 1$, so wird (19), unter Beachtung der Relationen

$$\frac{z}{(1-z)^2} * g = z g', \qquad \frac{z}{(1-z)^3} * g = z g' + \frac{1}{2} z^2 g'',$$

zu

$$\operatorname{Re} \left[z \frac{g''}{g'} + 1 \right] \geqq 0,$$

woraus man erkennt, daß (19) notwendig und hinreichend für die Funktionen aus K ist.

IV. Der Satz 4 und seine Folgerungen

Der Beweis von Satz 4 (s. Einführung) ist nun sehr einfach: Setzt man

$$h(z) := \begin{cases} \dfrac{1}{\beta - \gamma} \log \dfrac{1-\gamma z}{1-\beta z} & \text{für } \beta \neq \gamma \\[2ex] \dfrac{z}{1-\beta z} & \text{für } \beta = \gamma, \end{cases}$$

so ist $h(z)\in K$ und $zh'(z)=\dfrac{z}{(1-\beta z)(1-\gamma z)}$, und aus (10), (19) folgt $A_h\subset M$. In diesem Falle ist aber sogar für alle $g\in K$:

$$h*g=\begin{cases}\displaystyle\int_0^z \frac{zh'*g}{z}\,dz=\int_0^z \frac{g(\beta z)-g(\gamma z)}{z(\beta-\gamma)}\,dz\in K & \text{für } \beta\neq\gamma,\\[2ex] \dfrac{g(\beta z)}{\beta}\in K \quad \text{für } \beta=\gamma, & \end{cases}$$

wegen Lemma 2. Deshalb kann man wie in (15) schließen, und der Beweis ist vollendet.

Die in Satz 4 genannte Funktionenfamilie umfaßt zwei bekannte Teilmengen von C, nämlich die Menge der Funktionen $f\in S$, für die $\operatorname{Re} e^{i\alpha} f'(z)\geqq 0$ in EK ist, und diejenigen, die konvex in einer Richtung sind, also insbesondere auch K (s. Robertson [7]). Für diese Funktionen entspricht die Aussage des Satz 4 gerade der des Lemma 1.

Wir zeigen nun, daß noch zwei weitere bekannte Resultate in unserem Satz enthalten sind:

Lemma 4 (Rahmanov [6]). *Sei* $g\in K$, $\alpha\in\left[-\dfrac{\pi}{2},\dfrac{\pi}{2}\right]$. *Dann ist*

$$f(z):=\frac{1}{1+e^{i\alpha}}\,[g(z)+e^{i\alpha} z g'(z)]\in S.$$

Beweis. Wir beweisen, daß sogar $f(z)\in C$ gilt. Es ist nämlich $f(z)=f_0(z)*g(z)$ mit

$$f_0(z)=\frac{1}{1+e^{i\alpha}}\left[\frac{z}{1-z}+e^{i\alpha}\frac{z}{(1-z)^2}\right].$$

Für f_0 gilt aber

$$\operatorname{Re}[(1+e^{-i\alpha})(1-z)^2 f_0'(z)]=\operatorname{Re}\left[\frac{1+e^{-i\alpha}}{1+e^{i\alpha}}\right]+\operatorname{Re}\left[\frac{1+z}{1-z}\right]>0$$

in $\alpha\in\left[-\dfrac{\pi}{2},\dfrac{\pi}{2}\right]$, so daß Satz 4 anwendbar ist.

Lemma 5 (Sheil-Small [11]). *Sei* $g(z)=\sum_{k=1}^{\infty} a_k z^k\in K$. *Dann ist* $Q_n:=\sum_{k=1}^{n}\dfrac{a_k}{k}z^k\in C$, $n\in\mathbb{N}$.

Beweis. Es ist $Q_n(z)=\sum_{k=1}^{n}\dfrac{z^k}{k}*g$, und

$$\operatorname{Re}\left[(1-z)\left(\sum_{k=1}^{n}\frac{z^k}{k}\right)'\right]=\operatorname{Re}(1-z^n)>0,$$

woraus alles folgt.

Ebenso leicht findet man, unter Verwendung der Bezeichnungen von Lemma 5:

Satz 5. *Sei $g \in K$. Dann ist $\frac{1}{2}[Q_{2n-1}(z) - Q_{2n-1}(-z)] \in C$, $n \in \mathbb{N}$.*

Beweis. Es ist $\frac{1}{2}[Q_{2n-1}(z) - Q_{2n-1}(-z)] = \sum_{k=0}^{n-1} \frac{z^{2k+1}}{2k+1} * g$ und

$$\operatorname{Re}\left[(1-z^2)\left(\sum_{k=0}^{n-1} \frac{z^{2k+1}}{2k+1}\right)'\right] = \operatorname{Re}(1-z^{2n}) > 0.$$

Literatur

1. Epstein, B., Schoenberg, I.J.: On a conjecture concerning schlicht functions. Bull. Amer. Math. Soc. **65**, 273–275 (1959).
2. Herglotz, G.: Leipziger Berichte **63**, 501–511 (1911).
3. Kaplan, W.: Close-to-convex schlicht functions. Michigan Math. J. **1**, 169-185 (1952).
4. Mairhuber, J.C., Schoenberg, I.J., Williamson, R.E.: On variation diminishing transformations on the circle. Rend. Circ. Mat. Palermo (2), **8**, 241–270 (1959).
5. Pólya, G., Schoenberg, I.J.: Remarks on the de la Vallée Poussin means and convex conformal maps of the circle. Pacific J. Mat. Mech. **8**, 295–334 (1958).
6. Rahmanov, B.N.: On the theory of univalent functions. Doklady Akad. Nauk. SSSR, **78**, 209–211 (1951).
7. Robertson, M.S.: Analytic functions starlike in one direction. Amer. J. Math. **58**, 465–472 (1936).
8. Ruscheweyh, St.: Geometrische Eigenschaften der Lösungen der Differentialgleichung $(1-z\bar{z})^2\, w_{z\bar{z}} - n(n+1)\, w = 0$. Unveröffentlicht.
9. Ruscheweyh, St., Wirths, K.J.: Über die Faltung schlichter Funktionen. II. Unveröffentlicht.
10. Sheil-Small, T.: On convex univalent functions. J. London Math. Soc. (2), **1**, 483–492 (1969).
11. Sheil-Small, T.: A note on the partial sums of convex schlicht functions. Bull. London Math. Soc. **2**, 165–168 (1970).
12. Strohhäcker, E.: Beiträge zur Theorie der schlichten Funktionen. Math. Z. **37**, 356–380 (1933).
13. Suffridge, T.J.: Convolutions of convex functions. J. Math. Mech. **15**, 795–804 (1966).

Prof. Stephan Ruscheweyh
Math. Institut d. Univ. Bonn
D-5300 Bonn, Wegelerstraße 10
Deutschland

(Eingegangen am 8. März 1972)

Math. Z. 128, 93–108 (1972)

A-priori-Schranken für semi- und quasilineare parabolische Differentialgleichungen *

Wolf von Wahl

0. Einleitung und Bezeichnungen

Gegenstand dieser Untersuchung ist das Rand-Anfangswertproblem für semilineare und quasilineare parabolische Differentialgleichungen über einem zylindrischen Gebiet $[0,T]\times\bar{\Omega}$. Dabei ist T eine positive Zahl und Ω eine beschränkte offene Punktmenge des $\mathbb{R}^n$ mit glattem Rand.

Wenden wir uns zunächst den semilinearen parabolischen Gleichungen zweiter Ordnung zu, d.h.:

$$\frac{\partial u}{\partial t}-\sum_{i,j=1}^{n} a_{ij}(t,x)\frac{\partial^2 u}{\partial x_i\,\partial x_j}=f(t,x,u,\nabla u). \tag{I}$$

Dabei soll

$$M^{-1}|\xi|^2\leqq\sum_{i,j=1}^{n} a_{ij}(t,x)\,\xi_i\,\xi_j\leqq M|\xi|^2,\quad (t,x,\xi)\in[0,T]\times\bar{\Omega}\times\mathbb{R}^n$$

sowie

$$|f(t,x,u,p)|\leqq\mu(|\mu|)\,|p|^2+K,\quad (t,x,u,p)\in[0,T]\times\bar{\Omega}\times\mathbb{R}\times\mathbb{R}^n \tag{II}$$

sein. K und M sind positive Konstanten, μ ist eine monoton wachsende Funktion. Wir geben zunächst eine a-priori-Abschätzung der Größe

$$\sup_{0\leqq t\leqq T}\|\nabla u(t)\|_{C^\alpha(\Omega)},\qquad 0<\alpha<1,$$

ohne eine quantitative Einschränkung an $\|u\|_{C^0(\Omega)}$. Da die a_{ij} nur stetig von x und hölderstetig von t abzuhängen brauchen, versagen hier die Methoden aus [2].

Neben a-priori-Abschätzungen unter einer quantitativen Einschränkung an $\|u\|_{C^0(\bar{\Omega})}$, die im Falle der Gl. (I) auf das Studium eines allge-

---

* Dies ist der zweite Teil einer Arbeit, die von der Mathematisch-Naturwissenschaftlichen Fakultät der Georg-August-Universität zu Göttingen als Habilitationsschrift angenommen wurde.

meineren Systems zurückgehen, bildet die geeignete Übertragung einer von Tomi erfundenen Kontinuitätsmethode ein wichtiges Hilfsmittel zu dem Zweck, sich von der Kleinheitsbedingung an $\|u\|_{C^0(\Omega)}$ zu befreien. Dieser Autor hat die genannte Methode dazu benutzt, ein entsprechendes Problem für semilineare elliptische Gleichungen zu behandeln. Es sei darauf hingewiesen, daß für gewisse spezielle Systeme zweiter Ordnung ebenfalls eine quantitative Einschränkung an $\|u\|_{C^0(\Omega)}$ entbehrlich ist.

Mit tiefliegenden Hilfsmitteln aus der Theorie der quasilinearen elliptischen Gleichungen läßt sich auch der Fall einer quasilinearen parabolischen Gleichung

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{2} a_{ij}(x, \nabla u) \frac{\partial^2 u}{\partial x_i \partial x_j} = f(t, x, u, \nabla u)$$

in zwei Raumvariablen behandeln. Dabei wird von f die Wachstumsbedingung (II) ohne eine Kleinheitsbedingung an $\|u\|_{C^0(\Omega)}$ gefordert. Der nichtlineare Operator

$$\sum_{i,j=1}^{2} a_{ij}(x, \nabla u) \frac{\partial^2 u}{\partial x_i \partial x_j}$$

muß gleichmäßig elliptisch sein, weiter müssen gewisse Beschränktheitsaussagen für $\partial f/\partial u$ und $\partial f/\partial t$ gelten. Die Ableitungen der a_{ij} nach den Komponenten von ∇u brauchen dagegen nur einer wesentlich schwächeren Wachstumsbedingung als in [2] zu genügen. Eine kleine Variante unseres Beweises liefert ein entsprechendes Resultat für quasilineare elliptische Differentialgleichungen ohne eine Bedingung an $\partial f/du$.

An dieser Stelle sei darauf hingewiesen, daß die Ausnahmesituation wie sie in zwei Raumdimensionen bei nichtlinearen elliptischen Differentialgleichungen besteht, bei parabolischen i. allg. nur bei einer Raumdimension auftritt; beispielsweise lassen sich dann auch dort gewisse Ansätze von Bernstein verwenden ([2], S. 538).

Die Frage nach der Existenz im Großen von klassischen Lösungen wird in dieser Arbeit nicht behandelt. Mit Hilfe der von uns gewonnenen a-priori-Schranken läßt sie sich in geläufiger Weise erledigen, nämlich durch topologische Fixpunktsätze ([2, 1]) oder ein Iterationsverfahren ([1]). Die erforderliche Schranke für den Betrag der Lösung läßt sich etwa mit Hilfe eines Maximumprinzips gewinnen.

Wir führen einige Bezeichnungen ein. Sei X ein Banachraum. Unter $C^\nu([0, T], X)$ verstehen wir den Banachraum aller bis zur Ordnung ν stetig differenzierbaren Abbildungen von $[0, T]$ in X. Mit $C^{\nu+\alpha}([0, T], X)$, $0<\alpha<1$, bezeichnen wir den Banachraum aller bis zur Ordnung stetig differenzierbaren Abbildungen u, für die

$$\sup_{t,s\in[0,T]} \frac{1}{|t-s|^{\alpha}} \left\| \frac{d^{\nu}u}{dt^{\nu}}(t) - \frac{d^{\nu}u}{dt^{\nu}}(s) \right\|_X + \sum_{\lambda=0}^{\nu} \sup_{0\leqq t\leqq T} \left\| \frac{d^{\lambda}u}{dt^{\lambda}}(t) \right\|_X$$

endlich ist. $L^p([0,T],X)$, $1\leqq p\leqq\infty$, ist der Banachraum aller meßbaren Abbildungen u: $]0,T[\to X$ mit der Norm

$$\left(\int_0^T \|u(t)\|_X^p\,dt\right)^{1/p} \quad \text{bzw.} \quad \operatorname*{ess\,sup}_{0\leqq t\leqq T} \|u(t)\|_X.$$

Ω ist stets eine beschränkte offene Teilmenge des $\mathbb{R}^n$. $H^{\nu,p}(\Omega)$ ist der Banachraum aller auf Ω erklärten meßbaren komplexwertigen Funktionen mit in $L^p(\Omega)$ liegenden Distributionsableitungen bis zur Ordnung ν. Seine Norm wird mit $\|\cdot\|_{\nu,p}$ oder $\|\cdot\|_{H^{\nu,p}(\Omega)}$ bezeichnet. $\mathring{H}^{\nu,p}(\Omega)$ ist wie üblich die Vervollständigung von $C_0^{\infty}(\Omega)$ in der $H^{\nu,p}(\Omega)$-Norm. Die Norm von $C^0([0,T],H^{\nu,p}(\Omega))$ bezeichnen wir mit $|||\cdot|||_{\nu,p}$.

Wir werden häufig Funktionen betrachten, die auf $Q_T := [0,T]\times\bar{\Omega}$ stetig sind und für die

$$\sup_{t,s\in[0,T]} \frac{|f(x,t)-f(x,s)|}{|t-s|^{\gamma}}$$

mit einem γ, $1>\gamma\geqq 0$, endlich ist. Wir setzen dann

$$|||f|||_{\gamma,Q_T} := \sup_{(t,x)\in Q_T} |f(t,x)| + \sup_{(t,s)\in[0,T],\,x\in\Omega} \frac{|f(x,t)-f(x,s)|}{|t-s|^{\gamma}}$$

$$||||f||||(r) := \sup_{\substack{s,t\in[0,T],\,x\ y\in\{|x-y|\leqq r\}\cap\Omega,\\ |t-s|\leqq r}} |f(x,s)-f(y,t)|$$

$$||||f||||_x(r) := \sup_{t\in[0,T],\,x,y\in\{|x-y|\leqq r\}\cap\Omega} |f(x,t)-f(y,t)|.$$

Im Falle $\gamma=0$ schreiben wir auch $|||f|||$ statt $|||f|||_{0,Q_T}$.

$C^{\nu}(\bar{\Omega})$ bzw. $C^{\nu+\alpha}(\bar{\Omega})$ ist der Banachraum der auf Ω erklärten stetigen komplexwertigen Funktionen, die in Ω ν-mal stetig differenzierbar sind und deren partielle Ableitungen bis zur Ordnung ν auf $\bar{\Omega}$ stetig bzw. hölderstetig zum Exponenten α sind.

Wir sagen, daß Ω zur Klasse $C^{\nu+\alpha}$ gehört, wenn für $\partial\Omega$ lokal eine Darstellung

$$x_k = \varphi(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)$$

mit $\varphi\in C^{\nu+\alpha}$ gilt und Ω lokal auf einer Seite von $\partial\Omega$ liegt. Ist die Funktion g auf $\partial\Omega$ definiert, so sagen wir, daß g zu $C^{\nu+\alpha}(\partial\Omega)$ gehört, wenn es eine $\mathbb{R}^n$-Umgebung U von $\partial\Omega$ und $\tilde{g}\in C^{\nu+\alpha}(\bar{U})$ mit $\tilde{g}|\partial\Omega = g$ gibt. Wir setzen dann

$$\|g\|_{\nu+\alpha,\partial\Omega} := \inf_{(U,g)} \|\tilde{g}\|_{C^{\nu+\alpha}(\bar{U})}, \qquad \|g\|_{\partial\Omega} := \|g\|_{0,\partial\Omega}.$$

Die Norm von $C^{\nu}(\bar{\Omega})$ bzw. $C^{\nu+\alpha}(\bar{\Omega})$ wird mit $\|.\|_{\nu}$ bzw. $\|.\|_{\nu+\alpha}$, die von $C^{\alpha}([0,T], C^{\alpha}(\bar{\Omega}))$ mit $\|.\|_{\alpha}$ und die von $C^{0}([0,T], C^{\nu+\alpha}(\bar{\Omega}))$ mit $|||.|||_{\nu+\alpha}$ bezeichnet. Außerdem benutzen wir die Abkürzungen $|||.|||:=|||.|||_0$, $\|.\|_{L^p(\Omega)}:=\|.\|_p:=\|.\|_{0,p}$ und

$$D_j := \frac{1}{i}\frac{\partial}{\partial x_j},$$

$$D^{\tilde{\alpha}} := \prod_{j=1}^{n} D_j^{\alpha_j}, \qquad |\tilde{\alpha}| := \sum_{j=1}^{n} \alpha_j,$$

wobei $\tilde{\alpha}$ ein Multiindex des $\mathbb{R}^n$ mit den Komponenten $\alpha_j \in \mathbb{N} \cup \{0\}$ ist. $\mathbb{N}$ bedeutet stets die Menge aller positiven ganzen Zahlen, $\mathbb{R}^+$ ist die Menge aller nichtnegativen reellen Zahlen.

I. Semilineare parabolische Differentialgleichungen

T, m und α seien im folgenden beliebige feste positive Zahlen, $0<\alpha<1$, $m\in\mathbb{N}$. Ω sei von der Klasse C^{4m}. N sei die Anzahl aller n-Tupel $\tilde{\alpha}$ nichtnegativer ganzer Zahlen mit $|\tilde{\alpha}|\leqq 2m-1$. Es seien Abbildungen

$$A_{\tilde{\alpha}}\colon\ \mathbb{R}^+ \times \bar{\Omega} \to \mathbb{C}, \qquad |\tilde{\alpha}| = 2m,$$

gegeben mit folgenden Eigenschaften:

$$A_{\tilde{\alpha}}\colon\ \mathbb{R}^+ \times \bar{\Omega} \to \mathbb{R}, \qquad |\tilde{\alpha}| = 2m,$$

$$A_{\tilde{\alpha}}(.) \in C^{\alpha}([0,T], C^{0}(\bar{\Omega})),$$

$$M|\xi|^{2m} \geqq \sum_{|\tilde{\alpha}|=2m} A_{\tilde{\alpha}}(t,x)\,\xi^{\tilde{\alpha}} \geqq M^{-1}|\xi|^{2m}, \quad \xi\in\mathbb{R}^n, \quad (t,x)\in[0,T]\times\bar{\Omega} \tag{1}$$

mit einer positiven Konstanten M. Weiter setzen wir im folgenden

$$A_p(t)\,u := \sum_{|\tilde{\alpha}|=2m} A_{\tilde{\alpha}}(t,x)\,D^{\tilde{\alpha}} u, \qquad u\in H^{2m,p}(\Omega)\cap \mathring{H}^{m,p}(\Omega).$$

Sei $L\in\mathbb{N}$. Sei $\mathbf{u}\in(H^{2m,p}(\Omega)\cap\mathring{H}^{m,p}(\Omega))^L$. Wir setzen

$$\mathscr{A}_p(t)\,\mathbf{u} := (\mathscr{A}_p(t)\,u_1, \ldots, \mathscr{A}_p(t)\,u_L).$$

Es seien lokal hölderstetige Abbildungen f_λ gegeben, die der Wachstumsbedingung

$$\begin{gathered} |f_\lambda(t,x,\mathbf{u}(x), D^1\mathbf{u}(x), \ldots, D^{2m-1}\mathbf{u}(x))| \leqq \mu(\|u\|_0) \sum_{\nu=1}^{n} |D^{\nu}\mathbf{u}(x)|^{2m/\nu} + K, \\ (t,x)\in[0,T]\times\bar{\Omega}, \qquad \mathbf{u}\in(C^{2m-1}(\bar{\Omega}))^L, \qquad \lambda=1,\ldots,L, \end{gathered} \tag{2}$$

mit einer stetigen, monoton wachsenden Funktion μ und einer Konstanten $K>0$ genügen. Wir setzen

$$\mathbf{f}(t, x, \mathbf{u}, D^1 \mathbf{u}, \ldots, D^{2m-1} \mathbf{u})$$
$$:= \big(f_1(t, x, \mathbf{u}, D^1 \mathbf{u}, \ldots, D^{2m-1} \mathbf{u}), \ldots, f_L(t, x, \mathbf{u}, D^1 \mathbf{u}, \ldots, D^{2m-1} \mathbf{u})\big).$$

Nach [1], S. 108 ff. existiert genau eine Operatorenschar

$$\{U_p(t,s) \mid U_p(t,s) \in \mathscr{L}(L^p(\Omega)),\ T \geqq t \geqq s \geqq 0\}$$

mit folgenden Eigenschaften:

1. $U_p(t,s)$ ist stark stetig,
2. $U_p(t,s)$ ist stark stetig differenzierbar für $T \geqq t > s \geqq 0$,
3. $U_p(t,s)$ bildet für $T \geqq t > s \geqq 0$ den Raum $L^p(\Omega)$ in $H^{2m,p}(\Omega) \cap \mathring{H}^{m,p}(\Omega)$ ab,
4. es ist

$$\frac{\partial U_p(t,s)}{\partial t} + \mathscr{A}_p(t)\, U_p(t,s) = 0, \qquad T \geqq t > s \geqq 0, \qquad U_p(t,t) = I.$$

Weiter setzen wir voraus, daß die rechte Halbebene in der Resolventenmenge von $\mathscr{A}_p(t)$ liegt, $0 \leqq t \leqq T$ – dies ist für unsere Zwecke keine Einschränkung (vgl. [5], Kap. I).

Sei $1 > \rho > 0$. Dann erklärt man den Operator $\mathscr{A}_p^{-\rho}(t)$ durch

$$\mathscr{A}_p^{-\rho}(t) = \frac{\sin \pi \rho}{\pi} \int_0^\infty \lambda^{-\rho} (\lambda + \mathscr{A}_p(t))^{-1}\, d\lambda.$$

Dieser Operator ist eineindeutig, und seine Inverse wird mit $\mathscr{A}_p^{\rho}(t)$ bezeichnet (s. hierzu [1], Kap. 2.14).

Insbesondere gilt die Abschätzung

$$\|u\|_{2m-1+\alpha} \leqq c(p, M, \|\|\|A_{\tilde{\alpha}}\|\|\|_x) \|\mathscr{A}_p^{\rho}(t)\, u\|,$$
$$p > \frac{n}{1-\alpha}, \qquad u \in H^{2m,p}(\Omega) \cap \mathring{H}^{m,p}(\Omega),$$

wenn nur $1-\rho$ hinreichend nahe bei 0 liegt (s. [5], Kap. V). Endlich ist

$$\|\mathscr{A}_p^{\rho}(t)\, U_p(t,s)\| \leqq c(p, M, \|\|A_{\tilde{\alpha}}\|\|_{\alpha, Q_T}, \|\|\|A_{\tilde{\alpha}}\|\|\|_x) \frac{e^{-c(p, M, \|\|A_{\tilde{\alpha}}\|\|_{\alpha, Q_T}, \|\|\|A_{\tilde{\alpha}}\|\|\|_x)(t-s)}}{|t-s|^{\rho}},$$
$$T \geqq t > s \geqq 0,$$

(s. [1], S. 158 ff.). Hier bedeuten $c(p, \ldots)$ Konstanten, die außer möglicherweise von Ω, n und α auch noch von p, M und

$$\sup_{|\tilde{\alpha}| \leqq 2m} \|\|\|A_{\tilde{\alpha}}\|\|\|_x \quad \text{bzw.} \quad \sup_{|\tilde{\alpha}| \leqq 2m} \|\|A_{\tilde{\alpha}}\|\|_{\alpha, Q_T}$$

abhängen. Entsprechendes gilt im folgenden.

Weiter sei

$$\mathbf{g}(.)\in C^{\alpha}([0,T],(C^{2m}(\bar{\Omega}))^L)\cap C^{1+\alpha}([0,T],(C^0(\bar{\Omega}))^L), \tag{3}$$

$$\varphi\in(C^{2m}(\bar{\Omega}))^L, \tag{4}$$

$$D^{\tilde{\alpha}}(g_\lambda(0)-\varphi_\lambda)\,|\,\partial\Omega=0, \quad |\tilde{\alpha}|\leqq m-1, \quad \lambda=1,\dots,L.^1 \tag{5}$$

Dann gilt der folgende Störungssatz:

Satz 1. *Die Abbildung*

$$\mathbf{u}(.)\in C^0([0,T],(H^{2m,p}(\Omega))^L)\cap C^1([0,T],(L^p(\Omega))^L), \qquad p>\frac{n}{1-\alpha},$$

besitze die folgenden Eigenschaften:

$$\frac{d\mathbf{u}}{dt}+\mathscr{A}_p(t)\,\mathbf{u}=\mathbf{f}(t,\mathbf{u},D^1\mathbf{u},\dots,D^{2m-1}\mathbf{u}),$$

$$u_\lambda(t)-g_\lambda(t)\in\mathring{H}^{m,p}(\Omega), \qquad t\in[0,T], \quad \lambda=1,\dots,L,$$

$$\mathbf{u}(0)=\varphi.$$

Dann existiert ein $\varepsilon>0$ *derart, daß für alle* $\tilde{\varphi}\in(C^{2m}(\bar{\Omega}))^L$, $\tilde{\mathbf{g}}(.)\in C^{\alpha}([0,T],(C^{2m}(\bar{\Omega}))^L)\cap C^{1+\alpha}([0,T],(C^0(\bar{\Omega}))^L)$ *mit* $D^{\tilde{\alpha}}(\tilde{\varphi}-\tilde{\mathbf{g}}(0))\,|\,\partial\Omega=0$, $|\tilde{\alpha}|\leqq m-1$, $\|\varphi-\tilde{\varphi}\|_{2m}<\varepsilon$, $\left|\!\left|\!\left|\frac{\partial\mathbf{g}}{\partial t}-\frac{d\tilde{\mathbf{g}}}{\partial t}\right|\!\right|\!\right|<\varepsilon$, $|||\mathbf{g}-\tilde{\mathbf{g}}|||_{2m}<\varepsilon$ *das Problem*

$$\frac{d\mathbf{v}}{dt}+\mathscr{A}_p(t)\,\mathbf{v}=\mathbf{f}(t,\mathbf{v},D^1\mathbf{v},\dots,D^{2m-1}\mathbf{v}),$$

$$v_\lambda(t)-\tilde{g}_\lambda(t)\in\mathring{H}^{m,p}(\Omega), \qquad t\in[0,T], \quad \lambda=1,\dots,L,$$

$$\mathbf{v}(0)=\varphi$$

genau eine Lösung

$$\mathbf{v}(.)\in C^0([0,T],(H^{2m,p}(\Omega))^L)\cap C^1([0,T],(L^p(\Omega))^L), \qquad p>\frac{n}{1-\alpha},$$

besitzt, wenn nur $f_\lambda(t,x,.)$ *stetig differenzierbar ist,* $1\leqq\lambda\leqq L$.

---

[1] Für den folgenden Satz ist es auch hinreichend,

$$\mathbf{g}(.)\in C^{\delta}([0,T],(H^{2m,p}(\Omega))^L)\cap C^{1+\delta}([0,T],(L^p(\Omega))^L),$$

$$\varphi\in(H^{2m,p}(\Omega))^L, \qquad p>\frac{n}{1-\alpha},$$

zu fordern mit einem δ, $1>\delta>0$.

Beweis. Wir betrachten die Integralgleichung

$$v_\lambda(t) = \int_0^t U_p(t,\sigma)\Big(f_\lambda(\sigma, \mathbf{v}(\sigma), D^1\mathbf{v}(\sigma), \ldots, D^{2m-1}\mathbf{v}(\sigma))$$
$$-\mathscr{A}(\sigma)\,\tilde{g}_\lambda(\sigma) - \frac{\partial g_\lambda}{\partial t}(\sigma)\Big)\,d\sigma + g_\lambda(t) + U_p(t,0)\,(\varphi_\lambda - g_\lambda(0)).$$

Wenden wir auf beide Seiten $\mathscr{A}_p^\rho(t)$ an, $0<\rho<1$, ρ hinreichend nahe bei 1, bilden die Differenz $\mathscr{A}_p^\rho(t)\,(v_\lambda(t) - \tilde{g}_\lambda(t) - (u_\lambda(t) - g_\lambda(t)))$, bedenken, daß $u_\lambda(t) - g_\lambda(t)$ einer entsprechenden Integralgleichung genügt und beachten die Relation

$$|f_\lambda(\sigma, x, \mathbf{v}(\sigma,x), D^1\mathbf{v}(\sigma,x), \ldots, D^{2m-1}\mathbf{v}(\sigma,x)$$
$$-f_\lambda(\sigma, x, \mathbf{u}(\sigma,x), D^1\mathbf{u}(\sigma,x), \ldots, D^{2m-1}\mathbf{u}(\sigma,x)|$$
$$\leqq \sum_{\nu=1}^{2NL} \sup_{\substack{0\leqq y_\mu \leqq |D^{\bar{\alpha}_\mu}(\mathbf{v}(\sigma,x)-\mathbf{u}(\sigma,x))| + 2|D^{\bar{\alpha}_\mu}\mathbf{u}(\sigma,x)| \\ 1\leqq \mu \leqq L}} \left|\frac{\partial f_\lambda}{\partial y_\nu}(\sigma, x, y_1, \ldots, y_{2NL})\right|$$
$$\cdot\, |D^{\bar{\alpha}_\nu}(\mathbf{v}(\sigma,x) - \mathbf{u}(\sigma,x))|,$$

so folgt mit häufig verwendeten Argumenten die Behauptung unseres Satzes.

Eine Folgerung aus Satz 8, [5], ist der folgende Satz. Im folgenden setzen wir voraus, daß $m=1$ oder $m>n/2$ ist.

Satz 2. **f** *sei stetig und erfülle die Wachstumsbedingung in Voraussetzung* (2). **u** *sei eine Abbildung wie sie im vorigen Satz eingeführt worden ist. Dann existiert eine Schranke*

$$\delta = \delta(M,\mu) \qquad (\delta = \delta(M, \mu, |||\!|A_{\bar{\alpha}}|\!|||_x)$$

für $m>n/2$) *derart, daß*

$$|||\mathbf{u}|||_{2m-1+\alpha}$$
$$\leqq \begin{cases} c\left(T, M, \|\boldsymbol{\varphi}\|_{2m}, |||\mathbf{g}|||_{2m}, \left|\left|\left|\dfrac{\partial \mathbf{g}}{\partial t}\right|\right|\right|, K, |||A_{\bar{\alpha}}|||_{\alpha,Q_T}, |||\!|A_{\bar{\alpha}}|\!|||_x\right), & m>\dfrac{n}{2}, \\ c\left(p, T, M, \|\boldsymbol{\varphi}\|_{2,p}, |||\mathbf{g}|||_{2,p}, \left|\left|\left|\dfrac{\partial \mathbf{g}}{\partial t}\right|\right|\right|_{0,p}, K, |||\!|A_{\bar{\alpha}}|\!|||\right), & m=1 \end{cases}$$

ist, wenn nur $p>(n+2)/(1-\alpha)$ *und*

$$|||\mathbf{u}||| \leqq \delta(M,\mu) \quad \textit{bzw.} \quad \delta(M, \mu, |||\!|A_{\bar{\alpha}}|\!|||_x)$$

ist (*δ hängt nicht von K ab!*).

Der Einfachheit halber sei im folgenden Satz $K=0$. Dann besitzt das System

$$\frac{d\mathbf{u}}{dt}+\mathscr{A}_p(t)\,\mathbf{u}=\mathbf{f}(t,\mathbf{u},D^1\,\mathbf{u},\dots,D^{2m-1}\,\mathbf{u})$$

für $\mathbf{g}=0$, $\varphi=0$ bekanntlich die Lösung $\mathbf{u}\equiv 0$. Es gilt

Satz 3. *Voraussetzung: Es sei wieder jedes f_λ nach den letzten $2NL$ Variablen stetig differenzierbar. Weiter seien $\mathbf{u}_1$ und $\mathbf{u}_2$ zwei Lösungen des parabolischen Systems*

$$\frac{d\mathbf{u}_\eta}{dt}+\mathscr{A}_p(t)\,\mathbf{u}_\eta=\mathrm{f}(t,\mathbf{u}_\eta,D^1\,\mathbf{u}_\eta,\dots,D^{2m-1}\,\mathbf{u}_\eta),\qquad \eta=1,2,$$

mit den Daten φ, $\mathbf{g}(t)$ bzw. $\tilde{\varphi}$, $\tilde{\mathbf{g}}(t)$, die die am Anfang dieses Kapitels genannten Eigenschaften besitzen. Dann gelte ein Maximumprinzip in Form der Ungleichung

$$|||\mathbf{u}_1-\mathbf{u}_2|||\leqq c(T,M,|||A_{\tilde{\alpha}}|||_{\alpha,Q_T})\,(|||\mathbf{g}-\tilde{\mathbf{g}}|||_{m-1,\partial\Omega}+\|\varphi-\tilde{\varphi}\|_{m-1}).$$

Folgerung: Für alle φ, $\mathbf{g}(t)$ mit (3), (4), (5) besitzt dann das parabolische System

$$\frac{d\mathbf{u}}{dt}+\mathscr{A}_p(t)\,\mathbf{u}=\mathbf{f}(t,\mathbf{u},D^1\,\mathbf{u},\dots,D^{2m-1}\,\mathbf{u})$$

genau eine Lösung $\mathbf{u}$ mit den in Satz 11 genannten Eigenschaften mit den Daten φ, $\mathbf{g}(t)$. Für diese gilt:

$$|||\mathbf{u}|||_{2m-1+\alpha}$$
$$\leqq\begin{cases} c\left(T,M,\|\varphi\|_{2m},|||\mathbf{g}|||_{2m},\left|\left|\left|\dfrac{\partial\mathbf{g}}{\partial t}\right|\right|\right|,|||A_{\tilde{\alpha}}|||_{\alpha,Q_T},||||A_{\tilde{\alpha}}||||_x\right), & m>\dfrac{n}{2}\\[2ex] c\left(p,T,M,\|\varphi\|_{2,p},|||\mathbf{g}|||_{2,p},\left|\left|\left|\dfrac{\partial\mathbf{g}}{\partial t}\right|\right|\right|_{0,p},||||A_{\tilde{\alpha}}||||\right), & m=1,\quad p>\dfrac{n+2}{1-\alpha}.\end{cases}$$

Beweis. Der Beweis wird nach einer Kontinuitätsmethode geführt (vgl. [4], S. 361). Es sei Σ die Menge aller $\sigma\in[0,1]$, für die unser Rand-Anfangswertproblem mit den Daten $\sigma\varphi$, $\sigma\mathbf{g}(t)$ in dem genannten Sinn lösbar ist. Die Lösung sei mit $\mathbf{u}_\sigma$ bezeichnet.

Weiter sei Σ^* die Menge aller derjenigen $\sigma\in\Sigma$, für welche die folgenden Aussagen gelten:

I. $[0,\sigma]\subset\Sigma$,

II. Für alle $\tau\in[0,\sigma]$ gilt eine Abschätzung der Form

$$|||\mathbf{u}_\tau|||_{2m-1+\alpha}$$
$$\leqq\begin{cases} c\left(\sigma,T,M,\|\varphi\|_{2m},|||\mathbf{g}|||_{2m},\left|\left|\left|\dfrac{\partial\mathbf{g}}{\partial t}\right|\right|\right|,|||A_{\tilde{\alpha}}|||_{\alpha,Q_T},||||A_{\tilde{\alpha}}||||_x\right), & m>\dfrac{n}{2}\\[2ex] c\left(p,\sigma,T,M,\|\varphi\|_{2,p},|||\mathbf{g}|||_{2,p},\left|\left|\left|\dfrac{\partial\mathbf{g}}{\partial t}\right|\right|\right|_{0,p},||||A_{\tilde{\alpha}}||||\right), & m=1,\quad p>\dfrac{n+2}{1-\alpha}.\end{cases}$$

Unter Ausnutzung des Maximumprinzips unserer Voraussetzung kann man jetzt unter Verwendung der Sätze 1 und 2 so vorgehen, wie bei Tomi [4], S. 361/62 und zeigen, daß mit $\sigma_0 \in \Sigma^*$ die Abschätzung in II auch für $\Sigma \cap [\sigma_0, \sigma_0 + \delta_0]$ gilt, wobei δ_0 eine von σ_0 und $\mathbf{u}_{\sigma_0}$ unabhängige positive Zahl ist. Wie wir es in [5] mehrfach durchgeführt haben, kann man dann

$$\|D^{\tilde{\alpha}} \mathbf{u}_\sigma\|_\varepsilon, |\tilde{\alpha}| \leqq 2m-1, \quad \sigma \in \Sigma \cap [0, \sigma_0 + \delta_0]$$

für alle hinreichend kleinen ε, $\varepsilon > 0$, abschätzen. Mit erneuter Verwendung der Integralgleichung folgt dann die Abgeschlossenheit von $\Sigma \cap [0, \sigma_0 + \delta_0]$. Mit Satz 1 folgt nun in geläufiger Weise die Behauptung des Satzes.

Unser nächstes Ziel besteht darin, uns von der Voraussetzung bezüglich des Maximumprinzips zu befreien, falls m, $L = 1$ ist, d.h. im Fall einer einzigen Gleichung in n Raumdimensionen. Hierzu gilt der

Satz 4. *u sei eine Abbildung aus $C^\delta([0, T], C^2(\bar{\Omega})) \cap C^1([0, T], C^0(\bar{\Omega}))$ mit einem δ, $0 < \delta < 1$, und*

$$\frac{du}{dt} + \mathscr{A}_p(t)\, u = f(t, u, \nabla u),$$

$$u(t) - g(t) \in \mathring{H}^{1,p}(\Omega), \quad t \in [0, T],$$

$$u(0) = \varphi, \quad p > \frac{n+2}{1-\alpha},$$

$$g(t, x), \varphi(x), f(t, x, z_1, z_2) \in \mathbb{R}, \quad (t, x, z_1, z_2) \in \mathbb{R}^+ \times \bar{\Omega} \times \mathbb{C} \times \mathbb{C}^n.$$

Dann ist

$$|||u|||_{1+\alpha} \leqq c\left(T, p, M, \|\varphi\|_{2,p}, |||g|||_{2,p}, |||u|||, ||||A_{\tilde{\alpha}}||||, \left|\left|\left| \frac{\partial g}{\partial t} \right|\right|\right|, K\right).$$

Beweis. Sei

$$\mathscr{L}(t)\, u = f(t, x, u, \nabla u),$$

$$|f(t, u, \nabla u)| \leqq A + B|\nabla u|^2,$$

$$\tilde{f}(t, x, \mathbf{p}) = \frac{\mathscr{L}(t)\, u}{A + B|\nabla u|^2}(A + B|\mathbf{p}|^2), \quad \mathbf{p} \in \mathbb{R}^n,$$

mit von $|||u|||$ abhängigen Konstanten A, $B > 0$. Dann folgt der Satz aus der Betrachtung der Gleichung

$$\frac{du}{dt} + \mathscr{A}^*(t)\, u = \tilde{f}(t, \nabla u) + (\mathscr{A}^*(t) - \mathscr{A}(t))\, u,$$

wobei

$$\mathscr{A}^*(t)\, u = \sum_{i,j=1}^{n} a_{ij}^*(t, x) \frac{\partial^2 u}{\partial x_i \partial x_j}, \quad a_{ij}^*(t, x) \in \mathbb{R},$$

sei mit $a_{ij}^* \in C^1([0, T] \times \bar{\Omega})$, die die a_{ij} hinreichend genau approximieren, etwa mit $|||| a_{ij}^* |||| \leqq |||| a_{ij} ||||$. Nach Satz 3 ist (vgl. [2], Theorem 7.2, Kap. III)

$$|||u|||_{1+\alpha} \leqq c\left(T, p, M, \|\varphi\|_{2,p}, |||g|||_{2,p}, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{0,p}, |||u|||, K, ||||A_{\tilde{\alpha}}||||\right),$$

woraus unser Satz folgt.

Bemerkung 1 zum Beweis von Satz 3, 4. Falls $K \neq 0$ ist in Satz 3, so setze man die Existenz einer Lösung $\mathbf{u}^0$ des parabolischen Systems $\frac{d}{dt}\mathbf{u}^0 + \mathscr{A}_p(t)\,\mathbf{u}^0 = f(t, \mathbf{u}^0, D^1\mathbf{u}^0, \ldots, D^{2m-1}\mathbf{u}^0)$ zu den Daten $\boldsymbol{\varphi}^0$, $\mathbf{g}^0(t)$ voraus mit den zu Beginn dieses Kapitels angegebenen Eigenschaften. Die Schranke für $|||\mathbf{u}|||_{2m-1+\alpha}$ hängt dann zusätzlich von K, $|||\mathbf{u}^0|||_{2m-1+\alpha}$, $|||\mathbf{g}^0|||_{2,p}$ bzw. $|||\mathbf{g}^0|||_{2m}$ ab. Der Beweis bleibt unverändert, nur hat man die Probleme mit den Daten $\boldsymbol{\varphi}^0 + \sigma(\boldsymbol{\varphi} - \boldsymbol{\varphi}^0)$, $\mathbf{g}^0(t) + \sigma(\mathbf{g}(t) - \mathbf{g}^0(t))$ zu betrachten. Weiter haben wir benutzt, daß wir bei Satz 3 und 2 mit der Voraussetzung $\mathbf{g}(.) \in C^\alpha([0, T], (H^{2,p}(\Omega))^L) \cap C^{1-\alpha}([0, T], (L^p(\Omega))^L)$, $\boldsymbol{\varphi} \in (H^{2,p}(\Omega))^L$, auskommen $(m=1)$. — Im Fall $m=1$, $L=1$ läßt sich in der Tat ohne Einschränkung $u_0 = 0$ voraussetzen. Man kann dann nämlich die Gleichungsschar

$$\frac{du}{dt} - \sum_{i,j=1}^{n} a_{ij}(t, x)\frac{\partial^2 u}{\partial x_i \partial x_j} = f(t, x, u, \nabla u) + \sigma f(t, x, 0, 0)$$
$$- f(t, x, 0, 0), \qquad 0 \leqq \sigma \leqq 1,$$

mit verschwindenden Daten und mit $-\frac{\partial f}{\partial u} \geqq 0$ genau analog zum Beweis von Satz 3 behandeln (s. [2], Kap. I, Theorem 2.1 und 2.9 sowie Theorem 9.1 in Kap. IV, [2], und Kap. IV. B., [5]).

II. Quasilineare Gleichungen zweiter Ordnung in zwei Raumdimensionen

Es sei $n=2$ und Ω ein beschränktes Gebiet des $\mathbb{R}^2$, dessen Rand von der Klasse C^4 sei (vgl. Kap. I). Es sei

$$f\colon\ \bar{\Omega} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$$

eine zweimal stetig differenzierbare Abbildung, die bezüglich der letzten 3 Variablen dreimal stetig differenzierbar sei. Weiter seien zweimal stetig differenzierbare Abbildungen

$$a_{ij}\colon\ \bar{\Omega} \times \mathbb{R}^2 \to \mathbb{R}, \qquad 1 \leqq i, j \leqq 2,$$

gegeben, die in den letzten beiden Variablen dreimal stetig differenzierbar seien und einer Ungleichung

$$\begin{gathered}\nu\left(1+|(z_2,z_3)|\right)^{\tilde m-2}|\xi|^2\\ \leqq \sum_{i,j=1}^{2} a_{ij}(x,z_2,z_3)\,\xi_i\,\xi_j \leqq \mu\left(1+|(z_2,z_3)|\right)^{\tilde m-2}|\xi|^2, \qquad (6)\\ \tilde m \geqq 2,\ \nu,\mu \text{ positive Konstanten},\end{gathered}$$

sowie

$$\begin{gathered}\sum_{i,j=1}^{2}\left\{\left|\frac{\partial a_{ij}}{\partial z_2}(x,z_2,z_3)\right|+\left|\frac{\partial a_{ij}}{\partial z_3}(x,z_2,z_3)\right|\right\} \qquad (6\text{a})\\ \leqq \mu(1+|z_2|+|z_3|)^{1-\varepsilon_0+\tilde m-2}\end{gathered}$$

mit $\varepsilon_0>0$. Mit einer monoton wachsenden stetigen Funktion $\mu_0 \geqq 0$ sei

$$-\frac{\partial f}{\partial z_1}(t,x,z_1,z_2,z_3) \geqq -\mu_0(|z_1|), \qquad (6\text{b})$$

$$\left|\frac{\partial f}{\partial t}(t,x,z_1,z_2,z_3)\right| \leqq \mu_0(|z_1|), \qquad f(0,x,0,0,0)\equiv 0 \qquad (6\text{c})$$

für alle $t\in\mathbb{R}^+$. Bei (6), (6a) bis (6c) ist natürlich wie auch später – $(x,z_1,z_2,z_3,\xi)\in\bar\Omega\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2$.

Sei $g(.)\in C^1([0,T],C^3(\bar\Omega))\cap C^2([0,T],C^1(\bar\Omega))$ reell. Sei $u\in C^2(]0,T[\times\Omega)\cap C^1([0,T]\times\bar\Omega)\cap C^0([0,T],C^2(\bar\Omega))\cap C^1(]0,T[,C^2(\bar\Omega))$. u genüge in $]0,T[\times\Omega$ der parabolischen Differentialgleichung

$$\frac{\partial u}{\partial t}+\mathscr{A}(u)\,u=f(t,x,u,\nabla u) \qquad (7)$$

mit den Randbedingungen

$$u(t)|\partial\Omega=g(t), \qquad (8)$$

$$u(0)=g(0), \quad \frac{\partial g}{\partial t}(0,x)=(D^{\tilde\alpha}g)(0,x)=0=f(0,x,g(0,x),\nabla g(0,x)), \quad x\in\partial\Omega, \quad |\tilde\alpha|\leqq 2. \qquad (9)$$

Hierbei ist

$$\mathscr{A}(u)\,u=-\sum_{i,j=1}^{2} a_{ij}(x,\nabla u)\frac{\partial^2 u}{\partial x_i\,\partial x_j}$$

gesetzt. Unser Ziel ist, eine Abschätzung von $|||u|||_{1+\alpha}$ unter der Bedingung (K eine Konstante)

$$|f(t,x,z_1,z_2,z_3)| \leqq \mu_0(|z_1|)\,|(z_2,z_3)|^{\tilde m}+K \qquad (10)$$

an f sowie der Verwendung einer Schranke

$$M_1 \geqq |||u||| \tag{10a}$$

herzuleiten.

Hierzu gilt der vorbereitende

Satz 5. *Seien* (6), (6a) *bis* (6c), (10), (10a) *erfüllt. Dann gilt für eine Lösung u von* (7), (8) *und* (9) *die Abschätzung*

$$|||u|||_{1+\alpha} \leqq c\left(\nu, \mu, \hat{\mu}_0, M_1, T, |||g|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}, K\right)$$

mit $0<\alpha<1$, $\hat{\mu}_0 := \mu_0(M_1)$, *wenn nur* $|||u|||\ \mu_0(|||u|||)$ *hinreichend klein ist. Die Schranke hängt nur von* M_1, ν, μ, $\partial\Omega$, *nicht aber von* K *ab.*

Beweis. Wir differenzieren (7) nach t. Nach dem Maximumprinzip für parabolische Gleichungen ([2], Theorem 2.1, Kap. I), angewandt auf $\frac{\partial u}{\partial t}$, ist $\left(\left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega} := \sup_{0\leqq t\leqq T}\left\|\frac{\partial g}{\partial t}(t)\right\|_{\partial\Omega}\right)$

$$\left|\left|\left|\frac{\partial u}{\partial t}\right|\right|\right| \leqq c\left(\mu, \hat{\mu}_0, T, |||g|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}\right).$$

$\left(\text{Man beachte, daß } \frac{\partial u}{\partial t}(t,x) = \frac{\partial g}{\partial t}(t,x) \text{ ist, } (t,x)\in[0,T]\times\partial\Omega\right)$. Hieraus folgt (o. E. sei $\tilde{m}=2$):

$$\|\mathscr{A}(u(t))\,u(t)\|_{L^2(\Omega)} \leqq \|f(t, u(t), \nabla u(t))\|_{L^2(\Omega)} + Tc\left(\mu, \tilde{\mu}_0, T, |||g|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}\right)$$

und hieraus, wie in [3], S. 381, 380, die gewünschte Abschätzung.

Weiter haben wir

Satz 6. *Sei* (6), (6a) *bis* (6c), (10) *und* (10a) *erfüllt. Die parabolische Gl.* (7) *mit den Randbedingungen* (8) *und* (9) *besitze eine Lösung u mit den vorhin angegebenen Eigenschaften.*

Dann besitzt (7) *mit* (8) *und* (9) *für jedes reelle*

$$\tilde{g}(.)\in C^1([0,T], C^3(\bar{\Omega})) \cap C^2([0,T], C^1(\bar{\Omega}))$$

mit

$$\frac{\partial \tilde{g}}{\partial t}(0,x) = D^{\tilde{\alpha}}\tilde{g}(0,x) = 0, \qquad |\tilde{\alpha}|\leqq 2, \quad x\in\partial\Omega,$$

eine Lösung ũ mit den vorhin angegebenen Eigenschaften. Überdies gilt eine Abschätzung

$$\left|\left|\left|\frac{\partial \tilde{u}}{\partial t}\right|\right|\right| \leqq c\left(\mu, \hat{\mu}_0, T, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}\right),$$

$$|||\tilde{u}|||_{1+\alpha} \leqq c\left(\nu, \mu, \hat{\mu}_0, \varepsilon_0, M_1, T, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}, K\right).$$

Beweis. Der Beweis kann so durchgeführt werden wie der von Satz 3 in Kap. I: Wir betrachten dazu das Problem

$$\frac{\partial \tilde{u}_\sigma}{\partial t} + \mathscr{A}(\tilde{u}_\sigma)\,\tilde{u}_\sigma = f(t, x, \tilde{u}_\sigma, \nabla \tilde{u}_\sigma)$$

mit

$$\tilde{u}_\sigma(t)\,|\,\partial\Omega = g(t) + \sigma(\tilde{g}(t) - g(t))$$
$$\tilde{u}_\sigma(0) = g(0) + \sigma(\tilde{g}(0) - g(0)), \qquad 0 \leqq \sigma \leqq 1.$$

Für die Differenz zweier Lösungen gilt nach dem Maximumprinzip

$$|||\tilde{u}_{\sigma_1} - \tilde{u}_{\sigma_2}||| \leqq |\sigma_1 - \sigma_2|\,\{|||\tilde{g} - g|||_{\partial\Omega} + \|\tilde{g}(0) - g(0)\|\}\, e^{\hat{\mu}_0 T}.$$

Weiter ist

$$\left|\left|\left|\frac{\partial \tilde{u}_{\sigma_1}}{\partial t}\right|\right|\right| \leqq c\left(\mu, \hat{\mu}_0, T, |||g|||_2, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}\right)$$

sowie

$$\left|\left|\left|\frac{\partial \tilde{u}_{\sigma_2}}{\partial t}\right|\right|\right| \leqq c\left(\mu, \hat{\mu}_0, T, |||g|||_2, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}\right).$$

Wenn vorausgesetzt wird, daß

$$|||\tilde{u}_{\sigma_1}|||_{1+\alpha} \leqq c\left(\nu, \mu, \hat{\mu}_0, \varepsilon_0, M_1, T, \sigma_1, |||g|||_2, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}, K\right)$$

ist, so folgt, daß

$$|||\tilde{u}_{\sigma_1}|||_{2,\,p} \leqq c\left(\nu, \mu, \hat{\mu}_0, \varepsilon_0, M_1, T, \sigma_1, p, |||g|||_2, |||\tilde{g}|||_2, \left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}, \left|\left|\left|\frac{\partial \tilde{g}}{\partial t}\right|\right|\right|_{\partial\Omega}, K\right)$$

ist, $p \geqq 2$. Nun ist

$$-\frac{\partial \tilde{u}_{\sigma_1}}{\partial t} + \frac{\partial \tilde{u}_{\sigma_2}}{\partial t} + \mathscr{A}(\tilde{u}_{\sigma_2})(\tilde{u}_{\sigma_2} - \tilde{u}_{\sigma_1}) = -(\mathscr{A}(\tilde{u}_{\sigma_2}) - \mathscr{A}(\tilde{u}_{\sigma_1}))\,\tilde{u}_{\sigma_1}$$
$$+ f(t, x, \tilde{u}_{\sigma_2}, \nabla \tilde{u}_{\sigma_2}) - f(t, x, \tilde{u}_{\sigma_1}, \nabla \tilde{u}_{\sigma_1}).$$

Den letzten Term rechts behandeln wir wie im Beweis von Satz 3. Auf den ersten Term rechts wenden wir unsere Voraussetzung an. Es ist (o. E. sei $\tilde{m} = 2$)

$$|a_{ij}(x, \nabla \tilde{u}_{\sigma_2}) - a_{ij}(x, \nabla \tilde{u}_{\sigma_1})| \left|\frac{\partial^2 \tilde{u}_{\sigma_1}}{\partial x_i \partial x_j}(x)\right|$$
$$= \left|\int_0^1 \frac{d}{d\tilde{t}}\left(a_{ij}(x, \nabla \tilde{u}_{\sigma_1} + \tilde{t}(\nabla \tilde{u}_{\sigma_2} - \nabla \tilde{u}_{\sigma_1})\right) d\tilde{t}\right| \left|\frac{\partial^2 \tilde{u}_{\sigma_1}}{\partial x_i \partial x_j}(x)\right|,$$

$$=\left|\sum_{k=2}^{3}\int_0^1\left(\frac{\partial}{\partial z_k}a_{ij}\right)(x,\nabla\tilde u_{\sigma_1}+\tilde t(\nabla\tilde u_{\sigma_2}-\nabla\tilde u_{\sigma_1}))\,d\tilde t\right.$$
$$\left.\cdot\left(\frac{\partial}{\partial x_{k-1}}\tilde u_{\sigma_2}-\frac{\partial}{\partial x_{k-1}}\tilde u_{\sigma_1}\right)\right|\cdot\left|\frac{\partial^2\tilde u_{\sigma_1}}{\partial x_i\partial x_j}(x)\right|,$$
$$\leqq 2\mu(1+|\nabla\tilde u_{\sigma_1}|^{2-\varepsilon_0}+|\nabla\tilde u_{\sigma_2}-\nabla\tilde u_{\sigma_1}|^{2-\varepsilon_0})\left|\frac{\partial^2\tilde u_{\sigma_1}}{\partial x_i\partial x_j}(x)\right|$$
$$\leqq 2\mu(1+\|\nabla\tilde u_{\sigma_1}\|^{2-\varepsilon_0})\left|\frac{\partial^2\tilde u_{\sigma_1}}{\partial x_i\partial x_j}(x)\right|+\frac{2\mu(2-\varepsilon_0)}{2}|\nabla u_{\sigma_2}-\nabla u_{\sigma_1}|^2$$
$$+\frac{2\varepsilon_0\mu}{2}\left|\frac{\partial^2\tilde u_{\sigma_1}}{\partial x_i\partial x_j}(x)\right|^{2/\varepsilon_0}.$$

Der erste Term links ist bereits abgeschätzt. Wie in [3], S. 381, 380, folgt:

$$|||\tilde u_{\sigma_2}-\tilde u_{\sigma_1}|||_{1+\alpha}$$
$$\leqq c\left(\nu,\mu,\hat\mu_0,\varepsilon_0,M_1,T,\sigma_1,|||g|||_2,|||\tilde g|||_2,\left\|\left|\frac{\partial g}{\partial t}\right\|\right|_{\partial\Omega},\left\|\left|\frac{\partial\tilde g}{\partial t}\right\|\right|_{\partial\Omega},K\right)$$

falls $|\sigma_2-\sigma_1|$ hinreichend klein ist, wobei die Schranke von σ_1 unabhängig ist.

Hieraus und Satz 5 erhält man wie beim Beweis von Satz 3 in Kap. I die gewünschte a-priori Schranken für $|||\tilde u_\sigma|||_{1+\alpha}$, $\sigma\in[0,1]$. Es ist hierzu zu bemerken, daß für eine Folge $\{\sigma_\nu\}\subset[0,1]$ mit $\sigma_\nu\to\sigma_0$, bei der die Ausdrücke $|||\tilde u_{\sigma_\nu}|||_{1+\alpha}$ beschränkt bleiben

$$|||\nabla\tilde u_{\sigma_{\nu_\mu}}-\nabla v_{\sigma_0}|||_{\rho,Q_T}\to 0,$$
$$\|\nabla\tilde u_{\sigma_{\nu_\mu}}-\nabla v_{\sigma_0}\|_\rho\to 0,\qquad \rho>0$$

und hinreichend klein,

$$\left\|\left|\frac{\partial\tilde u_{\sigma_{\nu_\mu}}}{\partial t}-\frac{\partial v_{\sigma_0}}{\partial t}\right\|\right|_{\tilde\alpha}+|||\tilde u_{\sigma_{\nu_\mu}}-v_{\sigma_0}|||_{2+\tilde\alpha}\to 0,\qquad \alpha>\tilde\alpha>0,$$

mit einer Teilfolge σ_{ν_μ} und einem Element

$$v_{\sigma_0}\in C^0([0,T],C^{2+\tilde\alpha}(\bar\Omega))\cap C^{1+\tilde\alpha/2}([0,T]\times\bar\Omega).$$

Dies sieht man mit Lemma 3.1, Kap. II, in [2], bereits in [5] verwendeten Argumenten ((9) und (10)) und Theorem 5.2 in [2], Kap. IV. Die Betrachtung der linearen Differentialgleichung für den Differenzenquotienten

$$\frac{v_{\sigma_0}(t+h,x)-v_{\sigma_0}(t,x)}{h}$$

sowie die zur $L^p(\Omega)$-Theorie gehörige Integralgleichung und [5], Kap. IV.A, zeigen, daß $v_{\sigma_0} \in C^2(]0, T[\times \Omega)$ ist. Den noch erforderlichen Störungssatz kann sich der Leser anhand des Lemmas 3.1 in Kap. II und Theorems 5.2 in Kap. IV, [2], leicht selbst überlegen.

Eine einfachere Variante unseres Beweises zeigt, daß das entsprechende Problem für elliptische Gleichungen

$$\sum_{i,j=1}^{2} a_{ij}(x, \nabla u) \frac{\partial^2 u}{\partial x_i \partial x_j} = f(x, u, \nabla u) + a(x)\, u, \qquad a \in C^\alpha(\bar{\Omega}),\ a(x) > 0,\ x \in \bar{\Omega},$$

unter den Voraussetzungen (6), (6a) und (10) gelöst werden kann. Hierzu geht man zunächst wieder vom Fall $\partial f/\partial u \geqq 0$ aus. Unter der Annahme, daß zu den Randwerten 0 eine Lösung existiert, kann man dann so vorgehen wie beim Beweis von Satz 6. Von der Annahme, daß zu den Randwerten 0 eine Lösung existiert, befreit man sich, indem man die Gleichung

$$-\sum_{i,j=1}^{2} a_{ij}(x, \nabla u) \frac{\partial^2 u}{\partial x_i \partial x_j} + a(x)\, u = -f(x, u, \nabla u) + \sigma f(x, 0, 0) - f(x, 0, 0).$$

$$0 \leqq \sigma \leqq 1$$

betrachtet und von u fordert, daß u auf $\partial\Omega$ verschwindet. Für die Differenz zweier Lösungen u_{σ_1} und u_{σ_2} zu zwei Parameterwerten $\sigma_1, \sigma_2 \in [0, 1]$ gilt bekanntlich das Maximumprinzip in der Form

$$\|u_{\sigma_2} - u_{\sigma_1}\|_0 \leqq \frac{|\sigma_1 - \sigma_2|}{\varepsilon} \|f(x, 0, 0)\|_0, \qquad a(x) \geqq \varepsilon > 0,\ x \in \bar{\Omega},$$

so daß sich wieder die Kontinuitätsmethode anwenden läßt, die eine (von ε abhängige) Schranke liefert. Von der Bedingung $\partial f/\partial u \geqq 0$ befreit man sich, indem man als Inhomogenität einer neuen Gleichung

$$\frac{f(x, u, \nabla u)}{\mu_0(M_1)\, |\nabla u|^2 + K} (\mu_0(M_1)\, |p|^2 + K) = g(x) (\mu_0(M_1)\, |p|^2 + K), \qquad p \in \mathbb{R}^2,$$

betrachtet. (Im elliptischen Fall genügen erheblich schwächere Regularitätsvoraussetzungen an die a_{ij} und f, außerdem sei ohne Einschränkung $\tilde{m} = 2$.) Falls a in einem Punkt von $\bar{\Omega}$ verschwindet, so kann man die quasilineare elliptische Differentialgleichung unter der Voraussetzung $K = 0$ entsprechend behandeln.

Unser Resultat hinsichtlich quasilinearer parabolischer Differentialgleichungen fassen wir in dem folgenden Satz zusammen:

Satz 7. *Es seien die Funktionen a_{ij} sowie f in $\mathbb{R}^+ \times \bar{\Omega} \times \mathbb{R}^3$ zweimal stetig nach t, x und nach den restlichen Variablen dreimal stetig differenzierbar. Außerdem mögen sie den Ungleichungen* (6), (6a)–(6c) *und* (10) *genügen. Weiter sei*

$$u(.)\in C^0([0,T],H^{2,p}(\Omega))\cap C^1([0,T],L^p(\Omega)),\qquad p>\frac{4}{1-\alpha},$$

$$\frac{du}{dt}-\sum_{i,j=1}^{2}a_{ij}(x,\nabla u(t))\frac{\partial^2 u}{\partial x_i\,\partial x_j}(t)=f(t,u(t),\nabla u(t)),\qquad 0\leqq t\leqq T,$$

$$u(t,x)\,|\,\partial\Omega=g(t,x),\qquad t\in[0,T]\times\partial\Omega,$$

$$u(0)=g(0),\qquad \left(\frac{\partial g}{\partial t}\right)(0,x)=0=(D^{\tilde\alpha}g)(0,x),\qquad x\in\partial\Omega,\ |\tilde\alpha|\leqq 2,$$

mit einer Funktion $g(.)\in C^1([0,T],C^3(\bar\Omega))\cap C^2([0,T],C^1(\bar\Omega))$. *Dann ist* du/dt *stetig in* $[0,T]\times\bar\Omega$ *und*

$$\left|\left|\left|\frac{du}{dt}\right|\right|\right|\leqq c\left(\mu,\hat\mu_0,T,|||g|||_2,\left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega}\right),$$

$$|||u|||_{1+\alpha}\leqq c\left(v,\mu,\hat\mu_0,\varepsilon_0,|||u|||,T,|||g|||_2,\left|\left|\left|\frac{\partial g}{\partial t}\right|\right|\right|_{\partial\Omega},K\right).$$

Beweis. Man zeige so wie im Beweis von Satz 6, daß u die zu Beginn des Kapitels angegebenen Regularitätsvoraussetzungen erfüllt, wobei man Lemma 3.3 in [2], Kap. II, statt Lemma 3.1 benutze.

Endlich kann man statt $g(0)$ auch einen Anfangswert φ mit $(D^{\tilde\alpha}\varphi)\,|\,\partial\Omega=0$ vorschreiben, $|\tilde\alpha|\leqq 2$.

Nachtrag bei der Korrektur. Dem Autor ist es gelungen, sich von der Voraussetzung (6a) zu befreien. Wie bei Bemerkung 1 zu Satz 3 und 4 sieht man, daß man bei Satz 6 o.E. $u=0$ voraussetzen kann.

Literatur

1. Friedman, A.: Partial differential equations. New York-Chikago-San Francisco-Atlanta-Dallas-Montreal-Toronto-London-Sydney: Holt, Rinehart and Winston 1969.
2. Ladyzenskaja, O. A., Solonnikov, V. A., Uralceva, N. N.: Linear and quasilinear equations of parabolic type. Translations of Mathematical Monographs 23. Providence, R. I.: Amer. Math. Soc. 1968.
3. Ladyzhenskaja, O. A., Ural'tseva, N. N.: Linear and quasilinear elliptic equations. New York and London: Academic Press 1968.
4. Tomi, F.: Über semilineare elliptische Differentialgleichungen zweiter Ordnung. Math. Z. **111**, 350–366 (1969).
5. Wahl, W. v.: Gebrochene Potenzen eines elliptischen Operators und parabolische Differentialgleichungen in Räumen hölderstetiger Funktionen. Erscheint in den Nachr. Akad. Wiss. Göttingen.

Prof. W. v. Wahl
Mathematisches Institut der Universität
D-5300 Bonn, Wegelerstraße 10
Deutschland

(Eingegangen am 20. März 1972)

Math. Z. 128, 109–113 (1972)

Eine abstrakte Definition gewisser alternierender Gruppen

Hartmut Saß

1. Darstellungen der alternierenden Gruppen A_n durch Erzeugende und definierende Relationen sind lange bekannt ([1, 3]). Bei allen diesen Darstellungen wächst die Anzahl der Relationen mit dem Grade n. Die kürzeste Darstellung mit 2 Erzeugenden und $[n/2]+1$ Relationen findet man bei Coxeter ([2]). Es scheint aber keine abstrakte Definition bekannt zu sein, die mit einer bei variablem n festen Anzahl von Relationen auskäme. Wir können in dieser Note zeigen, daß man eine solche Darstellung für die A_{p+2} ($p>2$, p Primzahl) leicht angeben kann. Dazu betrachte man die A_{p+2} auf einer speziellen Menge, der Vereinigung des endlichen Körpers F_p mit einer 2elementigen Menge, operierend. Dann läßt sich ein Element der Ordnung $(p-1)/2$ angeben mit dessen Hilfe sich die Carmichaelsche Darstellung ([3]) auf eine feste Anzahl von Relationen reduzieren läßt. Wir wollen nicht diesen kürzesten Weg zum Ergebnis, dem Satz des 3. Abschnitts, einschlagen, sondern bevorzugen einen allgemeineren Rahmen. Wir betten die 2dimensionalen projektiven Gruppen $L_2(p)$, dargestellt als 2fach transitive Permutationsgruppe vom Grade $p+1$, in die A_{p+2} ein und gewinnen damit noch ein interessantes Ergebnis, das im 4. Abschnitt ausgesprochen wird. Diese spezielle Einbettung wird im 2. Abschnitt beschrieben. Alle Abkürzungen sind Standard (z.B. x^y für $y^{-1}xy$), kleine griechische Buchstaben bezeichnen Elemente des endlichen Körpers F_p.

2. Die über dem endlichen Körper F_p ($\mathsf{Z}/p\,\mathsf{Z}$) projektiven 2dimensionalen Gruppen $L_2(p)$, diese „kleinsten“ minimalen einfachen Gruppen mit übersichtlicher Struktur, deren Untergruppenverbände man z.B. vollständig kennt, besitzen eine Darstellung als 2fach transitive Gruppe auf den $p+1$ Punkten der projektiven Geraden oder auf den $p+1$ Nebenklassen nach der maximalen Untergruppe $H=[Z_p]\,Z_{(p-1)/2}$, dem semidirekten Produkt zweier zyklischer Gruppen. Ausgehend von folgender abstrakter Definition der $L_2(p)$, die auf Frasch ([5]) zurückgeht, wollen wir diese Darstellung jetzt ableiten.

$$\langle A, B, C;\ C^p=B^{(p-1)/2}=A^2=(AB)^2=(AC)^3=C^B C^{-\alpha^2}=(ABC^\alpha)^3=1\rangle \quad (1)$$

(α primitive Wurzel mod p, die letzte Relation ist im Falle -1 quadratischer Nichtrest, d.h. bei $p \equiv -1 \pmod 4$, entbehrlich).

Eine sehr verwandte Darstellung findet man bei Todd ([6]). Die $L_2(p)$, mit der Multiplikation als Verknüpfung, ist die Gruppe aller 2reihigen Matrizen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $a, b, c, d \in F_p$ $(ad-bc)$ ein von Null verschiedenes Quadrat, bei denen die Matrizen, die sich nur um einen Skalarfaktor unterscheiden, identifiziert werden. Dabei bilden die folgenden A, B, C entsprechenden 3 Matrizen ein Erzeugendensystem:

$$A \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B \to \begin{pmatrix} \alpha^2 & 0 \\ 0 & 1 \end{pmatrix}, \quad C \to \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \tag{2}$$

(α primitive Wurzel mod p, d.h. erzeugendes Element der multiplikativen zyklischen Gruppe Z_{p-1} des Körpers F_p).

Die Relationen von (1) sind leicht zu verifizieren. Aus den letzten 3 Relationen läßt sich B eliminieren, die Gruppe $L_2(p)$ wird mithin bereits von 2 Elementen erzeugt. Eine Darstellung als Permutationsgruppe vom Grade $p+1$ erhalten wir durch Anwendung der unter (2) gegebenen linearen Transformationen auf den Vektor aus den homogenen Koordinaten der $p+1$ Punkte der projektiven Geraden über F_p. Wir bekommen dadurch 3 erzeugende Permutationen vom Grade $p+1$. Dabei operiere die Matrix von rechts auf den Zeilenvektoren $(r\ 1)$ bzw. $(1\ 0)$. Wir bezeichnen sowohl die Elemente von F_p wie die Punkte mit den homogenen Koordinaten $(r\ 1)$ $(r \in F_p)$ mit $0, 1, \ldots, p-1$, den uneigentlichen Punkt der Geraden mit der homogenen Koordinate $(1\ 0)$ mit j. Die Gruppe $L_2(p)$ operiert also auf der Menge $\Sigma = \{j\} \cup F_p = \{j, 0, 1, \ldots, p-1\}$, $|\Sigma| = p+1$.

Die erzeugenden Permutationen lauten dann:

$$\begin{array}{ll} a\colon\ \alpha^i \leftrightarrow -\alpha^{-i},\ 0 \leftrightarrow j & \\ & \alpha^i \in F_p^*,\ \alpha \text{ prim. Element} \\ b\colon\ \alpha^i \to \alpha^{2+i},\quad 0 \to 0,\ j \to j & \\ c\colon\ x \to x+1,\quad j \to j & x \in F_p,\ F_p^* = F_p \setminus \{0\} \\ & (i = 1, \ldots, p-1)^1. \end{array} \tag{3}$$

Auch hier bestätigt man leicht die Relationen von (1) bei der Abbildung $A \to a$, $B \to b$, $C \to c$.

Wir betten diese Permutationsgruppe vom Grade $p+1$ in eine vom Grade $p+2$ ein, den permutierten Objekten ∞, den Permutationen ein

[1] Wir müßten korrekterweise zwischen den Restklassen, den Elementen von F_p, die wir mit $0, 1, \ldots, p-1$ bezeichnen und den Exponenten des prim. Elements, die wir ebenso bezeichnen, unterscheiden. Wir unterlassen es, da keine Mißverständnisse zu befürchten sind.

involutorisches Element d hinzufügend.

$$\Omega=\{\infty, j\} \cup F_p, \qquad \Omega=\{\infty, j, 0, 1, \ldots, p-1\}, \qquad |\Omega|=p+2. \tag{4}$$

Dabei sollen die Permutationen a, b, c ∞ festlassen auf den übrigen Punkten wie unter (3) operieren, d aber, wie folgt, definiert sein:

$$d:\ x \rightarrow -x, \quad \infty \leftrightarrow j \qquad x \in F_p.$$

Diese Permutationsgruppe (Γ_p, Ω), $\Gamma_p=\langle a, b, c, d\rangle$, ist also folgendermaßen definiert:

$$\begin{array}{lll}
a:\ \alpha^i \leftrightarrow -\alpha^{-i}, \quad 0 \leftrightarrow j, \quad \infty \rightarrow \infty & & \\
 & \alpha^i \in F_p^* \ (i=1, \ldots, p-1) & \\
b:\ \alpha^i \rightarrow \alpha^{2+i}, \quad 0 \rightarrow 0, \quad j \rightarrow j, \quad \infty \rightarrow \infty & & \\
 & & (5) \\
c:\ x \rightarrow x+1, \quad \infty \rightarrow \infty, \quad j \rightarrow j, & & \\
 & x \in F_p. & \\
d:\ x \leftrightarrow -x, \quad \infty \leftrightarrow j & &
\end{array}$$

Es ist $t=(d a)^2=(\infty\, j\, 0)$. Außerdem enthält die Gruppe Γ_p einen Zykel der Länge p, $c=(\infty)(j)(0\, 1 \ldots p-1)$. Da die erzeugenden Permutationen alle gerade sind, ist die Gruppe $\Gamma_p=A_{p+2}$. Dabei wird die Gruppe bereits von den Elementen t und c erzeugt. Man bestätigt leicht die folgenden Relationen:

$$t^b=t, \qquad (t c^{-i} t c^i)^2=1 \quad (i=1, 2, \ldots, p-1), \qquad b^d=b, \quad t^3=(d a)^6=1 \tag{6}$$

einmal wegen der schon in $L_2(p)$ geltenden Relation $(a b)^2=1$, zum anderen wegen $c^{-i} t c^i=(\infty\, j\, i)$ $(i=1, 2, \ldots, p-1)$ und $(\infty\, j\, 0)(\infty\, j\, i)=(\infty\, i)(j\, 0)$. Die Untergruppe $\langle b\rangle<\Gamma_p$ operiert auf 3 Bahnen der Länge 1 und 2 Bahnen der Länge $(p-1)/2$, von denen die eine die quadratischen Reste, die andere die quadratischen Nichtreste mod p enthält. Es ist $(\infty\, \mu)(\infty\, \nu) \cdot b=(0)(j)(\infty \ldots \mu \ldots \nu)$ $((\mu \nu/p)=-1)$. $(\infty\, \mu)(\infty\, \nu)=(\infty\, \mu\, \nu)=(\infty\, j\, \nu)(\infty\, j\, \mu)(\infty\, j\, \nu)(\infty\, j\, \nu)$, d.h. es gilt auch die Relation:

$$(c^{-\nu} t c^{\nu-\mu} t c^{\mu-\nu} t^2 c^\nu b)^p=1. \tag{7}$$

3. Wir behaupten, die alternierenden Gruppen A_{p+2} besitzen folgende abstrakte Definition:

Satz.

$$\begin{aligned} A_{p+2} \cong G_p=\langle T, C, B;\ & C^p=T^3=B^{(p-1)/2}=(TC)^{p+2}=(TC^{-1}TC)^2 \\ & =\underline{(TC^{-\lambda}TC^\lambda)^2}=C^B C^{-\alpha^2}=T^B T^{-1}=(C^{-\nu} TC^{\nu-\mu} TC^{\mu-\nu} T^2 C^\nu B)^p=1\rangle \end{aligned}$$

(λ, μ, ν *fest*, $(\mu \nu/p)=(\lambda/p)=-1$, α *primitive Wurzel* mod p, $p>2$ *prim, die unterstrichene Relation ist bei* $p \equiv -1$ (mod 4) *entbehrlich*).

Beweis. (i) Wie wir in 2. gesehen haben, sind alle Relationen der Gruppe G_p in der Permutationsgruppe $\Gamma_p \cong A_{p+2}$ bei der Abbildung

$T \to t$, $C \to c$, $B \to b$ erfüllt. A_{p+2} ist also homomorphes Bild von G_p, $G_p \xrightarrow{\sim} A_{p+2}$.

(ii) Carmichael ([3]) gibt für die A_{p+2} folgende Darstellung in 2 Erzeugenden:

$$\langle s, t;\ s^p = t^3 = (s\,t)^{p+2} = (t\,s^{-j}\,t\,s^j)^2 = 1\rangle \quad \left(1 \leqq j \leqq \frac{p-1}{2}\right). \tag{8}$$

Man bildet dabei $t \to (1\,2\,3)$, $s \to (3\,4 \ldots p+2)$ ab, sofern A_{p+2} auf der Menge $\{1, 2, \ldots, p+2\}$ operiert.

Aus den letzten 3 Relationen von G_p folgt $B \in \langle T, C\rangle$ $(B^p = B)$, die Gruppen $\langle T, C, B\rangle$ und $\langle T, C\rangle$ sind mithin identisch. Wir müssen nun zeigen, daß bei der Abbildung: $s \to C$, $t \to T$ die Relationen von (8) in der Gruppe G_p erfüllt sind. Für die ersten 3 Relationen von (8) ist das klar, für die übrigen unterscheiden wir 2 Fälle:

1) $(-1/p) = -1$. Aus der Relation $(TC^{-1}\,TC)^2 = 1$ folgt durch Transformation mit B^k $(k = 1, \ldots, (p-1)/2)$ $(TC^{-\beta}\,TC^{\beta})^2 = 1$ wegen der Relation $TB = BT$. Dabei durchläuft $\beta = \alpha^{2k}$ alle quadratischen Reste. Da die Nichtreste aber die negativen quadratischen Reste sind und mit $(TC^{-\beta}\,TC^{\beta})^2 = 1$ wegen

$$C^{-(p-\beta)}(TC^{-\beta}\,TC^{\beta})\,C^{p-\beta} = C^{\beta}\,TC^{-\beta}\,T = T^{-1}(TC^{\beta}\,TC^{-\beta})\,T$$

auch $(TC^{\beta}\,TC^{-\beta})^2 = 1$ gilt (s. Carmichael [3] p. 263), folgt die letzte Relation von (8) sogar für alle von 0 verschiedenen Exponenten von C.

2) $(-1/p) = 1$. In diesem Falle müssen wir die letzte Relation von (8) in G_p noch für einen bel. quadratischen Nichtrest als Exponenten von C fordern, um durch Transformation mit B^k die Gültigkeit dieser Relation wieder für alle quadratischen Nichtreste als Exponenten von C zu sichern, was durch Hinzufügen der unterstrichenden Relation in G_p geschehen ist. Damit ist nachgewiesen, daß G_p homomorphes Bild von A_{p+2} ist, $A_{p+2} \xrightarrow{\sim} G_p$. Aus (i) und (ii) folgt $A_{p+2} \cong G_p$, q. e. d.

4. Die Gruppen $G'_p = \langle a, b, c, d;\ a^2 = b^{(p-1)/2} = c^p = d^2 = (a\,b)^2 = (a\,c)^3 = c^b\,c^{-\alpha^2} = b^d\,b^{-1} = c^d\,c = (a\,b\,c^{\alpha})^3 = (a\,d)^6 = 1\rangle$ diese homomorphen Bilder des freien Produktes $L_2(p) * Z_2$ dürften von allgemeinem Interesse sein. Es sind nicht nur alternierende Gruppen, wie wir in 3. gesehen haben, bekannte homomorphe Bilder dieser Gruppen, sondern auch die kleinste Jankogruppe $J(11)$ als homomorphes Bild der G'_{11} (s. McKay [4]). Betrachtet man statt der Gruppe G'_p die Gruppe G''_p, die sich nur in der letzten Relation von G'_p unterscheiden soll, die wir durch $(d\,a)^3 = 1$ ersetzen, so muß zumindest für $p \geqq 13$ $G''_p = \langle e\rangle$ sein, da sich andernfalls die Gruppen $L_2(p)$ transitiv erweitern ließen. Eine Zerlegung von G''_p in Doppelnebenklassen nach der Untergruppe $L_2(p)$ ließe nur 2 solcher Klassen zu.

Literatur

1. Coxeter, H.S.M., Moser, W.O.J.: Generators and relations for discrete groups. Ergebnisse der Math. und ihrer Grenzgebiete Bd. 14 p. 66–67, 93–94, Berlin-Heidelberg-New York: Springer 1965.
2. Coxeter, H.S.M.: Abstract definitions for the symmetry groups of the regular polytopes in terms of two generators. Proc. Cambridge Philos. Soc. **33**, 315–324 (1937).
3. Carmichael, R.D.: Abstract definitions of the symmetric and alternating groups and certain other permutation groups. Quaterly J. Math. **49**, 226–270 (1923), insbesondere p. 262–264.
4. McKay, J.: The construction of the character table of a finite group from generators and relations. Computational Problems in Abstract Algebra. Proceedings of a conference held at Oxford (J. Leech Editor) p. 96 (1970).
5. Frasch, H.: Die Erzeugenden der Hauptkongruenzgruppen für Primzahlstufen. Math. Ann. **108**, 229–252 (1933).
6. Todd, J.A.: A note on the linear fractional group. J. London Math. Soc. **7**, 195–200 (1932).

Hartmut Saß
Hahn-Meitner-Institut
Sektor Mathematik
1 Berlin-39
Glienickerstraße 100
Deutschland

(Eingegangen am 29. März 1972)

Math. Z. 128, 115–128 (1972)

Zur Darstellungstheorie σ-vollständiger Vektorverbände

Wolfgang Hackenbroch

In dieser Arbeit wird die Darstellungstheorie eines σ-vollständigen Vektorverbandes E in weitgehender Analogie zu entsprechenden Resultaten von Schaefer über Banachverbände [5] entwickelt. An die Stelle der dort eingeführten (eindeutig bestimmten) Norm-erzeugenden vag kompakten Familie von Radonmaßen auf dem lokalkompakten Darstellungsraum tritt hier ein (ebenfalls eindeutig bestimmtes) E-wertiges Maß μ im Sinne von [7, 8].

Im zweiten Abschnitt wird E als ein Vektorverband $\hat{E}$ von stetigen numerischen Funktionen auf einem lokalkompakten Hausdorffraum, direkte Summe von Kompakten, realisiert, wobei $\hat{E}$ die stetigen reellen Funktionen mit kompakten Trägern als „dichtes" Ideal enthält (diese für einen Funktionalkalkül wichtige Eigenschaft ist bei entsprechenden Einbettungssätzen, etwa Theorem V, 4.2 von [6], nicht gewährleistet).

Im 3. Abschnitt führt die Integraldarstellung dieser Realisierung durch ein E-wertiges Maß μ mit Hilfe einer in [8] gegebenen Verallgemeinerung des Rieszschen Darstellungssatzes zunächst auf eine neue Darstellung von E durch den zugehörigen Raum $L^1(\mu)$ (Satz 4 und Korollar). Hierbei werden wichtige Prozesse in E sehr einfach in $L^1(\mu)$ repräsentiert, etwa die Ordnungskonvergenz durch majorierte Konvergenz punktweise lokal-μ-fast überall, oder Projektionen auf Hauptbänder durch Multiplikationen mit charakteristischen Funktionen.

Schließlich wird in Satz 5 der Zusammenhang mit der Freudenthalschen Spektraltheorie hergestellt.

1. Vorbemerkungen über σ-vollständige Vektorverbände

Wir bezeichnen in dieser Arbeit mit E stets einen σ-vollständigen (und also archimedisch geordneten) reellen Vektorverband mit positivem Kegel E_+; $\vee$ bzw. $\wedge$ bezeichnen das Verbandssupremum bzw. -infimum, $x_+ = x \vee 0$ bzw. $x_- = (-x)_+ = -x \wedge 0$ den positiven bzw. negativen Teil, $|x| = x_+ + x_-$ den Absolutbetrag von $x \in E$.

Für $u > 0$ aus E ist $E_u = \{x \in E : \exists\, \lambda > 0 \text{ mit } |x| \leqq \lambda u\}$ das von u erzeugte Ideal; E_u ist dann bezüglich der von E induzierten Ordnung wieder

σ-vollständig, außerdem vollständig bezüglich der durch $\|x\|_u := \inf\{\lambda > 0 \colon |x| \leqq \lambda u\}$ auf E_u definierten Norm $\| \ \|_u$, welche überdies die Eigenschaft $\|x \vee y\|_u = \max\{\|x\|_u, \|y\|_u\}$ für alle $x, y \in E$ hat, so daß $(E_u; \| \ \|_u)$ ein Kakutanischer M-Raum mit u als Ordnungseins ist.

Definition. Eine Abbildung $\varphi \colon E \to \overline{\mathbb{R}}_+ = [0, \infty]$ heißt eine *Bewertung* von E, wenn $\varphi(x) = \varphi(|x|)$ für alle $x \in E$ gilt und die Einschränkung $\varphi | E_+$ additiv, positiv-homogen und ein Verbandshomomorphismus ist (d.h. $\varphi(x \vee y) = \max\{\varphi(x), \varphi(y)\}$ und also auch $\varphi(x \wedge y) = \min(\varphi(x), \varphi(y))$ für $x, y \in E_+$; in $\overline{\mathbb{R}}$ gelten für das Rechnen mit $\pm\infty$ stets die üblichen Konventionen).

Nach [5], Lemma 1 gilt dann

Lemma 1. *Ist $u > 0$ in E, so läßt sich jede Bewertung φ von E_u mit $\varphi(u) = 1$ eindeutig zu einer Bewertung $\bar{\varphi}$ auf ganz E fortsetzen.*

Folgerung. $K_u := \{\varphi \colon E \to \overline{\mathbb{R}}_+ \colon \varphi$ Bewertung mit $\varphi(u) = 1\}$ ist offenbar kompakt als Teilraum des Cartesischen Produktes $(\overline{\mathbb{R}}_+)^E$. Die Restriktion auf $(E_u)_+ := E_u \cap E_+$ eines $\varphi \in K_u$ ist wegen $\varphi(u) = 1$ endlich und hat also eine eindeutig bestimmte Fortsetzung zu einem Verbandshomomorphismus $\varphi_0 \colon E_u \to \mathbb{R}$. Auf Grund von Lemma 1 definiert also $\varphi \mapsto \varphi_0$ eine Homeomorphie von K_u auf den schwach $*$-kompakten „Strukturraum" aller Verbandshomomorphismen $\varphi \colon E_u \to \mathbb{R}$ mit $\varphi(u) = 1$. Nach dem Kakutanischen Darstellungssatz ist demnach

$$x \mapsto \hat{x} \colon \hat{x}(\varphi) = \varphi_0(x) \left(= \varphi(x_+) - \varphi(x_-) \right)$$

ein M-Raum-Isomorphismus von E_u auf $\mathscr{C}(K_u)$, den (bezüglich punktweiser Ordnung und Sup-Norm) M-Raum aller stetigen $\mathbb{R}$-wertigen Funktionen auf K_u. Da E_u wieder σ-vollständig ist, ist K_u außerdem Quasi-Stone-Raum, d.h. die abgeschlossene Hülle jeder offenen F_σ-Menge ist offen ([6], S. 42).

Lemma 2. *Sei $U \subset E$ ein maximales System von paarweise disjunkten Elementen > 0 (d.h. $u \wedge v = 0$ für $u \neq v \in U$). Dann gilt*

1. *$E_U := \{x \in E \colon \exists\, n \in \mathbb{N}$ und $U' \subset U$ endlich mit $|x| \leqq \sum_{u \in U'} n u\} = \bigoplus_{u \in U} E_u$ ist das von U in E erzeugte Ideal.*

2. *E_U ist dicht in E in dem Sinne, daß für jedes $x \in E_+$ existiert und gilt*

$$x = \bigvee \{y \in E_U \colon 0 \leqq y \leqq x\} = \bigvee \Big\{ \sum_{u \in U'} (x \wedge n u) \colon n \in \mathbb{N};\ U' \subset U \text{ endlich} \Big\}.$$

Beweis. 1. Aus $u \wedge v = 0$ folgt $E_u \cap E_v = (0)$ für $u \neq v$ in U; die übrigen Aussagen sind offensichtlich.

2. Die Gleichheit der beiden Suprema ergibt sich aus $x \wedge (y+z) = x \wedge y + x \wedge z$ für $x, y, z \in E_+$ mit $y \wedge z = 0$ und dies aus der Rieszschen

Zerlegungseigenschaft. Bei beliebigem $z \in E_+$ mit $z \geqq y$ für jedes $y \in E_U$ mit $0 \leqq y \leqq x$ bleibt $z \geqq x$ zu zeigen. Andernfalls wäre aber $d := x - z \wedge x > 0$ und also wegen Maximalität von U auch $d \wedge u > 0$ für ein $u \in U$. Dann ist also für beliebiges $y \in E_U$ mit $0 \leqq y \leqq x$ einerseits $d \wedge u + y \leqq d \wedge u + z \wedge x \leqq d + z \wedge x = x$, andererseits $d \wedge u + y \in E_U$, also $z \geqq y + d \wedge u$. Daher wird $d' := x - (z - d \wedge u) \wedge x \geqq d$ sowie $d' \wedge u + y \leqq d' + (z - d \wedge u) \wedge x = x$ und $d' \wedge u + y \in E_U$ für jedes $y \in E_U$ mit $0 \leqq y \leqq x$, also $z - d \wedge u \geqq y + d' \wedge u \geqq y + d \wedge u$ usw., somit $z - n(d \wedge u) \geqq y$ für $n = 1, 2, \ldots$ und jedes solche y, im Widerspruch zur Archimedizität von E. ∎

In E betrachten wir die durch die Ordnung von E definierte Konvergenzstruktur $\lim\limits_{n \to \infty} x_n = x$ (kurz $x_n \to x$; bei isotonen Folgen auch $x_n \uparrow x$), wenn existiert und gilt

$$\limsup_{n \to \infty} x_n := \bigwedge_{k=1}^{\infty} \bigvee_{n=k}^{\infty} x_n = \liminf_{n \to \infty} x_n := \bigvee_{k=1}^{\infty} \bigwedge_{n=k}^{\infty} x_n = x .$$

Bekanntlich verhalten sich die Vektorverbandsoperationen stetig bezüglich dieser Konvergenz, ([4] Ch. I).

Zu E adjungieren wir gelegentlich ein abstraktes Element ∞ und definieren

$$\infty \geqq x; \; x + \infty = \infty \qquad \text{für jedes } x \in E .$$

Insbesondere schreiben wir für eine nach oben nicht majorierte Menge $A \subset E$ auch $\bigvee A = \infty$, und die Aussage $x_n \to \infty$ ist damit für eine Folge (x_n) in E durch $\liminf\limits_{n \to \infty} x_n = \infty$ definiert.

Für den im folgenden verwandten Maßbegriff vergleiche [7, 8].

Definition. Sei X lokalkompakter Hausdorffraum, $\Sigma = \Sigma(X)$ der von den kompakten G_δ-Mengen in X erzeugte σ-Ring der Baireschen Teilmengen von X. Eine Abbildung μ: $\Sigma \to E_+ \cup \{\infty\}$ heißt ein *Bairemaß*, wenn $\mu(\emptyset) = 0$, $\mu(A) \in E_+$ für $A \in \Sigma$ kompakt und $\mu\left(\bigcup\limits_{1}^{\infty} A_i\right) = \bigvee\limits_{n=1}^{\infty} \sum\limits_{i=1}^{n} \mu(A_i)$ für jede Folge von paarweise disjunkten Mengen $A_i \in \Sigma$ (also μ σ-additiv bezüglich der oben beschriebenen Konvergenzstruktur).

Lemma 3. *Sei X kompakt mit Σ der σ-Algebra der Bairemengen in X, μ: $\Sigma \to E_+$ ein Bairemaß; für eine Σ-Stufenfunktion $f = \sum \alpha_i \chi_{A_i}$: $X \to \mathbb{R}$ sei $\mu(f) := \sum \alpha_i \mu(A_i) \in E$. Dann gilt:*

1. *$\mu(f)$ ist unabhängig von der Darstellung von f und hängt in linearer Weise von f ab; $\mu(f) \geqq 0$ für $f \geqq 0$.*

2. *Sind (f_n), (g_n) zwei isotone Folgen von Stufenfunktionen mit demselben punktweisen Supremum, so gilt in $E \cup \{\infty\}$ auch $\bigvee\limits_{1}^{\infty} \mu(f_n) = \bigvee\limits_{1}^{\infty} \mu(g_n)$.*

Beweis. Beide Aussagen sind leichte Konsequenzen der Maßeigenschaften von μ (s. [7], S. 111–112). ∎

Definition. Sei X kompakt und μ: $\Sigma(X) \to E_+$ ein Bairemaß. Eine Σ-meßbare Funktion f: $X \to \overline{\mathbb{R}}_+$ heißt *μ-integrierbar*, wenn eine Folge (f_n) von Σ-Stufenfunktionen existiert mit $0 \leqq f_n \uparrow f$ punktweise und $\bigvee_1^\infty \mu(f_n) \in E$; nach Lemma 3 kann dann $\mu(f) = \bigvee_1^\infty \mu(f_n)$ definiert werden. Entsprechend heißt eine Σ-meßbare Funktion f: $X \to \overline{\mathbb{R}}$ *μ-integrierbar*, wenn ihre nicht-negativen Teile f_+, f_- μ-integrierbar sind; in diesem Falle wird das *Integral* durch

$$\int f\,d\mu \equiv \mu(f) = \mu(f_+) - \mu(f_-) \in E$$

definiert.

Definition. Sei X lokalkompakt und μ: $\Sigma(X) \to E_+ \cup \{\infty\}$ ein Bairemaß. Für kompakte $K \in \Sigma$ sei $\Sigma_K := K \cap \Sigma$, $\mu_K := \mu | \Sigma_K$. Eine Funktion f: $X \to \overline{\mathbb{R}}_+$ heißt *μ-integrierbar*, wenn für jedes kompakte $K \in \Sigma$ die Restriktion $f|K$ μ_K-integrierbar ist und $\bigvee\{\mu_K(f|K)$: $K \in \Sigma$ kompakt$\}$ in E existiert; dieses Supremum wird dann als das Integral $\mu(f)$ bezeichnet. Allgemein heißt f: $X \to \overline{\mathbb{R}}$ *integrierbar*, wenn f_+ und f_- integrierbar sind; in diesem Falle wird das *Integral* wieder durch

$$\int f\,d\mu \equiv \mu(f) = \mu(f_+) - \mu(f_-) \in E$$

definiert.

Bemerkung. Sind f, g: $X \to \overline{\mathbb{R}}$ lokal meßbar und $f = g$ lokal μ-fast überall, d.h. für jedes kompakte $K \in \Sigma$ die Restriktionen $f|K, g|K$ Σ_K-meßbar und $\mu_K(\{t \in K$: $f(t) \neq g(t)\}) = 0$, so impliziert offenbar die Integrierbarkeit von f diejenige von g sowie $\mu(f) = \mu(g)$. Wir betrachten daher im folgenden, mit $f \sim g \Leftrightarrow f = g$ lokal μ-fast überall, den Quotientenraum

$$L^1(\mu) := \{f\colon X \to \mathbb{R} : f\ \mu\text{-integrierbar}\}/\sim .$$

Unsere Definition hat, wenn E nicht sogar Dedekind-vollständig oder $\Sigma(X)$ sogar σ-Algebra ist, den offensichtlichen Mangel, daß eine lokal meßbare Funktion mit integrierbarer Betragsmajorante nicht notwendig selbst integrierbar sein muß. Trotzdem zeigen die folgenden Sätze, daß sie auf eine sinnvolle Klasse von Funktionen führt und für die Darstellungstheorie von σ-vollständigen Vektorverbänden geeignet ist.

Satz 1. *Sei X lokalkompakt, μ: $\Sigma(X) \to E_+ \cup \{\infty\}$ ein Bairemaß.*

1. $L^1(\mu)$ *ist ein linearer Raum und $f \mapsto \mu(f)$ induziert eine positive (d.h. $\mu(f) \geqq 0$ für $f \geqq 0$ lokal μ-fast überall) lineare Abbildung von $L^1(\mu)$ in E.*

2. (Beppo Levi). *Sei (f_n) eine isotone Folge von μ-integrierbaren nicht-negativen Funktionen mit punktweisem $\sup f_n = f$. Gilt $\mu(f_n) \leqq x$ mit einem $x \in E$ für alle $n \in \mathbb{N}$, so ist auch f integrierbar, und es gilt $\mu(f_n) \uparrow \mu(f)$ in E.*

3. (Majorierte Konvergenz). *Sei* (f_n) *eine Folge von* μ*-integrierbaren Funktionen mit punktweisem* $\lim f_n = f$. *Gilt dann mit einer integrierbaren Funktion* $g: X \to \mathbb{R}_+$, *daß* $|f_n| \leqq g$ *für alle* $n \in \mathbb{N}$ *ist, so ist auch* f *integrierbar, und es gilt* $\mu(f_n) \to \mu(f)$ *in* E.

Beweis. Ist $\bar{E}$ die Dedekind-Vervollständigung von E (vgl. [6], S. 109) und $i: E \to \bar{E}$ die zugehörige, Suprema und Infima erhaltende lineare Einbettung, so bezeichne $\bar{\mu} = i \circ \mu$ das zu μ gehörige Bairemaß mit Werten in $\bar{E}_+ \cup \{\infty\}$.

1. Offenbar ist f genau dann μ-integrierbar, wenn es $\bar{\mu}$-integrierbar und $\bar{\mu}(f) \in i(E)$ ist. Da die Aussagen in 1. bezüglich $\bar{\mu}$ klar sind, ergeben sie sich aus dieser Bemerkung unmittelbar für μ.

2. Sei zunächst X kompakt. Dann gibt es für jedes $n \in \mathbb{N}$ eine Folge $g_{nk} \uparrow f_n$ mit $k \to \infty$ von Σ-Stufenfunktionen g_{nk}. Sei $g_n = \sup\limits_{1 \leqq i \leqq n} g_{in}$; dann ist auch g_n Σ-Stufenfunktion und $0 \leqq g_n \leqq f_n$, mit $g_n \uparrow f$ punktweise. Somit existiert (wegen $\mu(f_n) \leqq x$) und gilt in E

$$\mu(f_m) \leqq \mu(f) = \vee \mu(g_n) \leqq \vee \mu(f_n), \qquad \text{jedes } m \in \mathbb{N}$$

und daher für $m \to \infty$ die Behauptung.

Im allgemeinen Falle genügt es wegen der σ-Vollständigkeit von E, d.h. der Folgenabgeschlossenheit von $i(E)$ in $\bar{E}$, die $\bar{\mu}$-Integrierbarkeit von f und $\bar{\mu}(f_n) \uparrow \bar{\mu}(f)$ in $\bar{E}$ zu zeigen. Nach dem soeben Bewiesenen gilt $\bar{\mu}_K(f_n|K) \uparrow \bar{\mu}_K(f|K)$ für jedes kompakte $K \in \Sigma$; insbesondere ist $\bar{\mu}_K(f|K) \leqq x$ und also f $\bar{\mu}$-integrierbar, und

$$\bar{\mu}(f) = \vee \{\bar{\mu}_K(f|K): K \in \Sigma \text{ kompakt}\} = \vee \bar{\mu}(f_n).$$

3. Wiederum braucht nur die $\bar{\mu}$-Integrierbarkeit von f und $\bar{\mu}(f_n) \to \bar{\mu}(f)$ in $\bar{E}$ gezeigt zu werden. Dabei ergibt sich wegen Vollständigkeit des Verbandes $\bar{E}$ die $\bar{\mu}$-Integrierbarkeit von f direkt aus $|f| \leqq g$ und der $\bar{\mu}$-Integrierbarkeit von g. Seien

$$g_n = \inf_{k \geqq n} (g - f_k), \qquad h_n = \sup_{k \geqq n} (g - f_k) \quad \text{für } n \in \mathbb{N};$$

dann gilt

$$0 \leqq g_n \leqq g - f_n \leqq h_n \leqq 2g \quad \text{und} \quad g_n \uparrow g - f, \quad h_n \downarrow g - f.$$

Die g_n, h_n sind also $\bar{\mu}$-integrierbar mit $\bar{\mu}(g_n) \uparrow \bar{\mu}(g - f) = \bar{\mu}(g) - \bar{\mu}(f)$ und $\bar{\mu}(2g - h_n) \uparrow \bar{\mu}(g + f)$, d.h. $\bar{\mu}(h_n) \downarrow \bar{\mu}(g) - \bar{\mu}(f)$, und somit $\bar{\mu}(f_n) \to \bar{\mu}(f)$. ∎

2. Darstellung σ-vollständiger Vektorverbände als Räume von stetigen numerischen Funktionen

Die Entwicklungen dieses Abschnitts schließen eng an die von Schaefer [5] gegebene Darstellung von Banachverbänden mit „topologischem Orthogonalsystem" an.

E bezeichnet wieder einen σ-vollständigen reellen Vektorverband. Für einen lokalkompakten Hausdorffraum X sei $\mathscr{C}(X)$ die Gesamtheit aller stetigen $\mathbb{R}$-wertigen, $\overline{\mathscr{C}}(X)$ die aller stetigen $\overline{\mathbb{R}}$-wertigen Funktionen auf X, und $\mathscr{K}(X)$ die Menge der Funktionen aus $\mathscr{C}(X)$ mit kompaktem Träger.

Für ein $u>0$ in E sei wieder

$$K_u := \{\varphi\colon E \to \overline{\mathbb{R}}_+ : \ \varphi \text{ Bewertung mit } \varphi(u)=1\}$$

der im Abschnitt 1 beschriebene kompakte Hausdorffraum, Teilraum von $(\overline{\mathbb{R}}_+)^E$.

Ist $U \subset E_+$ ein maximales System von paarweise disjunkten Elementen, so wird

$$K_U := \{\varphi\colon E \to \overline{\mathbb{R}}_+ : \ \varphi \text{ Bewertung mit } \varphi(u)=1 \text{ für ein } u \in U\}$$

als Teilraum von $(\overline{\mathbb{R}}_+)^E$ offenbar (wegen $\min\{\varphi(u), \varphi(v)\} = \varphi(u \wedge v) = 0$ für $u \neq v$ in U) die topologische direkte Summe der kompakten K_u, $u \in U$. Insbesondere ist also K_U ein lokalkompakter Hausdorffraum, und $\mathscr{K}(K_U) = \bigoplus_{u \in U} \mathscr{C}(K_u)$, wenn jeweils $f \in \mathscr{C}(K_u)$ mit $\tilde{f} \in \mathscr{K}(K_U)$: $\tilde{f}(t) = f(t)$ für $t \in K_u$ und $=0$ für $t \in K_U \setminus K_u$ identifiziert wird.

Für $x \in E$ ist (wegen $\min\{\varphi(x_+), \varphi(x_-)\} = \varphi(x_+ \wedge x_-) = 0$ für jede Bewertung φ) die Funktion $\hat{x}\colon K_U \to \overline{\mathbb{R}}$ durch $\hat{x}(\varphi) := \varphi(x_+) - \varphi(x_-)$ wohl definiert.

Satz 2. *Sei U ein maximales System von paarweise disjunkten Elementen $u>0$. Dann gilt*

1. *$\hat{x} \in \overline{\mathscr{C}}(K_U)$ und $\operatorname{int}\{\varphi\colon \hat{x}(\varphi) = \pm\infty\} = \emptyset$ für jedes $x \in E$.*

2. *$x \mapsto \hat{x}$ ist ein Vektorverbandsisomorphismus von E auf $\hat{E} := \{\hat{x} : x \in E\}$, wenn für $\hat{E} \subset \overline{\mathscr{C}}(K_U)$ die Vektorverbandsoperationen punktweise auf den gemeinsamen Endlichkeitsmengen der betreffenden Funktionen definiert werden. Dabei geht noch (das dichte Ideal) $E_U := \bigoplus_{u \in U} E_u$ auf $\mathscr{K}(K_U)$ über.*

Beweis. Daß $\hat{x} \in \overline{\mathscr{C}}(K_U)$ ist und $(\alpha x + \beta y)\hat{}(\varphi) = \alpha \hat{x}(\varphi) + \beta \hat{y}(\varphi)$ sowie $(x \vee y)\hat{}(\varphi) = \max\{\hat{x}(\varphi), \hat{y}(\varphi)\}$ gilt für $\varphi \in \{\varphi \in K_U : \hat{x}(\varphi) \text{ und } \hat{y}(\varphi) \in \mathbb{R}\}$, folgt unmittelbar aus den Definitionen. Damit ist $x \mapsto \hat{x}$ in der angegebenen Weise ein Verbandshomomorphismus von E auf $\hat{E}$, dessen Einschränkung auf ein E_u offenbar dem Kakutani-Isomorphismus von E_u auf $\mathscr{C}(K_u)$ entspricht; insbesondere geht also auch $E_U = \bigoplus_{u \in U} E_u$ injektiv auf $\mathscr{K}(K_U) = \bigoplus_{u \in U} \mathscr{C}(K_u)$ über. Wäre $x \mapsto \hat{x}$ nicht injektiv, so gäbe es Elemente $x \neq y$ in E_+ mit $\hat{x} = \hat{y}$, also auch $\hat{x} \wedge \hat{z} = \hat{y} \wedge \hat{z}$ und somit $x \wedge z = y \wedge z$ für jedes $0 \leqq z \in E_U$ wegen der schon bewiesenen Injektivität auf dem Ideal E_U; dies ergäbe aber nach Lemma 2. doch wieder $x = y$. Wäre schließlich für ein $x \in E$ und ein $0 < y \in E_U$

$$\text{Träger}\,(\hat{y}) \subset \operatorname{int}\{\varphi \in K_U : \ \hat{x}(\varphi) = \pm\infty\}$$

so hätte man

$$\varphi(|x| \wedge n\,y) = \min\{\varphi(|x|), \varphi(n\,y)\} = \varphi(n\,y); \qquad \text{alle } \varphi \in K_U,\ n \in \mathbb{N}.$$

Daraus folgt $|x| \geqq n\,y$ für alle $n \in \mathbb{N}$ (im Widerspruch zur Archimedizität von E), da andernfalls $|x| \wedge n\,y < n\,y$ für ein $n \in \mathbb{N}$ und also auch $\min\{\varphi(|x|), \varphi(n\,y)\} < \varphi(n\,y)$ für ein $\varphi \in K_U$ wäre. Also ist $\operatorname{int}\{\varphi\colon \hat{x}(\varphi) = \pm\infty\} = \varnothing$ für alle $x \in E$. ∎

Im folgenden betrachten wir zu einer Teilmenge $M \subset E$ das *disjunkte Komplement*

$$M^d := \{x \in E\colon |x| \wedge |y| = 0 \text{ für alle } y \in M\}.$$

Dann hat man für $u > 0$ in E, mit $u^d := \{u\}^d$, $u^{dd} := (u^d)^d$, die „Bandzerlegung" $E = u^d \oplus u^{dd}$, deren zugehörige Projektionen $P_u\colon E \to E$ auf das Band u^{dd} und $1 - P_u$ auf u^d Verbandshomomorphismen sind (s. [4]); insbesondere ist $0 \leqq P_u x \leqq x$ für $x \in E_+$.

Es gilt der folgende Eindeutigkeitssatz:

Satz 3. *Seien U und V zwei maximale Systeme von paarweise disjunkten positiven Elementen. Dann ist auch*

$$W := \{u \wedge v\colon\ u \in U, v \in V \text{ mit } u \wedge v > 0\}$$

ein maximales System, und K_W ist homeomorph einem offenen dichten Teil von K_U (oder K_V).

Beweis. Die Maximalität von W ist offensichtlich. Zu $u \in U$, $v \in V$ mit $u \wedge v > 0$ betrachten wir den abgeschlossenen Teil

$$K_{u,v} := \{\varphi \in K_u\colon\ \varphi(x) = 0 \text{ für alle } x \in v^d\}$$

von K_u. Hierfür gilt:

i) $K_u \setminus K_{u,v} = \{\varphi \in K_u\colon \varphi(y) = 0$ für alle $y \in v^{dd}\}$ (und somit $K_{u,v}$ offen und abgeschlossen in K_U), denn: Ist für $\varphi \in K_u$ jedes $\varphi(y) = 0$ für $y \in v^{dd}$, so kann φ nicht in $K_{u,v}$ sein, da sonst eine Zerlegung $u = u_1 + u_2$ mit $u_1 \in v^{dd}$, $u_2 \in v^d$ den Widerspruch $1 = \varphi(u) = \varphi(u_1) + \varphi(u_2) = 0$ ergäbe. Ist umgekehrt $\varphi \in K_u \setminus K_{u,v}$, so gibt es also $x \in v^d$ mit $\varphi(x) \neq 0$; für $y \in v^{dd}$ gilt daher

$$0 = \varphi(|x| \wedge |y|) = \min\{\varphi(x), \varphi(y)\} = \varphi(y).$$

ii) $u^{dd} \cap v^d = \{x \in u^{dd}\colon \varphi(x) = 0$ für jedes $\varphi \in K_{u,v}\}$, denn: Nach Definition von $K_{u,v}$ ist $\varphi(x) = 0$ für jedes $x \in u^{dd} \cap v^d$ und $\varphi \in K_{u,v}$. Sei umgekehrt $x \in u^{dd}$ mit $\varphi(x) = 0$ für jedes $\varphi \in K_{u,v}$, und $y \in v^{dd}$ beliebig; um $|x| \wedge |y| = 0$ zu beweisen, genügt es auf Grund der Injektivitätsaussage von Satz 2, $\varphi(|x| \wedge |y|) = 0$ für alle $\varphi \in K_u$ zu zeigen. Dies stimmt für $\varphi \in K_{u,v}$ wegen $\varphi(x) = 0$, und für $\varphi \in K_u \setminus K_{u,v}$ ist $\varphi(y) = 0$ nach i).

iii) $\bigcup_{v\in V} K_{u,v}$ ist dicht in K_u (und also $\bigcup_{\substack{u\in U\\ v\in V}} K_{u,v}$ offen und dicht in K_U)
denn: sonst existiert nach Urysohn ein $0 \neq x \in E_u \subset u^{dd}$ mit Träger $(\hat{x}) \subset K_u \setminus \bigcup_{v\in V} K_{u,v}$, also mit $|\hat{x}(\varphi)| = \varphi(x) = 0$ für alle $\varphi \in \bigcup_{v\in V} K_{u,v}$. Nach ii) wäre $x \in \bigcap_{v\in V} v^d$, also $x=0$ wegen Maximalität von V.

iv) $K_{u,v}$ ist homeomorph $K_{P_v u} \equiv \{\varphi\colon E \to \overline{\mathbb{R}}_+ \colon \varphi$ Bewertung mit $\varphi(P_v u) = 1\}$.

Eine Homeomorphie von $K_{P_v u}$ auf $K_{u,v}$ erhält man nämlich durch die Abbildung $\varphi \mapsto \varphi \circ P_v$: diese Abbildung ist offensichtlich stetig und surjektiv; ist $\varphi \neq \psi$ in $K_{P_v u}$, so existiert auf Grund der Eindeutigkeitsaussage von Lemma 1 ein $x \in E_{P_v u} \subset v^{dd}$ mit $\varphi(x) = \varphi(P_v x) \neq \psi(x) = \psi(P_v x)$.

v) $K_{P_v u}$ ist homeomorph $K_{v\wedge u}$; denn: Zunächst ist $(P_v u)^{dd} = (v \wedge u)^{dd}$, da $v \wedge u \leqq \bigvee_{n\in\mathbb{N}} (n v \wedge u) = P_v u$ und somit $(v\wedge u)^{dd} \subset (P_v u)^{dd} \subset v^{dd} \cap u^{dd}$; andererseits gilt für $0 \leqq x \in v^{dd} \cap u^{dd}$

$$\bigvee_{n\in\mathbb{N}} (n v \wedge x) = x = \bigvee_{n\in\mathbb{N}} (n u \wedge x) = \bigvee_{n\in\mathbb{N}} (n v \wedge n u \wedge x) = \bigvee_{n\in\mathbb{N}} \big(n(v\wedge u) \wedge x\big),$$

also auch $x \in (v \wedge u)^{dd}$. Es bleibt also zu zeigen, daß für beliebige $w, w' \in E$ mit $0 < w \leqq w'$ und $w^{dd} = (w')^{dd}$ notwendig K_w homeomorph $K_{w'}$ ist. Identifizieren wir nach Satz 2 (mit $(w')^{dd}$ an Stelle von E) schon $E_{w'}$ mit $\mathscr{C}(K_{w'})$, so wird w' die konstante Funktion 1 und w eine nichtnegative Funktion, deren Träger wegen $w^{dd} = (w')^{dd}$ offenbar ganz $K_{w'}$ sein muß. Die Multiplikationsabbildung $f \mapsto wf$ von $E_{w'}$ in E_w ist also injektiv, außerdem natürlich linear und in beiden Richtungen positiv, und w' geht auf w über. Ist ihre Surjektivität gezeigt, so sind $\mathscr{C}(K_{w'})$ und $\mathscr{C}(K_w)$ isometrisch isomorph und also $K_{w'}$ und K_w homeomorph. Um eine Funktion $g \in E_w$ als $g = wf$ mit $f \in \mathscr{C}(K_{w'})$ darzustellen kann o.B.d.A. $0 \leqq g \leqq w$ angenommen werden. Es gilt punktweise mit $n \to \infty$

$$w_n := \frac{1}{n}(n w \vee w') = w \vee \frac{w'}{n} \downarrow w$$

und also $w \dfrac{g}{w_n} \uparrow g$; andererseits existiert wegen σ-Vollständigkeit von $\mathscr{C}(K_{w'})$ das Supremum $f := \bigvee_{n\in\mathbb{N}} \dfrac{g}{w_n}$ in $\mathscr{C}(K_{w'})$. Dann ist $f(t) \geqq \sup \dfrac{g(t)}{w_n(t)} := f'(t)$, und $M := \{t \in K_{w'} \colon f(t) > f'(t)\} = \bigcup_{k\in\mathbb{N}} M_k$ mit

$$M_k = \left\{t\colon f'(t) \leqq f(t) - \frac{1}{k}\right\} = \bigcap_{n=1}^{\infty} \left\{t\colon \frac{g(t)}{w_n(t)} < f(t) - \frac{1}{k} + \frac{1}{n}\right\}.$$

Da $K_{w'}$ Quasi-Stone-Raum und M_k abgeschlossene G_δ-Menge ist, ist die charakteristische Funktion χ von $\operatorname{int} M_k$ stetig, und es gilt $\frac{g}{w_n} \leqq f - \frac{1}{k}\chi$ für alle n, so daß $\chi = 0$ sein muß. Also ist $\operatorname{int} M_k = \varnothing$ für alle k und somit nach dem Baireschen Satz auch $\operatorname{int} M = \varnothing$. Da $\{t: (wf)(t) > g(t)\} \subset M$ und offen ist, gilt $wf = g$. ∎

Korollar. *Sei $u > 0$ in E. Dann existiert ein bis auf Homeomorphie eindeutig bestimmter kompakter Hausdorffraum K mit einem Verbandsisomorphismus von u^{dd} auf einen (bezüglich punktweiser Operationen auf Endlichkeitsmengen) Vektorverband $F \subset \bar{\mathscr{C}}(K)$, dessen Elemente jeweils außerhalb einer nirgends dichten Menge endliche Werte haben, und der $\mathscr{C}(K)$ enthält.*

Beweis. Nach Satz 2 für u^{dd} statt E erfüllt $K = K_u$ sämtliche Behauptungen des Satzes. Sei umgekehrt K ein beliebiger kompakter Raum der im Satz beschriebenen Art und T der Verbandsisomorphismus von u^{dd} auf F. Wegen $1^{dd} = F$ gilt für das Urbild $v = T^{-1}1$ der konstanten Funktion 1 auf K also $v^{dd} = u^{dd}$, und T bildet E_v auf $\mathscr{C}(K)$ ab. Da andererseits $x \mapsto \hat{x}$ eine Ordnungsisomorphie mit v auf 1 von E_v auf $\mathscr{C}(K_v)$ definiert, sind K und K_v homeomorph. Nach Satz 3 (für u^{dd} statt E) sind weiter K_v und K_u homeomorph. Also ist K homeomorph K_u. ∎

Bemerkung. Daß für nicht-endliche maximale disjunkte Systeme U oder V die nach Satz 3 bestehende Homeomorphie zwischen offenen dichten Teilen von K_U und K_V im allgemeinen nicht die Homeomorphie von ganz K_U und K_V zur Folge hat, zeigt das Beispiel $E = l_1$, wo für $U = \{(1,0,0,\ldots), (0,1,0,\ldots), (0,0,1,0,\ldots),\ldots\}$ offenbar $K_U = \mathbb{N}$ wird, während man für $V = \{v\}$ ein-elementig mit einem beliebigen strikt positiven $v \in l_1$ leicht sieht, daß K_V mit der Stone-Čech-Kompaktifizierung von $\mathbb{N}$ identifiziert werden kann.

3. Spektraltheorie

Die im 2. Abschnitt beschriebene Darstellungstheorie steht in weitgehender Analogie zur Gelfandtheorie kommutativer C^*-Algebren. Eine Integraldarstellung der inversen Gelfandabbildung liefert dort ein Spektralmaß P auf dem Strukturraum K der Algebra, dessen Bildmaße $P \circ f^{-1}$ auf $\mathbb{C}$ für stetige $f: K \to \mathbb{C}$ gerade die Gesamtheit der Spektralmaße der einzelnen Elemente der Algebra liefern. In ähnlicher Weise läßt sich im Falle eines σ-vollständigen Vektorverbandes E die Freudenthalsche Spektraldarstellung der einzelnen Elemente von E (vgl. [2] und [6] Seite 105) simultan aus einer Integraldarstellung der in Satz 2 beschriebenen Abbildung $\hat{x} \mapsto x$ gewinnen (und in eine flexible Integrationstheorie einordnen, vgl. [8] Abschnitt 2).

Wir fixieren wieder ein maximales System U von paarweise disjunkten Elementen $u>0$ und betrachten den zugehörigen lokalkompakten Hausdorffraum K_U, die topologische direkte Summe der kompakten K_u, $u\in U$. Nach der in [8], Theorem 2 Seite 199 bewiesenen Verallgemeinerung des Rieszschen Darstellungssatzes gibt es ein eindeutig bestimmtes Bairemaß $\mu\colon \Sigma(K_U)\to E_+\cup\{\infty\}$ mit

$$x=\int \hat{x}\,d\mu\equiv\mu(\hat{x}) \qquad \text{für jedes } x\in E_U.$$

Satz 4. 1. *μ ist ein Spektralmaß (d.h. $\mu(A\cap B)=\mu(A)\wedge\mu(B)$ für beliebige $A,B\in\Sigma(K_U)$ mit $\mu(A)\neq\infty\neq\mu(B)$). Träger $(\mu)=K_U$, und für eine kompakte G_δ-Menge A gilt*

$$\mu(A)=0 \Leftrightarrow \operatorname{int} A=\varnothing.$$

2. *Für $x\in E$ ist $\hat{x}$ μ-integrierbar und $\mu(\hat{x})=x$. Umgekehrt gilt für eine beliebige μ-integrierbare Funktion $f\colon K_U\to\overline{\mathbb{R}}$, daß $(\mu(f))\hat{}=f$ lokal μ-fast überall ist.*

Beweis. i) Daß Träger $(\mu)=K_U$ ist, folgt sofort aus dem Urysohnschen Lemma und der Injektivität der durch μ dargestellten Abbildung $\hat{x}\mapsto x$; insbesondere ist für eine μ-Nullmenge A notwendig $\operatorname{int} A=\varnothing$. Ist umgekehrt A kompakte G_δ-Menge mit $\operatorname{int} A=\varnothing$ und (x_n) eine fallende Folge in E_+, für die $\hat{x}_n\in\mathscr{K}(K_U)$ und $\lim \hat{x}_n=\chi_A$ ist, so gilt (majorierte Konvergenz) $\mu(A)=\bigwedge_1^\infty x_n$ und also $(\mu(A))\hat{}\leqq\chi_A$; da $(\mu(A))\hat{}$ stetig ist, muß $\mu(A)=0$ sein.

ii) Ist $x\in E_+$ und $u\in U$, so ist nach Satz 2 $(x\wedge nu)\hat{}$ das punktweise $\inf\{\hat{x}, n\chi_{K_u}\}$ für jedes $n\in\mathbb{N}$ und also nach Levi

$$P_u x=\bigvee_{n\in\mathbb{N}} x\wedge nu=\mu(\hat{x}\,\chi_{K_u}).$$

Summation über endlich viele $u\in U$ und anschließende Supremumsbildung über alle endlichen Teilmengen von U ergibt (links nach Lemma 2.2, rechts nach Definition des Integrals) die Integrierbarkeit von $\hat{x}$ und $x=\mu(\hat{x})$. Gleiches gilt damit auch für beliebige $x\in E$ durch Zerlegung in nicht-negative Teile.

iii) Gilt $x_n\uparrow x$ in E_+, so konvergiert $\hat{x}_n\uparrow\hat{x}$ punktweise lokal μ-fast überall; insbesondere gilt bei beliebigem $u\in U$ für die Bandprojektion P_u auf u^{dd}

$$(P_u x)\hat{}=\chi_{K_u}\hat{x} \qquad \mu\text{-fast überall;} \quad \text{jedes } x\in E.$$

Denn: Die letzte Beziehung folgt für $x\geqq 0$ schon aus der Konvergenzaussage für Folgen in einem Band u^{dd} durch Anwendung auf $x\wedge nu\uparrow P_u x$, wenn man noch $(x\wedge nu)\hat{}=\inf\{\hat{x}, n\chi_{K_u}\}$ beachtet. Sei nun zunächst $x_n\uparrow x$ in $(u_{dd})_+$ gegeben. Dann ist natürlich, mit $f=\sup\hat{x}_n$ (punktweise),

$\hat{x} \geqq f$ mit trivialer Gleichheit außerhalb K_u. Weiter ist

$$\{t \in K_u\colon\ \hat{x}(t) > f(t)\} = \bigcup_{n \in \mathbb{N}} \left\{t \in K_u\colon\ \hat{x}(t) \geqq f(t) + \frac{1}{n}\right\}$$

und für festes $n \in \mathbb{N}$

$$\left\{t \in K_u\colon\ \hat{x}(t) \geqq f(t) + \frac{1}{n}\right\} = \bigcap_{i \in \mathbb{N}} \left\{t \in K_u\colon\ \hat{x}(t) > \hat{x}_i(t) + \frac{1}{n} - \frac{1}{i}\right\} := A_n.$$

Da K_u Quasi-Stone-Raum und A_n hiernach abgeschlossene G_δ-Menge in K_u ist, ist $\operatorname{int} A_n$ abgeschlossen und daher $\chi_{\operatorname{int} A_n} = \hat{y}$ für ein $y \in E_+$.

Somit wird für $i \geqq 2n$

$$\hat{x}(t) \geqq \hat{x}_i(t) + \left(\frac{1}{n} - \frac{1}{i}\right) \hat{y}(t) \geqq \hat{x}_i(t) + \frac{1}{2n} \hat{y}(t), \qquad \text{alle } t \in K_U.$$

Dies bedeutet, daß

$$x = \bigvee_{k \in \mathbb{N}} x_k \geqq x_i + \frac{1}{2n} y, \qquad \text{alle } i \geqq 2n, \quad n \in \mathbb{N},$$

somit $y = 0$ und also $\operatorname{int} A_n = \varnothing$ sein muß.

Es folgt $\mu(A_n) = 0$ $\left(\text{und also auch } \mu\left(\bigcup_1^\infty A_n\right) = 0\right)$ nach i). Schließlich ergibt sich die allgemeine Konvergenzaussage durch Kombination der beiden Teilresultate: Es gilt $x_n \wedge n u \uparrow P_u x$ für jedes $u \in U$ und also punktweise μ-fast überall

$$(x_n \wedge n u)\hat{}\uparrow (P_u x)\hat{} = \chi_{K_u} \hat{x};$$

andererseits hat man μ-fast überall für $f = \sup \hat{x}_n$

$$(x_n \wedge n u)\hat{} = \inf\{\hat{x}_n, n \chi_{K_u}\} \uparrow \chi_{K_u} f \quad \text{bei} \quad n \to \infty.$$

Da jeder kompakte Teil von K_U von endlich vielen K_u, $u \in U$, überdeckt wird, folgt die Behauptung.

iv) Sei K kompakte G_δ-Menge; dann gilt $\chi_A = (\mu(A))\hat{}$ lokal μ-fast überall für jede Bairemenge $A \subset K$, denn: Nach iii) enthält das System

$$\mathscr{D} := \{A \in K \cap \Sigma\colon\ \chi_A = (\mu(A))\hat{} \text{ lokal } \mu\text{-fast überall}\}$$

kompakte G_δ-Mengen und ist außerdem abgeschlossen gegen isotone Folgenlimites. Da offenbar für $A, B \in \mathscr{D}$ im disjunkten Falle auch $A \cup B$, im Falle $A \subset B$ auch $B \setminus A$ in $\mathscr{D}$ ist, ist $\mathscr{D}$ Dynkin-System und also $\mathscr{D} = K \cap \Sigma$ (vgl. [1] S. 18).

v) μ ist Spektralmaß: Offenbar genügt es, für disjunkte A, $B \in \Sigma$ mit endlichem $\mu(A)$, $\mu(B)$ noch $\mu(A) \wedge \mu(B) = 0$ zu zeigen. Dies gilt zunächst, wenn A, B in einer kompakten G_δ-Menge K enthalten sind, da für $x=\mu(A)\wedge\mu(B)$ nach iv) $\hat{x}\leqq\chi_{A\cap B}$ lokal μ-fast überall und daher $x\leqq\mu(A\cap B)=0$ gilt. Im allgemeinen gibt es jedenfalls eine $A\cup B$ überdeckende aufsteigende Folge kompakter G_δ-Mengen K_n, so daß gilt

$$0=\mu(A\cap B\cap K_n)\uparrow\mu(A\cap B).$$

vi) Für eine integrierbare Funktion $f\colon K_U\to\overline{\mathbb{R}}_+$ gilt

$$\mu(\chi_{K_u}f)=P_u\mu(f), \qquad \text{jedes } u\in U.$$

Denn auf Grund der Definition des Integrals einerseits und der Tatsache andererseits, daß für ein beliebiges isotones Netz (x_α) in E_+ mit $\vee x_\alpha=x$ auch $\vee P_u x_\alpha=P_u x$ ist, kann $f=0$ außerhalb einer kompakten G_δ-Menge $K\supset K_u$ und zugleich auf K als Σ-Stufenfunktion angenommen werden. Die Behauptung ergibt sich dann aus der Beziehung

$$\mu(\inf\{f, n\,\chi_{K_u}\})=\mu(f)\wedge n\,u$$

mit $n\to\infty$. Diese Beziehung wiederum oder allgemeiner $\mu(\inf\{g,h\}=\mu(g)\wedge\mu(h)$ für Stufenfunktionen $g=\Sigma\,\alpha_i\chi_{A_i}$ $h=\Sigma\,\beta_k\chi_{B_k}$: $K\to\mathbb{R}_+$ (mit endlich vielen paarweise disjunkten $A_1,\dots,A_r$ bzw. $B_1,\dots,B_s$) folgt aus v), da ohne Einschränkung $\inf\{g,h\}=0$ angenommen werden kann, so daß $\mu(g)\wedge\mu(h)=\sum_{i,k}(\alpha_i\,\mu(A_i))\wedge(\beta_k\,\mu(B_k))=0$ wird.

vii) Für eine integrierbare Funktion $f\colon K_U\to\overline{\mathbb{R}}$ gilt $f=(\mu(f))^\wedge$ lokal μ-fast überall, denn: O.B.d.A. kann $f\geqq0$ angenommen werden. Bei beliebigem $u\in U$ ist

$$\chi_{K_u}f=\chi_{K_u}(\mu(f))^\wedge \qquad \mu\text{-fast überall}$$

zu zeigen. Nach iii) und vi) ist $\chi_{K_u}(\mu(f))^\wedge=(\mu(\chi_{K_u}f))^\wedge$ μ-fast überall. Nach iv) gilt weiter $\chi_{K_u}f=(\mu(\chi_{K_u}f))^\wedge$ lokal μ-fast überall für eine Stufenfunktion f und damit nach iii) allgemein, da $\chi_{K_u}f$ isotoner Grenzwert von Stufenfunktionen ist. ∎

Korollar. *Zu jedem σ-vollständigen Vektorverband E gibt es einen lokalkompakten Hausdorffraum X, topologische direkte Summe von Kompakten, und ein Bairemaß μ: $\Sigma(X)\to E_+\cup\{\infty\}$ so, daß das μ-Integral $f\mapsto\mu(f)$ einen Isomorphismus des Vektorverbandes $L^1(\mu)$, dessen Operationen punktweise lokal μ-fast überall definiert sind, auf E darstellt.*

Beweis. Nach Satz 4 hat für ein beliebiges maximales disjunktes System U der Raum $X=K_U$ die behaupteten Eigenschaften. Es genügt hierbei, Klassen $\mathbb{R}$-wertiger Funktionen zu betrachten, da für eine integrierbare Funktion $f\colon X\to\overline{\mathbb{R}}$ nach Satz 4 $f=(\mu(f))^\wedge$ lokal μ-fast überall und nach Satz 2 $\{t\in X\colon(\mu(f))^\wedge(t)=\pm\infty\}$ eine abgeschlossene G_δ-Menge mit leerem Inneren und somit nach Satz 4 eine lokale μ-Nullmenge ist. ∎

Der Zusammenhang mit der Freudenthalschen Spektraltheorie ergibt sich in folgendem Satz:

Satz 5. *Sei E ein σ-vollständiger Vektorverband, $x \in E$.*

1. *Zu jedem $u>0$ mit $x \in u^{dd}$ gibt es genau ein Spektralmaß $\pi: \Sigma(\mathbb{R}) \to E_+$ mit*

$$\pi(\mathbb{R})=u \quad \textit{und} \quad x=\int t\, \pi(dt).$$

2. *Für jedes Spektralmaß π: $\Sigma(\mathbb{R}) \to E_+$ mit $x=\int t\, \pi(dt)$ gilt $x \in (\pi(\mathbb{R}))^{dd}$, und die Integrationsabbildung π: $L^1(\pi) \to E$ ist ein Verbandshomomorphismus. (Operationen in $L^1(\pi)$ punktweise (lokal) π-fast überall.) Für die zugehörige (linksseitig stetige) Verteilungsfunktion e: $\mathbb{R} \to E_+$ mit $e(t)=\pi((-\infty, t))$ gilt die Freudenthalsche Formel (mit $u:=\pi(\mathbb{R})$)*

$$e(t)=\bigvee_{n=1}^{\infty}\big(u \wedge n(t\,u-x)_+\big).$$

Beweis. 1. Sei $\hat{x} \in \overline{\mathscr{C}}(K_u)$ die zu u gehörige Darstellung von x und μ: $\Sigma(K_u) \to E_+$ das zugehörige Spektralmaß nach Satz 4 (für u^{dd} statt E); $\pi:=\mu \circ \hat{x}^{-1}$: $\Sigma(\mathbb{R}) \to E$ sei das zugehörige Bildmaß. Da $\hat{x}$ μ-fast überall endlich ist, gilt

$$\pi(\mathbb{R})=\mu(K_u)=u.$$

Bezeichnet $\hat{x}_0$ die Einschränkung von $\hat{x}$ auf seinen Endlichkeitsbereich $\{t \in K_u: \hat{x}(t) \in \mathbb{R}\}$, so existiert und gilt

$$\int f \circ \hat{x}_0\, d\mu=\int f\, d\pi$$

zunächst für $\Sigma(\mathbb{R})$-Stufenfunktionen f: $\mathbb{R} \to \mathbb{R}$ nach Definition von π und daher nach Definition der Integrale auch für jede $\Sigma(\mathbb{R})$-meßbare Funktion f, für welche $f \circ \hat{x}_0$ μ-integrierbar ist. Insbesondere ergibt sich, mit $\mathrm{id}(t)=t$ für $t \in \mathbb{R}$

$$x=\int \hat{x}\, d\mu=\int \hat{x}_0\, d\mu=\int \mathrm{id}\, d\pi.$$

Die Eindeutigkeit von π ergibt sich aus der weiter unten bewiesenen Freudenthalschen Formel: Hiernach enthält, wenn π' ein weiteres Spektralmaß mit $\pi'(\mathbb{R})=u$ und $\pi'(\mathrm{id})=x$ ist, das Dynkinsystem $\{A \in \Sigma(\mathbb{R}): \pi(A)=\pi'(A)\}$ insbesondere halboffene Intervalle, also einen Durchschnitts-abgeschlossenen Erzeuger von $\Sigma(\mathbb{R})$, und ist mithin $=\Sigma(\mathbb{R})$.

2. Die Aussage $x \in \pi(\mathbb{R})^{dd}$ folgt unmittelbar aus der Definition des Integrals $\int \mathrm{id}\, d\pi=x$. Die behauptete Verbandshomomorphie der Integration als Abbildung π: $L^1(\pi) \to E$ für ein Spektralmaß π ergibt sich, wenn für f, g: $\mathbb{R} \to \mathbb{R}_+$ mit $\inf\{f, g\}=0$ π-fast überall noch $\pi(f) \wedge \pi(g)=0$ in E nachgewiesen ist. Dies gilt für charakteristische Funktionen von Mengen aus $\Sigma(\mathbb{R})$ nach Definition des Spektralmaßes, und damit auch

für Stufenfunktionen f, g (s. Punkt vi) des Beweises von Satz 4). Im allgemeinen Falle gibt es aufsteigende Folgen (f_n), (g_n) von Stufenfunktionen mit $0 \leqq f_n \uparrow f$ und $0 \leqq g_n \uparrow g$ und somit

$$\pi(f) \wedge \pi(g) = \bigvee_{n=1}^{\infty} \big(\pi(f_n) \wedge \pi(g_n)\big) = 0.$$

Schließlich ergibt sich die Freudenthalsche Formel für die Verteilungsfunktion e eines Spektralmaßes π mit $\pi(\mathbb{R}) = u$ und $\pi(\mathrm{id}) = x$ aus der soeben bewiesenen Homomorphie und dem Satz von Beppo Levi wegen

$$\chi_{(-\infty,t)} = \sup_{n \in N} \inf\{1, n(t-\mathrm{id})_+\}. \quad \blacksquare$$

Bemerkung. Natürlich braucht in einer Darstellung $x = \int t\,\pi(dt)$ nur über den Träger von π integriert zu werden. Die im Beweis gegebene Darstellung von π in der Form $\pi = \mu \circ \hat{x}^{-1}$ mit Hilfe einer Darstellung von u^{dd} (mit $u = \pi(\mathbb{R})$) gemäß Satz 4 zeigt wegen Träger$(\mu) = K_u$, daß der Träger(π) existiert und gleich $\hat{x}(K_u) \cap \mathbb{R}$ ist. Diese abgeschlossene Teilmenge von $\mathbb{R}$ heißt aus (im Hinblick auf die durch das vorangehende Korollar beschriebene multiplikative Struktur von u^{dd}) offensichtlichen Gründen das *Spektrum* von x relativ u. Aus der gegebenen Darstellung des Spektrums von x als wesentlicher Wertebereich von $\hat{x}$ ergeben sich unmittelbar weitere Eigenschaften des Spektrums, insbesondere ein Spektralabbildungssatz.

Literatur

1. Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. Berlin: Walter de Gruyter 1968.
2. Freudenthal, H.: Teilweise geordnete Moduln. Proc. Acad. Amsterdam **39**, 641–651 (1936).
3. Johnson, D. G., Kist, J. E.: Prime ideals in vector lattices. Canadian J. Math. **14**, 517–528 (1962).
4. Luxemburg, W. A. J., Zaanen, A. C.: Riesz spaces (linear vector lattices) I. Preprint.
5. Schaefer, H. H.: On the representation of Banach lattices by continuous numerical functions. Math. Z. **125**, 215–232 (1972).
6. Vulikh, B. Z.: Introduction to the theory of partially ordered spaces. Groningen: Wolters-Noordhoff 1967.
7. Wright, J. D. Maitland: Stone algebra valued measures and integrals. Proc. London Math. Soc. **19**, 107–122 (1969).
8. Wright, J. D. Maitland: Vector lattice measures on locally compact spaces. Math. Z. **120**, 193–203 (1971).

Dr. W. Hackenbroch
Math. Institut der Universität
D-6600 Saarbrücken
Deutschland

(Eingegangen am 10. April 1972)

Math. Z. 128, 129–151 (1972)

4-dimensionale Translationsebenen

Dieter Betten

Einleitung und Definitionen

Eine „topologische projektive Ebene $\mathbf{P}=(P, \mathfrak{L}, I)$" besteht aus einem Punktraum P, einem Geradenraum $\mathfrak{L}$ und einer Inzidenzrelation $I \subseteq P \times \mathfrak{L}$ derart, daß je zwei verschiedene Punkte mit genau einer Geraden inzidieren, je zwei verschiedene Geraden sich in genau einem Punkt schneiden, und Verbinden und Schneiden stetige Operationen sind. Zwei Ebenen $\mathbf{P}=(P, \mathfrak{L}, I)$ und $\mathbf{P}'=(P', \mathfrak{L}', I')$ heißen isomorph, wenn es Homöomorphismen π: $P \to P'$ und λ: $\mathfrak{L} \to \mathfrak{L}'$ gibt, so daß für alle $p \in P$, $L \in \mathfrak{L}$ gilt: $p\,I\,L$ genau wenn $p^{\pi}\,I'\,L^{\lambda}$. Die Ebene $\mathbf{P}$ heißt 2-dimensional, wenn P und $\mathfrak{L}$ homöomorph zur Fläche der reellen projektiven Ebene sind und die Punktreihen $[L]$ (Menge der Punkte auf der Geraden L) und Geradenbüschel $[p]$ (Menge der Geraden durch den Punkt p) homöomorph zur Kreislinie sind. Die volle Kollineationsgruppe Γ einer 2-dimensionalen Ebene ist eine höchstens 8-dimensionale Lie-Gruppe, und die 2-dimensionalen (oder auch ebenen) projektiven Ebenen mit $\dim \Gamma \geqq 3$ sind vollständig klassifiziert. Insbesondere gibt es vier Scharen nicht desarguesscher Ebenen: die Moulton-Ebenen ($\dim \Gamma = 4$), die Ebenen mit einfacher Kollineationsgruppe $(\Gamma = PSL_2(R))$, die Ebenen für die Γ eine 3-dimensionale Gruppe ist und genau zwei Punkte und zwei Geraden festläßt, und schließlich die Ebenen, deren volle Kollineationsgruppe 3-dimensional ist und genau eine Fixfahne hat. Eine Übersicht über diese Klassifikation gibt [21].

Auf einer 3-dimensionalen Punktmannigfaltigkeit gibt es keine topologischen projektiven Ebenen, und ohne auf axiomatische Feinheiten einzugehen – siehe dazu [5, 23] – definieren wir: Eine „4-dimensionale Ebene" ist eine topologische projektive Ebene, deren Punktraum und Geradenraum homöomorph zur komplexen projektiven Ebene sind und deren Punktreihen $[L]$ und Geradenbüschel $[p]$ 2-Sphären sind. Eine 4-dimensionale Ebene verhält sich also topologisch wie die gewöhnliche projektive Ebene über den komplexen Zahlen, ist aber als topologisch-geometrische Struktur möglicherweise nicht zur komplexen Ebene isomorph. Auf dem Weg zur Klassifikation 4-dimensionaler Ebenen bewies Salzmann unter anderem folgendes [24, 25]: Die volle Kollineationsgruppe Γ einer 4-dimensionalen Ebene ist eine Lie-Gruppe einer Dimension $\leqq 16$. Die Ebene ist desarguessch, falls sie eine zusammenhängende punkttransitive Gruppe von Kollineationen zuläßt, oder wenn $\dim \Gamma \geqq 10$ ist. Jede 4-dimensionale Ebene mit 9-dimensionaler Kollineationsgruppe ist eine Translationsebene, das heißt, es gibt eine Gerade W derart, daß zu je zwei verschiedenen nicht auf W liegenden Punkten p und q eine Kollineation existiert, die jeden Punkt von W und jede Gerade durch $(p \cup q) \cap W$ festhält und p nach q überführt. Die ersten Beispiele nicht desarguesscher 4-dimensionaler Ebenen finden sich in [4, 17], es sind durchweg Translationsebenen.

In der vorliegenden Arbeit wird folgendes bewiesen: Jede 4-dimensionale Translationsebene entsteht, indem man eine geeignete Partition $\mathfrak{B}$ des R^4 in 2-dimensionale Teilräume nimmt, diese Partition starr im R^4 verschiebt, und die so entstehende affine Ebene topologisch projektiv abschließt. Jede solche Partition kann mit Hilfe eines „transversalen Homöomorphismus" τ der reellen Ebene konstruiert werden, und die zugehörige Translationsebene $\mathbf{P}(\tau)$ ist genau dann desarguessch, wenn τ linear ist (Satz 1). Der Kern einer 4-dimensionalen Translationsebene ist isomorph zu C oder zu R je nachdem, ob die Geometrie desarguessch ist oder nicht. Es folgt, daß im nicht desarguesschen Fall die volle Kollineationsgruppe Γ die Translationsachse festhält und daß die Standgruppe Γ_0 auf einem eigentlichen Punkt als lineare Gruppe des R^4 wirkt (Satz 2). Man kann somit die Theorie der Darstellungen linearer Gruppen anwenden. Da Γ_0 ferner auf der zur 2-Sphäre homöomorphen Partition als Transformationsgruppe wirkt, hat man als zweites Hilfsmittel die Theorie der Transformationsgruppen zur Hand, insbesondere den schon bei ebenen Ebenen so nützlichen Satz von Brouwer über die Transformationsgruppen auf eindimensionalen Mannigfaltigkeiten. Wir beweisen, daß aus der Transitivität von Γ auf der Translationsachse der Satz von Desargues folgt (Satz 3). Hieraus ergibt sich, daß 4-dimensionale Translationsebenen mit $\dim \Gamma \geqq 9$ desarguessch sind (Satz 4). Zusammen mit dem oben erwähnten Resultat von Salzmann hat man damit: Jede 4-dimensionale Ebene mit $\dim \Gamma \geqq 9$ ist desarguessch. Die Zahl 9 ist Grenzdimension in dem Sinn, daß schon für $\dim \Gamma = 8$ nicht desarguessche Ebenen existieren. Wir geben nämlich eine Schar nicht desarguesscher Translationsebenen mit 8-dimensionaler Kollineationsgruppe an. Diese Ebenen entstehen, indem man eine Hälfte des desarguesschen Büschels mit Hilfe eines Parameters $w > 1$ abändert. Verschiedene Werte $w \neq w'$ liefern dabei nicht isomorphe Geometrien (Satz 5).

Während bei 2-dimensionalen projektiven Ebenen jede Translationsebene desarguessch ist, gibt es im 4-dimensionalen Fall eine Fülle nicht desarguesscher Translationsebenen (auch für $\dim \Gamma = 7$ ergeben sich nicht desarguessche Translationsebenen, die aber an anderer Stelle behandelt werden sollen). Diese Tatsache verleitet zu der Hoffnung, daß es erst recht eine große Vielfalt von 4-dimensionalen Nicht-Translationsebenen gibt. Allerdings scheint bisher noch kein Beispiel einer solchen Ebene angegeben worden zu sein.

Hilfsmittel und Bezeichnungen

Wir zitieren einige bekannte Sätze, die ständig benutzt werden und geben einige Hilfssätze, die wir vorziehen, um später den Beweisgang nicht zu unterbrechen. Zur Terminologie sei auch auf [21] verwiesen.

Die topologische Gruppe G wirke als Transformationsgruppe auf dem topologischen Raum M. Für $S \subseteq M$ setzen wir $G_S = \{g \in G;\ S^g = S\}$ und $G_{[S]} = \{g \in G;\ s^g = s$ für alle $s \in S\}$. Falls $S = \{b\}$ nur aus einem Punkt besteht, schreiben wir kurz G_b für $G_{\{b\}}$. Für $b \in M$ sei $b^G = \{b^g;\ g \in G\}$ die Bahn von b unter G. Die Abbildung $\Phi_b = (g \mapsto b^g)$: $G \to b^G$ bildet G stetig auf die Bahn b^G ab. Es gilt

Lemma (über homogene Räume). *Wenn G lokalkompakt und von abzählbarer Basis und b^G lokalkompakt ist, dann ist Φ_b offen.*

Dieses Lemma findet man in allgemeinerer Form in [3]. Es folgt, daß Φ_b einen Homöomorphismus vom Restklassenraum $G:G_b$ auf die Bahn b^G induziert.

Dimensionsformel. *Die Lie-Gruppe G wirke als Transformationsgruppe auf dem Raum M. Dann gilt für $b \in M$:*

$$\dim G_b \geqq \dim G - \dim b^G .$$

Beweis. Die Bahn b^G ist eineindeutiges stetiges Bild des lokalkompakten Restklassenraumes $G:G_b$, und daher ist $\dim(G:G_b) \leqq \dim b^G$. Zusammen mit $\dim G - \dim G_b = \dim(G:G_b)$ folgt die Behauptung. In vielen Fällen gilt sogar das Gleichheitszeichen, zum Beispiel bei lokalkompakter Bahn (vgl. voriges Lemma).

Unter einem Viereck verstehen wir ein 4-Tupel von Punkten, von denen keine drei auf einer Geraden liegen. In [24, 4.1] wird bewiesen:

Viereck-Lemma. *Läßt eine Kollineation γ aus der Zusammenhangskomponente Γ^1 die Ecken eines Vierecks fest, so ist $\gamma = 1$.*

Satz von Brouwer. *Jede lokalkompakte, zusammenhängende, transitive und effektive Transformationsgruppe der Kreislinie ist als Transformationsgruppe isomorph zur Rotationsgruppe SO_2 oder zu einer endlichblättrigen Überlagerung der Gruppe $\Omega = PSL_2(R)$. Jede lokalkompakte, zusammenhängende, transitive und effektive Transformationsgruppe der Zahlengeraden ist isomorph zur Translationsgruppe $R = \{x \mapsto x + b;\ b \in R\}$,* *zur linearen Gruppe* $L_2=\{x\mapsto ax+b;\ a>0,\ b\in R\}$ *oder zur einfach zusammenhängenden Überlagerungsgruppe* Ω^∞ *von* Ω.

Für diesen Satz aus [7] wird in [20] ein moderner Beweis gegeben. Es gilt außerdem: Die Standgruppen der Gruppe L_2 auf einem Punkt sind genau die eindimensionalen nicht normalen Untergruppen von L_2. Die Standgruppen von Ω auf einem Punkt sind genau die 2-dimensionalen Untergruppen von Ω.

Hilfssatz 1. *Es sei* $\varphi: R^2\to R^2$ *stetig und injektiv, dann ist* φ *offen.*

Der Beweis ergibt sich als einfache Folgerung aus dem Jordanschen Kurvensatz [8, Chap. XVII, 5.4]. Als Korollare erhalten wir: Ist $\varphi: R^2\to R^2$ stetig und bijektiv, so ist φ topologisch. Ist $\varphi: R^2\to R^2$ stetig und injektiv, dann ist $\varphi(R^2)$ offen in R^2.

Hilfssatz 2. *Es sei T eine offene, konvexe Teilmenge des* R^2, *und T werde von jeder gewöhnlichen Geraden getroffen. Dann gilt* $T=R^2$.

Beweis. Angenommen, es gäbe einen Punkt $p\in R^2-T$. Dann sei $\mathfrak{S}$ die Kreislinie der von p ausgehenden Halbgeraden, und wir setzen $\mathfrak{A}=\{H\in\mathfrak{S};\ H\cap T\neq\varnothing\}$ und $\mathfrak{B}=\{H\in\mathfrak{S};\ H\cap T=\varnothing\}$. Dann ist $\mathfrak{A}$ offen in $\mathfrak{S}$, $\mathfrak{A}$ und $\mathfrak{B}$ liegen diametral zueinander, und wegen $\mathfrak{A}\cup\mathfrak{B}=\mathfrak{S}$ ergibt sich ein Widerspruch zum Zusammenhang von $\mathfrak{S}$.

Hilfssatz 3. *Die zusammenhängende Gruppe G wirke als Transformationsgruppe auf der Kreislinie S. Dann wirkt G transitiv auf S oder hat einen Fixpunkt.*

Beweis. Falls G nicht transitiv auf S ist, hat G auf S eine zu einem offenen Intervall homöomorphe Bahn. Ein Randpunkt dieses Intervalles bleibt dann unter G fest.

Hilfssatz 4. *Die Gruppe* SO_2 *wirke auf der 2-Sphäre und halte drei Punkte fest. Dann ist die Wirkung trivial.*

Beweis. Aus [12, 6.7.1] folgt, daß die Gruppe SO_2 auf der 2-Sphäre folgendermaßen wirkt: SO_2 hält zwei Punkte (etwa den Nordpol und Südpol) fest und wirkt als Gruppe von Rotationen um die Nord-Süd-Achse. Mit einem weiteren Punkt bleibt daher alles fest.

Konstruktion 4-dimensionaler Translationsebenen

Eine Partition $\mathfrak{B}$ des R^4 in 2-dimensionale Teilräume besteht aus einem System 2-dimensionaler Teilräume, so daß jeder von Null verschiedene Punkt auf genau einem Teilraum des Systems liegt. Insbesondere sind dann je zwei verschiedene Teilräume des Systems komplementär. Sei Σ die Gruppe der gewöhnlichen starren Schiebungen des R^4, dann gilt offensichtlich, vgl. auch [2]: $\mathbf{A}(\mathfrak{B})=(R^4,\mathfrak{B}^\Sigma)$ ist eine affine

Ebene mit Σ als Translationsgruppe. Wir nennen die Partition $\mathfrak{B}$ desarguessch (nicht desarguessch), wenn die Ebene $\mathbf{A}(\mathfrak{B})$ desarguessch (nicht desarguessch) ist. Die Partition $\mathfrak{B}$ heiße topologisch, wenn die projektive Abschließung $\mathbf{P}(\mathfrak{B})$ von $\mathbf{A}(\mathfrak{B})$ eine 4-dimensionale Translationsebene ist.

Lemma 1. *Jede 4-dimensionale Translationsebene ist darstellbar als* $\mathbf{P}(\mathfrak{B})$ *für eine geeignete Partition* $\mathfrak{B}$.

Beweis. Sei W die Translationsachse und $E=P-[W]$, dann ist E homöomorph zum R^4. Nach [24, 3.2−7] ist die Translationsgruppe G eine Lie-Gruppe, die eine abzählbare Basis hat und keine echten kompakten Untergruppen besitzt. G wirkt als topologische Transformationsgruppe auf dem R^4, und nach dem Lemma über homogene Räume ist für $x\in R^4$ die bijektive stetige Abbildung $\Phi_x=(g\mapsto x^g)$: $G\to R^4$ offen, und daher ist G homöomorph zum R^4. G ist also eine zusammenhängende und (als Translationsgruppe) abelsche Lie-Gruppe und folglich das Produkt einer Torusgruppe und einer Vektorgruppe. Da G keine nicht trivialen kompakten Untergruppen besitzt (oder aus Homotopiegründen) entfällt der Torusanteil, und G ist eine Vektorgruppe. Da die unterliegende Mannigfaltigkeit der R^4 ist, handelt es sich um die Vektorgruppe R^4. Sei nun $w\in W$, dann ist der topologische Raum von $G_{[w]}$ homöomorph zu R^2, und daher ist $G_{[w]}$ isomorph zur Vektorgruppe R^2. Die Kongruenz von $G=R^4$ [2] besteht also aus einer Partition des R^4 in 2-dimensionale Teilräume.

Als nächstes wollen wir ein Verfahren angeben, mit dem man jede Partition des R^4 in 2-dimensionale Teilräume gewinnen kann. Dazu definieren wir: Die Bijektion τ der reellen (α, β)-Ebene heiße transversal, wenn für je zwei gewöhnliche parallele Geraden H, K gilt $|H^\tau\cap K|=1$. In Koordinaten sei τ gegeben durch $\tau=((\alpha,\beta)\mapsto(f(\alpha,\beta), g(\alpha,\beta)))$. Im $R^4=\{(x,y,u,v);\ x,y,u,v\in R\}$ sei $V_{\tau,\alpha,\beta}$ der durch $u=\alpha x+\beta y$, $v=f(\alpha,\beta)\,x+g(\alpha,\beta)\,y$ gegebene 2-dimensionale Teilraum, ferner sei S der Teilraum, der durch $x=y=0$ definiert wird. Dann gilt:

Lemma 2. *Sei* τ *eine transversale Bijektion der reellen* (α,β)*-Ebene, dann bildet das System* $\mathfrak{B}(\tau)=S\cup\{V_{\tau,\alpha,\beta};\ \alpha,\beta\in R\}$ *eine Partition des* R^4 *in 2-dimensionale Teilräume.*

Beweis. Die von Null verschiedenen Punkte mit $x=y=0$ liegen genau auf S. Jetzt sei ein von Null verschiedener Punkt (x,y,u,v) mit $(x,y)\neq(0,0)$ gegeben. Wir betrachten die parallelen Geraden $H=\{(\alpha,\beta);\ u=\alpha x+\beta y\}$, $K=\{(\alpha,\beta);\ v=\alpha x+\beta y\}$ der (α,β)-Ebene. Da τ transversal ist, existiert genau ein $(\alpha,\beta)\in H$ mit $(f(\alpha,\beta), g(\alpha,\beta))\in K$, der Punkt (x,y,u,v) liegt also genau auf der Ebene $V_{\tau,\alpha,\beta}$ des Systems $\mathfrak{B}(\tau)$.

Bemerkungen. Lemma 2 gilt entsprechend auch für andere Körper. Beispiele für transversale Bijektionen liefern die affinen elliptischen Abbildungen, das sind die affinen Abbildungen, die keine Parallelschar in sich überführen. Die transversale Bijektion $\tau=((\alpha,\beta)\mapsto(-\beta,\alpha))$ liefert die Teilräume $u=\alpha x+\beta y$, $v=-\beta x+\alpha y$, zusammen mit S also die gewöhnliche desarguessche Partition der Geraden mit komplexer Steigung in der komplexen affinen Ebene.

Eine transversale Bijektion τ definiert nach Lemma 2 eine Partition $\mathfrak{B}(\tau)$, und diese Partition liefert durch Verschieben und projektives Abschließen eine Translationsebene $\mathbf{P}(\mathfrak{B}(\tau))$, die wir kurz mit $\mathbf{P}(\tau)$ bezeichnen. Es gilt

Satz 1. (a) *Wenn τ ein transversaler Homöomorphismus der reellen (α,β)-Ebene ist, dann ist $\mathbf{P}(\tau)$ eine 4-dimensionale Translationsebene.*

(b) *Jede 4-dimensionale Translationsebene ist darstellbar als $\mathbf{P}(\tau)$ für einen geeigneten transversalen Homöomorphismus der reellen (α,β)-Ebene.*

(c) *Die Ebene $\mathbf{P}(\tau)$ ist genau dann desarguessch, wenn τ linear ist.*

Bevor wir Satz 1 beweisen, führen wir in der Ebene $\mathbf{P}(\tau)=\mathbf{P}(\mathfrak{B}(\tau))$ Koordinaten ein und berechnen den Ternärkörper: Durch Koordinatentransformation im R^4 können wir annehmen, daß $\tau(0,0)=(0,0)$ ist, daß also der Teilraum $W=\{(x,y,0,0);\ x,y\in R\}=V_{\tau,0,0}$ zur Partition $\mathfrak{B}(\tau)$ gehört. Wir wählen W als waagrechte Achse, S als senkrechte Achse eines Koordinatensystems, die Translationsachse sei die uneigentliche Gerade, $(1,0,1,0)$ der Einheitspunkt. Die Einheitsgerade I, die 0 und $(1,0,1,0)$ verbindet, wird gegeben durch $u=x+\beta_1 y$, $v=g_1 y$. Dabei ist $g_1\neq 0$, sonst wäre $\tau(1,\beta_1)=(0,0)=\tau(0,0)$ im Widerspruch zur Injektivität von τ. Die Ebene der Steigung $\sigma=(\sigma_1,\sigma_2)$, das ist die Verbindung von 0 und $(1,0,\sigma_1,\sigma_2)$, hat die Form $u=\sigma_1 x+\beta_\sigma y$, $v=\sigma_2 x+g_\sigma y$. Die Komponenten von $\sigma\circ\xi$ erhält man, indem man die Parallele zu W durch $(0,0,\xi_1,\xi_2)$ mit I schneidet und den x- und y-Wert des Schnittpunktes in die Gleichung der Ebene der Steigung σ einsetzt: Das Gleichungssystem $\xi_1=x+\beta_1 y$, $\xi_2=g_1 y$ liefert. $x=\xi_1-\dfrac{\beta_1}{g_1}\xi_2$, $y=\dfrac{\xi_2}{g_1}$. Einsetzen ergibt

$$(\sigma_1,\sigma_2)\circ(\xi_1,\xi_2)$$
$$=\left(\sigma_1\xi_1-\frac{\beta_1}{g_1}\sigma_1\xi_2+\frac{1}{g_1}\beta_\sigma\xi_2,\ \sigma_2\xi_1-\frac{\beta_1}{g_1}\sigma_2\xi_2+\frac{1}{g_1}g_\sigma\xi_2\right).$$

Ferner ergibt sich die Addition komponentenweise, und die Ternäroperation ist linear.

Beweis von Satz 1a. Um zu zeigen, daß die Ternärmultiplikation stetig ist, genügt es laut Formel nachzuweisen, daß β_σ und g_σ stetig von σ_1 und σ_2 abhängen: Sei φ die Abbildung der (α, β)-Ebene in die Ebene $S_{1,0} = \{(1, 0, u, v);\ u, v \in R\}$, gegeben durch $\varphi = ((\alpha, \beta) \mapsto V_{\tau, \alpha, \beta} \cap S_{1,0})$. Die Abbildung φ läßt sich in Komponenten ausdrücken als $\varphi = (\varphi_1, \varphi_2)$ mit $\varphi_1 = \alpha$, $\varphi_2 = f(\alpha, \beta)$, ist also stetig, da τ ein Homöomorphismus ist. Da ferner jeder Punkt $(1, 0, \sigma_1, \sigma_2)$ auf genau einem Teilraum der Partition $\mathfrak{B}(\tau)$ liegt, ist φ eine Bijektion, und nach Hilfssatz 1 ist φ^{-1} stetig, das heißt $\alpha = \alpha_\sigma$ und $\beta = \beta_\sigma$ hängen stetig von $\sigma = (\sigma_1, \sigma_2)$ ab. Da τ stetig ist, hängt dann g_σ stetig von α_σ und β_σ, also auch stetig von σ_1 und σ_2 ab. Die Ternärmultiplikation ist folglich stetig, und da die Addition komponentenweise erklärt und die Ternäroperation linear ist, folgt die Stetigkeit der Ternäroperation.

Zur Stetigkeit der inversen Operationen hat man folgendes nachzuweisen:

α) In $(u, v) = (\sigma_1, \sigma_2) \circ (x, y) + (t_1, t_2)$ hängt (t_1, t_2) stetig von (σ_1, σ_2), (x, y) und (u, v) ab.

β) In $(\sigma_1, \sigma_2) \circ (x, y) + (t_1, t_2) = (\xi_1, \xi_2) \circ (x, y) + (r_1, r_2)$ mit $(\sigma_1, \sigma_2) \neq (\xi_1, \xi_2)$ hängt die eindeutig bestimmte Lösung (x, y) stetig von (σ_1, σ_2), (ξ_1, ξ_2), (t_1, t_2) und (r_1, r_2) ab.

γ) In $(r_1, r_2) = (\sigma_1, \sigma_2) \circ (x, y)$, $(s_1, s_2) = (\sigma_1, \sigma_2) \circ (x', y')$ mit $(x, y) \neq (x', y')$ hängt die eindeutig bestimmte Lösung (σ_1, σ_2) stetig von (r_1, r_2), (s_1, s_2), (x, y) und (x', y') ab.

Die Richtigkeit von α) folgt aus der Stetigkeit der Ternärmultiplikation und der komponentenweise Definition der Addition. Das gesuchte (x, y) in β) ist Lösung des Gleichungssystems

$$(\sigma_1 - \xi_1)\, x + \left[-\frac{\beta_1}{g_1}(\sigma_1 - \xi_1) + \frac{1}{g_1}(\beta_\sigma - \beta_\xi) \right] y = r_1 - t_1$$

$$(\sigma_2 - \xi_2)\, x + \left[-\frac{\beta_1}{g_1}(\sigma_2 - \xi_2) + \frac{1}{g_1}(g_\sigma - g_\xi) \right] y = r_2 - t_2 .$$

Hier hängen die Koeffizienten der Matrix und die absoluten Glieder stetig von (σ_1, σ_2), (ξ_1, ξ_2), (r_1, r_2) und (t_1, t_2) ab, also auch die Lösung (x, y).

Im Fall γ) ergibt sich folgendes Gleichungssystem für σ_1, σ_2:

$$r_1 - s_1 = \sigma_1 \left[(x - x') - \frac{\beta_1}{g_1}(y - y') \right] + \frac{1}{g_1} \beta_\sigma (y - y')$$

$$r_2 - s_2 = \sigma_2 \left[(x - x') - \frac{\beta_1}{g_1}(y - y') \right] + \frac{1}{g_1} g_\sigma (y - y').$$

Wegen $(x-x', y-y') \neq (0,0)$ ist auch

$$\left((x-x')-\frac{\beta_1}{g_1}(y-y'), \frac{1}{g_1}(y-y')\right) \neq (0,0)$$

und zum Beweis von γ) genügt es nachzuweisen, daß die Abbildung $\pi = \big((x,y,u,v) \mapsto ((x,y,u,v) \cup 0) \cap S_{1,0}\big): R^4 - S \to S_{1,0}$ stetig ist. Wir setzen $S_{r,s} = \{(x,y,u,v);\ x=r,\ y=s\}$ und zeigen zunächst, daß für $(r,s) \neq (0,0)$ die Beschränkung $\pi/S_{r,s}$: $S_{r,s} \to S_{1,0}$ ein Homöomorphismus ist. Wir betrachten die inverse Abbildung φ: Da β_σ und g_σ stetig von $(1,0,\sigma_1,\sigma_2)$ abhängen und folglich auch $u=\sigma_1 r + \beta_\sigma s$, $v=\sigma_2 r + g_\sigma s$, ist φ stetig, und nach Hilfssatz 1 ist $\pi/S_{r,s}$ ein Homöomorphismus. Jetzt sei (x_0, y_0, u_0, v_0) mit $(x_0, y_0) \neq (0,0)$ gegeben, und es sei $(1, 0, \bar\sigma_1, \bar\sigma_2) = (0 \cup (x_0, y_0, u_0, v_0)) \cap S_{1,0}$. Sei K eine offene Kreisscheibe um $(\bar\sigma_1, \bar\sigma_2)$ in $S_{1,0}$ mit Rand D, sei ferner V eine abgeschlossene Kreisscheibe um (x_0, y_0) mit $(0,0) \notin V$. Sei f: $D \times V \to R$ definiert durch

$$\begin{aligned} f = \big(((\sigma_1,\sigma_2),(x,y)) \mapsto ((\rho_1 x + \beta_\sigma y - \bar\sigma_1 x + \beta_{\bar\sigma} y)^2 \\ + (\sigma_2 x + g_\sigma y - \bar\sigma_2 x + g_{\bar\sigma} y)^{\frac{1}{2}}\big). \end{aligned}$$

Die Funktion f ist stetig und nimmt auf dem kompakten Definitionsbereich $D \times V$ ihr Minimum μ an. Es ist $\mu > 0$, sonst läge der Punkt $(\bar\sigma_1, \bar\sigma_2)$ auf D. Wir setzen

$$\begin{aligned} W = \{&(x,y,u,v); \\ &(x,y) \in \mathrm{Inn}(V) \text{ und } ((u - \bar\sigma_1 x + \beta_{\bar\sigma} y)^2 + (v - \bar\sigma_2 x + g_{\bar\sigma} y)^2)^{\frac{1}{2}} < \mu\}. \end{aligned}$$

Dann ist W eine offene Umgebung von (x_0, y_0, u_0, v_0), und da die Beschränkung $\pi/S_{r,s}$ ein Homöomorphismus ist, gilt $\pi(W \cap S_{r,s}) \subset K$ für jedes $(r,s) \in \mathrm{Inn}(V)$, also auch $\pi(W) \subset K$. Die Abbildung π: $R^4 - S \to S_{1,0}$ ist also stetig.

Damit ist gezeigt, daß der Ternärkörper der Ebene $\mathbf{P}(\tau)$ topologisch ist, und nach einem Satz von Skornjakov [26, 21, 7.16] ist die Translationsebene $\mathbf{P}(\tau)$ topologisch.

Beweis von Satz 1b. Sei eine 4-dimensionale Translationsebene gegeben, dann wird sie nach Lemma 1 von einer Partition $\mathfrak{B}$ des R^4 erzeugt. Wir wählen ein (x,y,u,v)-Koordinatensystem des R^4 so, daß der Teilraum $S = \{(0,0,u,v);\ u, v \in R\}$ zur Partition gehört. Jeder andere Teilraum der Partition hat dann die Form $u = \alpha x + \beta y$, $v = f x + g y$. Da je zwei Teilräume der Partition komplementär sind, gilt für einen zweiten Teilraum $u = \alpha' x + \beta' y$, $v = f' x + g' y$, daß $\det \begin{pmatrix} \alpha - \alpha' & \beta - \beta' \\ f - f' & g - g' \end{pmatrix} \neq 0$ ist, insbesondere liefern verschiedene Paare $(\alpha, \beta) \neq (\alpha', \beta')$ verschiedene Teilräume. Wir setzen $T = \{(\alpha, \beta);\ (\alpha, \beta)$ kommt in der Partition vor$\}$ und beweisen α) T wird von jeder gewöhnlichen Geraden H der (α, β)-Ebene R^2 getroffen.

Zum Beweis sei die Gerade gegeben durch

$$H = \{(\alpha, \beta);\ u_0 = \alpha x_0 + \beta y_0,\ (x_0, y_0) \neq (0,0)\}.$$

Sei $v_0 \in R$ beliebig gewählt, dann geht durch (x_0, y_0, u_0, v_0) genau eine Ebene der Partition, es gibt also Zahlen α, β, f, g mit $u_0 = \alpha x_0 + \beta y_0$, $v_0 = f x_0 + g y_0$, und es folgt $(\alpha, \beta) \in T \cap H$.

β) T ist offen in R^2.

Beweis. Da die Partition eine topologische Translationsebene erzeugt, ist die Projektion der Ebene $S_{1,0}$ von 0 aus auf die Ebene $S_{0,1}$ topologisch. Daraus folgt, daß die injektive Abbildung $\varphi\colon S_{1,0} \to R^2$, die jedem Paar $(\alpha, f) \in S_{1,0}$ das Paar $(\alpha, \beta(\alpha, f)) \in R^2$ zuordnet, stetig ist. Nach Hilfssatz 1 folgt, daß $T = \varphi(S_{1,0})$ offen in R^2 ist.

γ) T ist konvex.

Beweis. Sei H die Gerade $H = \{(\alpha, \beta);\ u_0 = \alpha x_0 + \beta y_0,\ (x_0, y_0) \neq (0,0)\}$. Sei $v \in R$, dann schneidet die Verbindung $0 \cup (x_0, y_0, u_0, v)$ die Ebene $S_{1,0}$ in (α, f) und die Ebene $S_{0,1}$ in (β, g). Da die Translationsebene topologisch ist, hängen α und β stetig von v ab. Wir haben also eine stetige Abbildung $\varphi\colon R \to H$ mit $\varphi(R) = H \cap T$. Als stetiges Bild von R ist $H \cap T$ zusammenhängend, und da H beliebig war, folgt die Konvexität von T.

Aus α), β) und γ) folgt nach Hilfssatz 2, daß $T = R^2$ ist, daß also jedes Paar (α, β) in der Partition vorkommt. Die Abbildung $\varphi = ((\alpha, f) \mapsto (\alpha, \beta))\colon S_{1,0} \to R^2$ ist also ein Homöomorphismus. Entsprechend ist auch $\psi = ((\alpha, f) \to (f, g))\colon S_{1,0} \to R^2$ ein Homöomorphismus, und es folgt $\tau = \varphi^{-1}\psi = ((\alpha, \beta) \mapsto (f, g))$ ist eine topologische Abbildung der reellen Ebene auf sich. Um zu zeigen, daß τ transversal ist, seien H und K parallele Geraden: $H = \{(\alpha, \beta);\ u_0 = \alpha x_0 + \beta y_0\}$, $K = \{(f, g);\ v_0 = f x_0 + g y_0\}$, mit $(x_0, y_0) \neq (0,0)$. Dann geht durch (x_0, y_0, u_0, v_0) genau eine Ebene der Partition, es gibt also genau ein Viertupel reeller Zahlen α, β, f, g, welches das Gleichungssystem $u_0 = \alpha x_0 + \beta y_0$, $v_0 = f x_0 + g y_0$ löst, das heißt, das τ-Bild von H schneidet die zu H parallele Gerade K genau einmal.

Beweis von Satz 1c. Sei zunächst τ linear, wobei wir wieder annehmen können, daß $\tau(0,0) = (0,0)$ ist. τ hat also die Form $f = a\alpha + b\beta$, $g = c\alpha + d\beta$, und wegen der Transversalität von τ gilt $S^2 - 4D < 0$, wobei $D = ad - bc$ und $S = a + d$ gesetzt ist. Aus $S^2 - 4D < 0$ folgt $D \neq 0$, und wegen $S^2 - 4D = (a-d)^2 + 4bc < 0$ gilt $b \neq 0$. Die dem Paar (α, β) entsprechende Ebene hat die Gestalt $u = \alpha x + \beta y$, $v = (a\alpha + b\beta)x + (c\alpha + d\beta)y$. Für den Teilraum $0 \cup (1, 0, \sigma_1, \sigma_2)$ gilt $\sigma_1 = \alpha_\sigma$, $\sigma_2 = a\alpha_\sigma + b\beta_\sigma$, und dies liefert $\alpha_\sigma = \sigma_1$,

$f_\sigma = \sigma_2$, $\beta_\sigma = (\sigma_2 - a\sigma_1)/b$, $g_\sigma = c\alpha_\sigma + d\beta_\sigma = -\frac{D}{b}\sigma_1 + \frac{d}{b}\sigma_2$. Falls $(\sigma_1, \sigma_2) = (1, 0)$ ist, erhalten wir $\beta_1 = -a/b$, $g_1 = -D/b$, und Einsetzen in die Formel für die Ternärmultiplikation liefert

$$(\sigma_1, \sigma_2) \circ (\xi_1, \xi_2) = \left(\sigma_1 \xi_1 - \frac{1}{D}\sigma_2 \xi_2, \sigma_1 \xi_2 + \sigma_2 \xi_1 - \frac{S}{D}\sigma_2 \xi_2\right).$$

Wir wollen jetzt zeigen, daß dieser Ternärkörper isomorph zum Körper C der komplexen Zahlen ist. Für das Element, das der Zahl i entspricht, muß gelten $(\sigma_1, \sigma_2) \circ (\sigma_1, \sigma_2) = \left(\sigma_1^2 - \frac{1}{D}\sigma_2^2, 2\sigma_1\sigma_2 - \frac{S}{D}\sigma_2^2\right) = (-1, 0)$.

Es ist $\sigma_2 \neq 0$, sonst wäre $\sigma_1^2 = -1$, und es folgt $\sigma_1^2 - \frac{1}{D}\sigma_2^2 = -1$, $2\sigma_1 - \frac{S}{D}\sigma_2 = 0$. Es ergibt sich $\sigma_1 = S(4D - S^2)^{-\frac{1}{2}}$, $\sigma_2 = 2D(4D - S^2)^{-\frac{1}{2}}$, und $\varphi = ((x, y) \mapsto (x + S(4D - S^2)^{-\frac{1}{2}} y, 2D(4D - S^2)^{-\frac{1}{2}} y))$: $C \to TK$ liefert einen Isomorphismus vom Körper C der komplexen Zahlen auf den Ternärkörper.

Nun sei umgekehrt die Geometrie $\mathbf{P}(\tau)$ desarguessch. Dann ist die Ternärmultiplikation kommutativ, und es gilt für alle $\xi = (\xi_1, \xi_2)$, $\sigma = (\sigma_1, \sigma_2)$:

$$(-\beta_1 \sigma_1 + \beta_0)\xi_2 = (-\beta_1 \xi_1 + \beta_\xi)\sigma^2, \tag{1}$$

$$g_1 \sigma_2 \xi_1 + g_\sigma \xi_2 = g_1 \xi_2 \sigma_1 + g_\xi \sigma_2. \tag{2}$$

Da $\varphi = ((\xi_1, \xi_2) \mapsto (\xi_1, \beta_\xi))$: $S_{1,0} \to R^2$ eine topologische Abbildung und insbesondere surjektiv ist, gibt es ein $\xi = (\xi_1, \xi_2)$ mit $\xi_2 \neq 0$ und $-\beta_1 \xi_1 + \beta_\xi \neq 0$. Dann folgt aus (1)

$$\sigma_2 = \frac{-\beta_1 \xi_2}{-\beta_1 \xi_1 + \beta_\xi}\sigma_1 + \frac{\xi_2}{-\beta_1 \xi_1 + \beta_\xi}\beta_\sigma,$$

und aus (2) ergibt sich

$$g_\sigma = \frac{\beta_\xi g_1 - \beta_1 g_\xi}{-\beta_1 \xi_1 + \beta_\xi}\sigma_1 + \frac{g_\xi - g_1 \xi_1}{-\beta_1 \xi_1 + \beta_\xi}\beta_\sigma.$$

σ_2 und g_σ hängen also linear von σ_1 und β_σ ab, das heißt, der transversale Homöomorphismus τ ist linear.

Bemerkung. Im letzten Beweisschritt haben wir nur die Kommutativität der Ternärmultiplikation benutzt. Daher gilt: Eine 4-dimensionale Translationsebene ist genau dann desarguessch, wenn sie einen Ternärkörper mit kommutativer Ternärmultiplikation besitzt.

Die Kollineationsgruppe

Die Partition $\mathfrak{B}$ des R^4 in 2-dimensionale Teilräume liefert das Büschel der Geraden durch den Nullpunkt. Dieses Büschel ist ein zur 2-Sphäre homöomorpher Unterraum des Geradenraumes. In diesem Abschnitt beweisen wir einige Tatsachen über Büschel und untersuchen die Kollineationsgruppe der Translationsebene.

Lemma 3. *Die Translationsebene* $\mathbf{P}(\mathfrak{B})$ *ist genau dann nicht desarguessch, wenn es außer den 4-reihigen reellen Diagonalmatrizen keine lineare Abbildung des* R^4 *auf sich gibt, die jeden Teilraum der Partition* $\mathfrak{B}$ *in sich überführt.*

Beweis. Die komplexen Diagonalmatrizen im C^2 lassen jede Gerade mit komplexer Steigerung fest und bilden eine 2-dimensionale Gruppe. Um die andere Richtung zu beweisen, wählen wir ein Koordinatensystem im R^4 so, daß W, S und $E=\{(x, y, u, v); u=x, v=y\}$ zur Partition gehören. Die Partition $\mathfrak{B}$ sei durch die transversale Bijektion $\tau=((\alpha, \beta)\mapsto (f(\alpha, \beta), g(\alpha, \beta)))$ gegeben, und es sei mit $\begin{pmatrix}\alpha & \beta\\ f & g\end{pmatrix}$ der durch $u=\alpha x+\beta y$, $v=fx+gy$ gegebene Teilraum bezeichnet. Jede lineare Abbildung, die W, S und E in sich überführt, hat die Form $\begin{pmatrix}A & 0\\ 0 & A\end{pmatrix}$. Sie hält genau dann den Teilraum $D=\begin{pmatrix}\alpha & \beta\\ f & g\end{pmatrix}$ fest, wenn $AD=DA$ gilt. Es folgt für alle α, β:

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ f(\alpha, \beta) & g(\alpha, \beta)\end{pmatrix}=\begin{pmatrix}\alpha & \beta\\ f(\alpha, \beta) & g(\alpha, \beta)\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

also

$$\begin{pmatrix}a\alpha+bf(\alpha, \beta) & a\beta+bg(\alpha, \beta)\\ c\alpha+df(\alpha, \beta) & c\beta+dg(\alpha, \beta)\end{pmatrix}=\begin{pmatrix}\alpha a+\beta c & \alpha b+\beta d\\ f(\alpha, \beta)a+g(\alpha, \beta)c & f(\alpha, \beta)b+g(\alpha, \beta)d\end{pmatrix}.$$

Wäre $b\neq 0$ und $c\neq 0$, dann folgte aus der Gleichheit an der Stelle 1, 1 und an der Stelle 2, 1, daß τ linear ist, ein Widerspruch. Sei nun zunächst $c=0$, dann folgt aus der Gleichheit an der Stelle 2, 2 und wegen $f(\alpha, \beta)\neq 0$, daß $b=0$ ist, und Vergleich der Stellen 2, 1 liefert $a=d$. Ähnlich folgt für $b=0$ aus der Gleichheit an der Stelle 1, 1, daß $c=0$ ist, und die Gleichheit an der Stelle 1, 2 ergibt $a=d$.

Lemma 4. *Der Kern einer 4-dimensionalen Translationsebene ist entweder der Körper R der reellen Zahlen oder der Körper C der komplexen Zahlen. Genau im zweiten Fall ist die Geometrie desarguessch.*

Beweis. Der Kern [2; 16, 8.2] besteht aus allen Endomorphismen der abelschen Trägergruppe R^4 auf sich, die die Kongruenz respektieren, also jeden 2-dimensionalen Teilraum der Partition in sich überführen. Nach Lemma 3 ergibt sich im nicht desarguesschen Fall der Körper R. Andererseits liest man an der gewöhnlichen Darstellung der 4-dimensionalen desarguesschen Ebene ab, daß ihr Kern der Körper C ist.

Lemma 5. *Die volle Kollineationsgruppe Γ einer nicht desarguesschen Translationsebene hält die Translationsachse W fest.*

Beweis. Nach Lemma 4 ist die Charakteristik einer 4-dimensionalen Translationsebene $\neq 2$. Angenommen, es gäbe eine Kollineation $\gamma \in \Gamma$ mit $W^\gamma \neq W$, dann wäre nach [16, Seite 188 und Seite 207] die Ebene eine Moufang-Ebene und nach [10; 21, 7.23, 25] desarguessch.

Sei Γ_0 die Untergruppe der vollen Kollineationsgruppe Γ, die den eigentlichen Punkt $0 \in R^4$ festhält. Dann gilt

Satz 2. *Für eine nicht desarguessche Translationsebene $P(\mathfrak{B})$ stimmt Γ_0 überein mit der Gruppe der linearen Abbildungen des R^4 auf sich, welche $\mathfrak{B}$ in sich überführen.*

Beweis. Nach [2, 3. Satz 19; 16, 8.3.8] sind die Kollineationen, die den Nullpunkt festlassen, genau die semilinearen Abbildungen der Trägergruppe, aufgefaßt als Vektorraum über dem Kern. Da nach Lemma 4 im nicht desarguesschen Fall der Kern der Körper R ist und keine echten Automorphismen hat, ist jede semilineare Abbildung linear.

Korollar. *Die nicht desarguesschen Partitionen $\mathfrak{B}_1$ und $\mathfrak{B}_2$ des R^4 sind genau dann linear isomorph, wenn die Translationsebenen $\mathbf{P}(\mathfrak{B}_1)$ und $\mathbf{P}(\mathfrak{B}_2)$ isomorph sind.*

Die volle Kollineationsgruppe Γ einer 4-dimensionalen projektiven Ebene ist nach [24] eine höchstens 16-dimensionale Lie-Gruppe. Für nicht desarguessche 4-dimensionale Translationsebenen ist Γ semidirektes Produkt von Γ_0 und der Translationsgruppe R^4 und besteht nur aus stetigen Kollineationen. Ferner gilt $\dim \Gamma = \dim \Gamma_0 + 4$. Wir werden im folgenden nur nicht desarguessche Translationsebenen betrachten und Γ_0 als lineare Gruppe des R^4 auffassen.

Lemma 6. *Es sei G eine mindestens 2-dimensionale Kollineationsgruppe einer 4-dimensionalen Translationsebene, und G halte drei Geraden durch einen eigentlichen Punkt fest. Dann ist die Geometrie desarguessch.*

Beweis. Wir können annehmen, daß G zusammenhängend ist, linear auf dem R^4 wirkt und die 2-dimensionalen Teilräume S, W und $E = \{(x, y, u, v);\ u = x,\ v = y\}$ festhält. Angenommen, G hielte einen eindimensionalen Teilraum von W fest, dann lieferte Festhalten eines Punktes dieses Teilraumes wegen des Viereck-Lemmas einen Widerspruch. Nach Hilfssatz 3 folgt, daß G transitiv auf dem Raum der eindimensionalen Teilräume von W wirkt. Da Γ_0 alle 4-dimensionalen reellen Diagonalmatrizen enthält, folgt, daß $\Gamma_{W,S,E}$ die Gruppe der Drehstreckungen von W enthält. Insbesondere gibt es eine Gruppe SO_2, die S, W, E und nach Hilfssatz 4 alle Geraden durch den Nullpunkt festhält. Nach Lemma 3 ist die Geometrie desarguessch.

Lemma 7. *Die Kollineationsgruppe G einer 4-dimensionalen Translationsebene halte zwei Geraden durch $0 \in R^4$ fest und wirke auf der einen Geraden trivial und auf der anderen Geraden als Gruppe von Drehstreckungen. Dann ist die Ebene desarguessch.*

Beweis. Wir können das Koordinatensystem des R^4 so wählen, daß G die Teilräume W und S festhält und aus den Matrizen der Form

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & a & b \\ & & -b & a \end{pmatrix}$$

besteht, und daß ferner der Teilraum $E = \{(x, y, u, v);\ u = x,\ v = y\}$ zum Geradenbüschel durch Null gehört. Dann entsprechen die Geraden des Büschels durch Null den Matrizen $\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, die Geometrie ist also desarguessch.

Homogene 4-dimensionale Translationsebenen

Wir folgern aus Homogenitätsvoraussetzungen den Satz von Desargues, und zwar beweisen wir: Ist Γ transitiv auf der Translationsachse oder ist $\dim \Gamma \geqq 9$, so ist die Ebene desarguessch.

Satz 3. *Die Kollineationsgruppe Γ einer 4-dimensionalen Translationsebene wirke transitiv auf der Translationsachse. Dann ist die Translationsebene desarguessch.*

Beweis. Wir können annehmen, daß Γ_0 den Punkt $0 \in R^4$ festhält und linear auf dem R^4 wirkt. Nach Voraussetzung ist Γ_0 transitiv auf dem zur 2-Sphäre homöomorphen Büschel der Geraden durch den Nullpunkt, und wir zeigen:

(a) Es gibt eine kompakte, zusammenhängende Untergruppe K von Γ_0, die transitiv auf dem Büschel $\mathfrak{B}$ wirkt.

Beweis. Γ_0 ist als lineare Gruppe von abzählbarer Basis, daher hat nach dem Lemma über homogene Räume die offene Untergruppe $\Delta = (\Gamma_0)^1$ auf $\mathfrak{B}$ nur offene Bahnen, ist also transitiv auf $\mathfrak{B}$. Da die 2-Sphäre $\mathfrak{B}$ kompakt und einfach zusammenhängend ist, gibt es nach

[11; 12, 5.6] eine kompakte Untergruppe K von Δ, die auch transitiv auf $\mathfrak{B}$ wirkt. Wie oben bleibt die Transitivität beim Übergang zur Zusammenhangskomponente erhalten.

(b) K ist als topologische Gruppe isomorph zu Spin(3), der einfach zusammenhängenden Überlagerungsgruppe der Gruppe SO_3.

Beweis. Die effektiv gemachte kompakte zusammenhängende Gruppe $K/K_{[\mathfrak{B}]}$ wirkt auf der 2-Sphäre $\mathfrak{B}$ wie die Gruppe SO_3 [13; 14; 18]. Nach Lemma 3 ist $K_{[\mathfrak{B}]}$ eine kompakte Untergruppe der multiplikativen Gruppe der reellen Zahlen und folglich isomorph zu 1 oder zu Z_2. Wäre $K_{[\mathfrak{B}]}=1$, dann würde $K=SO_3$ linear auf dem R^4 wirken. Nach [9, 35.4] ist die Wirkung der kompakten Gruppe SO_3 vollständig reduzibel, und da nach [15, Seite 269] die Gruppe SO_3 irreduzible reelle Darstellungen nur in ungeraden Dimensionen hat, folgt, daß K einen eindimensionalen Teilraum des R^4 festläßt. Dann bleibt auch die durch diesen Teilraum bestimmte Gerade fest, im Widerspruch zur Transitivität von K auf $\mathfrak{B}$. Es folgt $K_{[\mathfrak{B}]}=Z_2$, und daraus ergibt sich die Behauptung.

(c) $K=\mathrm{Spin}(3)$ wirkt als lineare Gruppe wie die reell aufgefaßte spezielle unitäre Gruppe $SU(2, C)$.

Beweis. Dies entnimmt man entweder der allgemeinen Theorie der Darstellungen von Lie-Gruppen [9; 28] oder sieht es unter Benutzung der Geometrie wie folgt: Die Standgruppe K_W auf einer Geraden $W\in\mathfrak{B}$ ist isomorph zu SO_2 und hält nach Hilfssatz 4 noch eine Gerade $S\in\mathfrak{B}$ fest. Wäre K_W nicht transitiv auf dem Raum P_W der eindimensionalen Teilräume von W, so würde K_W auf W die Identität induzieren, und nach Lemma 7 oder nach dem Viereck-Lemma ergäbe sich ein Widerspruch. Es folgt, daß K transitiv auf dem Raum der eindimensionalen Teilräume des R^4 ist. Nach [9, 35.1] kann man annehmen, daß $K=$ Spin(3) orthogonal auf dem R^4 wirkt und damit transitiv auf der Einheitssphäre M des R^4 operiert. Sei $m\in M$, dann ist K_m 0-dimensional und folglich zentral, das heißt, $K_m=0$ oder $K_m=Z_2$. Im zweiten Fall würde folgen, daß der reelle 3-dimensionale projektive Raum homöomorph zur 3-Sphäre ist, ein Widerspruch. Es gilt also $K_m=0$, und die Wirkung von Spin(3) auf M ist äquivalent zur Wirkung der Spin(3) durch Rechtstranslationen auf sich selbst. Hieraus ergibt sich die Behauptung.

(d) Die Translationsebene ist desarguessch.

Beweis. Die Gruppe $SU(2, C)$ können wir darstellen auf dem Raum H der Quaternionen als Gruppe der Rechtsmultiplikationen mit Quaternionen der Norm 1: $K=\mathrm{Spin}(3)=\{\rho_s;\ s\in H,\ |s|=1\}$. Um zu zeigen, daß Spin(3) normal in SO_4 ist, betrachten wir die Überlagerungsabbildung $\varphi=((a,b)\mapsto(x\mapsto a\,x\,b^{-1},\ x\in H))$: $K\times K\to SO_4$ der einfach zusammenhängenden Überlagerungsgruppe $K\times K$ von SO_4 auf SO_4 und benutzen die Surjektivität von φ [27, Seite 74]: Sei $s \in H$, $|s|=1$ und $\varphi(a,b) \in SO_4$ gegeben, dann ist $[\varphi(a,b)]^{-1} \rho_s \varphi(a,b) = \varphi(a^{-1}, b^{-1}) \rho_s \varphi(a,b) = \rho_{bsb^{-1}}$ wieder Rechtsmultiplikation mit einer Quaternion der Norm 1. Sei nun $w \in SO_4$ so, daß für eine Gerade E des Büschels $\mathfrak{B}$ gilt $E^w = W = \{(x,y,u,v);\ u=v=0\}$. Wir weisen nach, daß unter w die K-Bahn von E übergeht in die K-Bahn von W: Da K Normalteiler von SO_4 ist, gilt: $(E^K)^w = (E^{wKw^{-1}})^w = W^K$. Das Büschel $\mathfrak{B}$ ist also linear isomorph zum desarguesschen Büschel der Geraden mit komplexer Steigung in C^2 und daher selbst desarguessch.

Satz 4. *Jede 4-dimensionale Translationsebene mit* $\dim \Gamma \geqq 9$ *ist desarguessch.*

Beweis. Für die Zusammenhangskomponente $\Delta = (\Gamma_0)^1$ gilt $\dim \Delta \geqq 5$, und wir zeigen:

(a) Δ hält eine Gerade $S \in \mathfrak{B}$ fest und wirkt transitiv auf $\mathfrak{C} = \mathfrak{B} - \{S\}$.

Beweis. Δ hat auf $\mathfrak{B}$ keine eindimensionale Bahn, sonst lieferte Fixieren dreier Geraden dieser Bahn nach Lemma 6 den Satz von Desargues. Hätte Δ nur 2-dimensionale Bahnen auf $\mathfrak{B}$, so wären diese offen und Δ wäre transitiv auf $\mathfrak{B}$. Dann liefert Satz 3, daß die Geometrie desarguessch ist. Δ hält folglich eine Gerade $S \in \mathfrak{B}$ fest. Hätte Δ auf $\mathfrak{C}$ eine 0-dimensionale oder eine 1-dimensionale Bahn, so ergäbe sich mit Lemma 6 der Satz von Desargues. Daraus folgt, daß Δ auf $\mathfrak{C}$ nur 2-dimensionale Bahnen hat und folglich transitiv auf $\mathfrak{C}$ ist.

(b) $\Delta_{[S]}$ hat auf $\mathfrak{C}$ eine eindimensionale Bahn $\mathfrak{X}$.

Beweis. Es gilt $\dim \Delta_{[S]} \geqq 1$, da die lineare Wirkung auf S höchstens 4-dimensional ist. Sei zunächst $\dim \Delta_{[S]} = 1$, dann hat $\Delta_{[S]}$ auf $\mathfrak{C}$ nur 0- oder 1-dimensionale Bahnen. Wären alle Bahnen 0-dimensional, so ergäbe sich ein Widerspruch zum Viereck-Lemma. Jetzt sei $\dim \Delta_{[S]} \geqq 2$. Falls $\Delta_{[S]}$ transitiv auf $\mathfrak{C}$ wirkt, dann ist nach [16, 3.5.40] der Ternärkörper distributiv und nach [10; 21, 7.25] isomorph zu C. Es gibt also noch 0- oder 1-dimensionale Bahnen. Es können aber keine 0-dimensionale Bahnen mehr auftreten, sonst hielte $(\Delta_{[S]})^1$ noch eine Gerade $W \in \mathfrak{B}$ fest, und mit Hilfssatz 3, Lemma 7 und dem Viereck-Lemma würde folgen, daß die Geometrie desarguessch ist.

(c) Da $\Delta_{[S]}$ Normalteiler von Δ ist, bilden die 1-dimensionalen Mengen $\{\mathfrak{X}^\delta, \delta \in \Delta\}$ Imprimitivitätsgebiete für die Wirkung von Δ auf $\mathfrak{C}$. Die Standgruppe Δ_W auf $W \in \mathfrak{C}$ ist mindestens 3-dimensional und hält das eindimensionale Imprimitivitätsgebiet $[W]$ fest. Fixieren einer Geraden $K \in [W]$ liefert nach Lemma 6, daß die Geometrie desarguessch ist.

Lemma 8. *Die 2-dimensionale Torusgruppe T wirke als Kollineationsgruppe einer 4-dimensionalen Translationsebene und halte den Punkt $0 \in R^4$ fest. Dann ist die Geometrie desarguessch.*

Beweis. Sei $\mathfrak{B}$ das zur 2-Sphäre homöomorphe Geradenbüschel durch 0. Eine zu SO_2 isomorphe Untergruppe von T hält nach Hilfssatz 4 zwei Geraden $S, W \in \mathfrak{B}$ fest und hat sonst nur Kreisbahnen. Die Gruppe T hat auf $\mathfrak{B}$ die gleichen Bahnen, und Festhalten einer Geraden einer solchen Kreisbahn liefert eine eindimensionale Gruppe, die $\mathfrak{B}$ elementweise festhält und keine Diagonalmatrizen enthält. Nach Lemma 3 folgt die Behauptung.

Eine Klasse nicht desarguesscher 4-dimensionaler Translationsebenen mit 8-dimensionaler Kollineationsgruppe

Satz 5. *Gegeben sei eine nicht desarguessche* 4-*dimensionale Translationsebene mit* $\dim \Gamma_0 = 4$. *Ferner halte* $(\Gamma_0)^1$ *keine Gerade des Büschels* $\mathfrak{B}$ *fest und wirke reduzibel auf dem* R^4. *Dann wird die Ebene bis auf Isomorphie von folgendem Büschel erzeugt:*

$$\mathfrak{B}_w = \{S\} \cup \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}; \beta \geqq 0 \right\} \cup \left\{ \begin{pmatrix} \alpha & \beta \\ -w\beta & \alpha \end{pmatrix}; \beta \leqq 0 \right\} \quad \text{mit } w > 1.$$

Umgekehrt gibt es zu jedem $w > 1$ *eine Translationsebene* $\mathbf{P}_w = \mathbf{P}(\mathfrak{B}_w)$ *der genannten Art, und zwei Ebenen* $\mathbf{P}_w$ *und* $\mathbf{P}_{w'}$, $w, w' > 1$, *sind genau dann isomorph, wenn* $w = w'$ *ist. Die Gruppe* $(\Gamma_0)^1$ *wirkt als Matrizengruppe*

$$\left\{ \begin{pmatrix} a & & b & \\ & a & & b \\ c & & d & \\ & c & & d \end{pmatrix}; ad - bc > 0 \right\}$$

und hat auf $\mathfrak{B}$ *drei Bahnen, von denen genau eine homöomorph zur Kreislinie* $\mathfrak{S}$ *ist. Es gilt* $\Gamma_0/(\Gamma_0)^1 = Z_2 \times Z_2$, *und die Nebenklassen entsprechen den Kollineationen, welche die Orientierung von* $\mathfrak{S}$ *ändern oder die beiden Teilbüschel vertauschen oder beides gleichzeitig bewirken.*

Beweis. (a) $\Delta = (\Gamma_0)^1$ hat auf $\mathfrak{B}$ drei Bahnen, von denen genau eine homöomorph zur Kreislinie $\mathfrak{S}$ ist, und es gilt $\Delta/\Delta_{[\mathfrak{S}]} \cong PSL_2(R)$ auf $\mathfrak{S}$.

Beweis. Nach Satz 3 ist Δ nicht transitiv auf $\mathfrak{B}$, und da Δ keine Gerade des Büschels $\mathfrak{B}$ festläßt, hat Δ auf $\mathfrak{B}$ mindestens eine eindimensionale Bahn. Wäre die effektive Wirkung von Δ auf dieser Bahn wie R auf R, wie L_2 auf R, wie Ω^∞ auf R, wie SO_2 auf S oder wie eine echte Überlagerung von Ω auf S, dann ergäbe sich durch Festhalten zweier Geraden dieser Bahn eine 2-dimensionale Gruppe, die mindestens drei Geraden der Bahn festhält, und die Geometrie wäre nach Lemma 6 desarguessch. Nach dem Satz von Brouwer folgt, daß die eindimensionale Bahn $\mathfrak{S}$ homöomorph zur Kreislinie ist und daß $\Delta/\Delta_{[\mathfrak{S}]} \cong \Omega$ auf $\mathfrak{S}$.

Angenommen, es gäbe eine zweite Kreisbahn $\mathfrak{S}'$, dann wäre auch $\Delta/\Delta_{[\mathfrak{S}']} \cong \Omega$ auf $\mathfrak{S}'$ und Festhalten einer Geraden $L \in \mathfrak{S}$ hielte auch eine Gerade $L' \in \mathfrak{S}'$ fest. Durch Fixieren einer weiteren Geraden $K \in \mathfrak{S}$ ergäbe sich eine 2-dimensionale Gruppe mit drei Fixgeraden, und die Geometrie wäre nach Lemma 6 desarguessch.

Nun sei $H = \{\delta \in \Delta, \det \delta = 1\}$, dann gilt

(b) $H \cong SL_2(R)$, und wir können das (x, y, u, v)-Koordinatensystem des R^4 so wählen, daß H als Matrizengruppe

$$\left\{ \begin{pmatrix} a & & b & \\ & a & & b \\ c & & d & \\ & c & & d \end{pmatrix};\ a\,d - b\,c = 1 \right\}$$

wirkt, und daß der durch $x = y = 0$ gegebene Teilraum S zur Geradenbahn $\mathfrak{S}$ gehört und die Standgruppe H_S durch $b = 0$ gegeben wird.

Beweis. Da Γ_0 die Streckungsgruppe R^* und folglich $\Delta = (\Gamma_0)^1$ die positive Streckungsgruppe $R^{>0}$ enthält, gilt $\dim H = 3$. Ferner ist die Zusammenhangskomponente H^1 transitiv auf $\mathfrak{S}$ und $H^1/H^1_{[\mathfrak{S}]} \cong PSL_2(R)$. Es folgt, daß H^1 eine Überlagerungsgruppe von $\Omega = PSL_2(R)$ ist, also $H^1 = \Omega^{\infty}/nZ$ für ein $n \in Z$, insbesondere ist $H = H^1$ und H ist halbeinfach. Nach dem Conducible Theorem [9, 35.4] ist die Darstellung von H auf R^4 vollständig reduzibel. Die Gruppe H läßt keinen eindimensionalen Teilraum des R^4 fest, sonst bliebe auch die durch diesen Teilraum bestimmte Gerade fest, im Widerspruch zur Wirkung von Δ (und H) auf $\mathfrak{B}$. Da nach Voraussetzung Δ und folglich H auf dem R^4 reduzibel wirkt, hält H zwei 2-dimensionale komplementäre Teilräume F und G des R^4 fest. Seien φ_F und φ_G die Darstellungen von H auf F und auf G, und sei δ die Determinantenabbildung. Wegen des Zusammenhangs von H gilt dann $H^{\varphi_F \circ \delta} = (0, \infty)$ oder $= \{1\}$. Der erste Fall scheidet aus, sonst wäre der Kern von $\varphi_F \circ \delta$ ein 2-dimensionaler Normalteiler von H. Die Gruppe H wirkt nicht trivial auf F, sonst blieben drei eindimensionale Teilräume von F und die durch sie bestimmten Geraden fest, und die Geometrie wäre nach Lemma 6 desarguessch. Da alle echten Normalteiler von H 0-dimensional sind, ist die zusammenhängende Gruppe H^{φ_F} 3-dimensional und stimmt mit $SL(F)$ überein. Daraus folgt für die Zentren $Z(H)^{\varphi_F} \subset Z(SL(F)) = \{-1, 1\}$. Wäre $Z(H)^{\varphi_F} = \{1\}$, dann würde folgen Kern $\varphi_F = Z(H)$ und $H/Z(H) = PSL_2(R) = SL_2(R)$, ein Widerspruch. Daher ist $Z(H)^{\varphi_F} = \{-1, 1\}$, und es folgt Kern $\varphi_F \cap Z(H) = Z(H)/2$. Entsprechend gilt Kern $\varphi_G \cap Z(H) = Z(H)/2$. Wegen Kern $\varphi_F \cap$ Kern $\varphi_G = 1$ folgt $Z(H)/2 = 1$, also $|Z(H)| = 2$ und folglich $H = SL_2(R)$.

Wir erhalten also $H=SL_2(R)$, und H wirkt als Matrizengruppe

$$\left\{\begin{pmatrix} a & b & & \\ c & d & & \\ & & a & b \\ & & c & d \end{pmatrix};\ a\,d-b\,c=1\right\}.$$

Man beachte jedoch, daß die beiden ausgezeichneten Teilräume F und G nicht zu $\mathfrak{B}$ gehören, denn das widerspräche der Wirkung von Δ auf $\mathfrak{B}$. Die Gruppe $SL_2(R)$ wirkt transitiv auf dem Raum P_F der eindimensionalen Teilräume von F und transitiv auf P_G. Da jedes Element von P_F genau eine Gerade bestimmt, und zwar verschiedene Elemente verschiedene Geraden, bestimmt P_F eine zur Kreislinie homöomorphe Geradenbahn. Entsprechendes gilt für P_G, und da es genau eine zur Kreislinie homöomorphe Geradenbahn gibt, stimmen die beiden Geradenbahnen mit $\mathfrak{S}$ überein, das heißt, die Geraden aus $\mathfrak{S}$ verbinden die eindimensionalen Teilräume von F und diejenigen von G paarweise. Indem wir eine dieser Geraden als ausgezeichnete Gerade S nehmen und ein (x, y, u, v)-Koordinatensystem des R^4 geeignet wählen, können wir die Wirkung von $H=SL_2(R)$ annehmen als

$$\left\{\begin{pmatrix} a & & b & \\ & a & & b \\ c & & d & \\ & c & & d \end{pmatrix};\ a\,d-b\,c=1\right\},$$

wobei die Standgruppe auf S durch $b=0$ gegeben wird.

(c) Konstruktion des Büschels $\mathfrak{B}$ aus der Kollineationsgruppe.

Die Bahn der Geraden S unter H besteht aus S und den durch $u=r\,x$, $v=r\,y$, $r\in R$, gegebenen Teilräumen, das heißt

$$\mathfrak{S}=\{S\}\cup\left\{\begin{pmatrix} r & \\ & r \end{pmatrix},\ r\in R\right\}.$$

Die Bahn des Teilraumes $\begin{pmatrix} 0 & 1 \\ r & s \end{pmatrix}$, $r\neq 0$, unter H_S besteht aus den Matrizen $\left[\begin{pmatrix} c & \\ & c \end{pmatrix}+\begin{pmatrix} d & \\ & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ r & s \end{pmatrix}\right]\begin{pmatrix} a & \\ & a \end{pmatrix}^{-1}$, und wegen $a\,d=1$ ergibt sich die Bahn

$$\mathfrak{E}_1=\left\{\begin{pmatrix} c\,d & d^2 \\ d^2\,r & d^2\,s+c\,d \end{pmatrix};\ \begin{matrix} r, s\in R, & r\neq 0 \text{ fest} \\ c, d\in R, & d\neq 0 \text{ variabel} \end{matrix}\right\}.$$

Entsprechend ergibt sich als H_S-Bahn der Geraden $\begin{pmatrix} 0 & -1 \\ w & z \end{pmatrix}$:

$$\mathfrak{E}_2 = \left\{ \begin{pmatrix} c\,d & -d^2 \\ w\,d^2 & d^2 z + c\,d \end{pmatrix};\ \begin{matrix} w, z \in R, & w \neq 0 \text{ fest} \\ c, d \in R, & d \neq 0 \text{ variabel} \end{matrix} \right\}.$$

Das System $\mathfrak{S} \cup \mathfrak{E}_1 \cup \mathfrak{E}_2$ liefert folgende Abbildung τ der (α, β)-Ebene: Für $\beta \geqq 0$ stimmt τ überein mit der linearen Abbildung

$$\tau_1 = ((\alpha, \beta) \mapsto (r\,\beta, \alpha + s\,\beta)), \qquad r \neq 0,$$

für $\beta \leqq 0$ mit

$$\tau_2 = ((\alpha, \beta) \mapsto (-w\,\beta, \alpha - z\,\beta)), \qquad w \neq 0.$$

Da die Geometrie nach Voraussetzung nicht desarguessch ist, sind die beiden Beschränkungen von τ_1 und τ_2 bei der $(\beta = 0)$-Achse schief zusammengesetzt.

(d) τ ist ein transversaler Homöomorphismus genau dann, wenn $s^2 + 4r < 0$ und $z^2 - 4w < 0$ gilt. (Das heißt $r < 0$, $w > 0$ und τ_1 und τ_2 elliptisch.)

Beweis. Sei zunächst (etwa) τ_1 nicht elliptisch, dann hält τ_1 eine Gerade durch den Nullpunkt fest, und diese Gerade wird unter τ nicht transversal abgebildet. Jetzt seien umgekehrt beide Bedingungen erfüllt. Dann zeigen wir zunächst, daß τ „schlicht" ist, daß also für je zwei gewöhnliche parallele Geraden H, K gilt $|H^\tau \cap K| \leqq 1$, das heißt, daß für alle $(\alpha_1, \beta_1) \neq (\alpha_2, \beta_2)$ gilt $\left| \begin{pmatrix} \alpha_1 & \beta_1 \\ f_1 & g_1 \end{pmatrix} - \begin{pmatrix} \alpha_2 & \beta_2 \\ f_2 & g_2 \end{pmatrix} \right| \neq 0$. Da τ_1 und τ_2 elliptisch sind, können wir uns auf den Fall $\beta_1 > 0$, $\beta_2 < 0$ beschränken. Außerdem können wir $A = \alpha_1 - \alpha_2 \neq 0$ annehmen. Wir setzen $\beta_1 = \beta > 0$, $\beta_2 = -p\,\beta$ mit $p > 0$, dann ergibt sich

$$\left| \begin{pmatrix} \alpha_1 & \beta_1 \\ \beta_1 r & \alpha_1 + \beta_1 s \end{pmatrix} - \begin{pmatrix} \alpha_2 & \beta_2 \\ -\beta_2 w & \alpha_2 - \beta_2 z \end{pmatrix} \right| = f(\beta)$$
$$= A^2 + A(s - z\,p)\,\beta + (1 + p)(p\,w - r)\,\beta^2.$$

Wir zeigen, daß die Diskriminante D dieser quadratischen Funktion in β negativ ist: $D/A^2 = p^2(z^2 - 4w) + p(-2s\,z + 4r - 4w) + s^2 + 4r$. Falls $-2s\,z + 4r - 4w \leqq 0$, dann gilt wegen $p > 0$ und wegen $z^2 - 4w < 0$, $s^2 + 4r < 0$, daß $D/A^2 < 0$ ist. Wir können also annehmen, daß $-2s\,z + 4r - 4w > 0$ ist. Dann berechnen wir die Diskriminante d der quadratischen Gleichung für p und schätzen diese mit Hilfe der Ungleichungen $-2s\,z + 4r - 4w > 0$, $s^2 + 4r < 0$, $z^2 - 4w < 0$ ab:

$$d/4 = (-s\,z + 2r - 2w)^2 - (z^2 - 4w)(s^2 + 4r) < -4(r + w)^2 \leqq 0.$$

Die Diskriminante d ist negativ, und daher ist D/A^2 wegen $s^2+4r<0$ stets negativ, und folglich $f(\beta)$ für $p>0$ stets positiv. Damit ist die Schlichtheit von τ gezeigt.

Wäre τ nicht transversal, dann gäbe es eine Gerade H, die nicht transversal abgebildet wird. Das Bild H besteht aus zwei geraden Linien, die bei $\alpha=0$ schief zusammengesetzt sind, daher schneidet H einige Parallelen von H zweimal im Widerspruch zur Schlichtheit von τ.

Schließlich ist τ als Zusammenheftung der Beschränkungen $\tau_{1/\beta\geqq 0}$ und $\tau_{2/\beta\leqq 0}$ ein Homöomorphismus.

(e) Überführung in die Normalform $\mathbf{P}_w$, $w>1$.

Auf die Partition $\mathfrak{B}(r, s, w, z)$ wenden wir die lineare Abbildung

$$\lambda_1 = \begin{pmatrix} 1 & 0 & & \\ -s(-s^2-4r)^{-\frac{1}{2}} & 2(-s^2-4r)^{-\frac{1}{2}} & & \\ & & 1 & 0 \\ & & -s(-s^2-4r)^{-\frac{1}{2}} & 2(-s^2-4r)^{-\frac{1}{2}} \end{pmatrix}.$$

Diese Abbildung führt das Halbbüschel $\left\{\begin{pmatrix} \alpha & \beta \\ r\beta & \alpha+s\beta \end{pmatrix};\ \beta\geqq 0\right\}$ über in das Halbbüschel $\left\{\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix};\ \beta\geqq 0\right\}$, wir erhalten also $r'=-1$, $s'=0$, und w, z gehen unter λ_1 über in Zahlen w', z' mit $(w', z')\neq(1, 0)$. Nun wenden wir die Abbildung

$$\lambda_2 = \begin{pmatrix} 1 & b & & \\ -b & 1 & & \\ & & 1 & b \\ & & -b & 1 \end{pmatrix}$$

an, dann geht das durch $r'=-1$ und $s'=0$ bestimmte Halbbüschel in sich über und z' geht über in $z''=\left(z' b^2+2(w'-1)\, b-z'\right)/\neq 0$. Die Diskriminante der quadratischen Gleichung für b im $\cdot$ Zähler ist $D= 4(w'-1)^2+4z'^2>0$, daher läßt sich die Zahl b in der Abbildung λ_2 so finden, daß $z''=0$ wird. Falls $w''>1$, dann sind wir fertig, und für $w''<1$ führen wir noch die lineare Abbildung

$$\lambda_3 = \begin{pmatrix} 0 & 1 & & \\ -1 & 0 & & \\ & & 0 & -1 \\ & & 1 & 0 \end{pmatrix}$$

aus. Sie führt das durch $r=-1$, $s=0$ gegebene Halbbüschel in sich über, und $w''<1$ geht über in $w'''=1/w''>1$, während $z'''=z''=0$ gilt. Bei allen Abbildungen geht die Geradenbahn $\mathfrak{S}$ in sich über.

(f) Berechnung der vollen Standgruppe Γ_0.

Zunächst zeigen wir, daß

$$\Delta=\left\{\begin{pmatrix} a & & b & \\ & a & & b \\ c & & d & \\ & c & & d \end{pmatrix};\ ad-bc>0\right\}$$

in Γ_0 enthalten ist: Anwendung eines Elementes aus Δ auf die Gerade $\begin{pmatrix} \alpha & \beta \\ -w\beta & \alpha \end{pmatrix}$, $\beta<0$, ergibt die Gerade

$$\begin{pmatrix} \alpha' & \beta' \\ f' & g' \end{pmatrix}=\left[\begin{pmatrix} a & \\ & a \end{pmatrix}+\begin{pmatrix} b & \\ & b \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ -w\beta & \alpha \end{pmatrix}\right]\left[\begin{pmatrix} c & \\ & c \end{pmatrix}+\begin{pmatrix} d & \\ & d \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ -w\beta & \alpha \end{pmatrix}\right]^{-1}.$$

Man rechnet nach, daß $f'=-w\beta'$ und $g'=\alpha'$ gilt, die Bildgerade liegt also wieder im „unteren" Halbbüschel. Entsprechend ergibt sich, daß das obere Halbbüschel in sich übergeht und daß $\mathfrak{S}$ orientierungstreu auf sich abgebildet wird.

Nun beweisen wir, daß jedes Element $\gamma\in\Gamma_0$, welches $\mathfrak{S}$ elementweise festhält und $\mathfrak{E}_1$ und $\mathfrak{E}_2$ jeweils in sich überführt, Diagonalform hat: Das Element γ hat die Form $\begin{pmatrix} a & b & & \\ c & d & & \\ & & a & b \\ & & c & d \end{pmatrix}$ und führt die Gerade $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, $\beta>0$, über in

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$$
$$=\begin{pmatrix} ad\alpha-bd\beta-ac\beta-bc\alpha & b^2\beta+a^2\beta \\ -d^2\beta-c^2\beta & -bc\alpha+bd\beta+ac\beta+ad\alpha \end{pmatrix}\cdot\frac{1}{D}.$$

Die Bildgerade soll wieder im oberen Halbbüschel $\mathfrak{E}_1$ liegen, und dies liefert $bd+ac=0$, $e^2+b^2=c^2+d^2$, $ad-bc>0$. Die Gerade $\begin{pmatrix} \alpha & \beta \\ -w\beta & \alpha \end{pmatrix}$, $\beta<0$, des unteren Halbbüschels $\mathfrak{E}_2$ geht über in

$$\begin{pmatrix} ad\alpha-bdw\beta-ac\beta-bc\alpha & b^2w\beta+a^2\beta \\ -d^2w\beta-c^2\beta & -bc\alpha+bdw\beta+ac\beta+ad\alpha \end{pmatrix}\cdot\frac{1}{D}.$$

Sie liegt wieder im unteren Halbbüschel, wenn $-wa^2-w^2b^2+c^2+wd^2=0$, $bdw+ac=0$ und $ad-bc>0$ gilt. Aus diesen Gleichungen ergibt sich $bd=bdw$, also wegen $w>1$ $bd=0$. Wäre $d=0$, dann folgte

$b \neq 0$, $c \neq 0$, $a=0$, $b^2=c^2$ und $b^2=w^2 b^2$. Dies lieferte $w^2=1$, ein Widerspruch. Es folgt $d \neq 0$, $b=0$, $a \neq 0$, $c=0$ und daraus $a^2=d^2$. Wegen $D=ad>0$ ergibt sich die Diagonalität von γ.

Aus der letzten Behauptung folgt, daß Δ der Normalteiler von Γ_0 ist, welcher aus den Elementen besteht, die $\mathfrak{S}$ orientierungstreu abbilden und $\mathfrak{E}_1$ und $\mathfrak{E}_2$ jeweils in sich überführen. Da für jedes Element $\gamma \in \Gamma_0$ nur die vier Möglichkeiten bestehen, $\mathfrak{S}$ orientierungstreu oder orientierungsändernd auf sich abzubilden und $\mathfrak{E}_1$ und $\mathfrak{E}_2$ zu belassen oder zu vertauschen, folgt $|\Gamma_0:(\Gamma_0)^1| \leq 4$. Daß alle drei Nebenklassen auftreten, zeigen wir durch Angabe von Repräsentanten: Seien

$$\gamma_1=\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \quad \gamma_2=\begin{pmatrix} 0 & 1 & & \\ w & 0 & & \\ & & 0 & 1 \\ & & w & 0 \end{pmatrix} \quad \gamma_3=\begin{pmatrix} 0 & 1 & & \\ -w & 0 & & \\ & & 0 & -1 \\ & & w & 0 \end{pmatrix},$$

dann ändert γ_1 die Orientierung von $\mathfrak{S}$ und führt $\mathfrak{E}_1$ und $\mathfrak{E}_2$ jeweils in sich über, γ_2 führt $\mathfrak{S}$ orientierungserhaltend in sich über und vertauscht $\mathfrak{E}_1$ und $\mathfrak{E}_2$, und γ_3 ändert die Orientierung von $\mathfrak{S}$ und vertauscht $\mathfrak{E}_1$ und $\mathfrak{E}_2$. Wir ersparen uns die einfachen Rechnungen.

Literatur

1. Adams, J. F.: Lectures on Lie groups. New York, Amsterdam: Benjamin 1969.
2. André, J.: Über nicht-Desarguessche Ebenen mit transitiver Translationsgruppe. Math. Z. **60**, 156–186 (1954).
3. Arens, R.: Topologies for homeomorphism groups. Amer. J. Math. **68**, 593–610 (1946).
4. Betten, D.: Nicht-desarguessche 4-dimensionale Ebenen. Arch. der Math. **21**, 100–102 (1970).
5. Breitsprecher, S.: Projektive Ebenen, die Mannigfaltigkeiten sind. Math. Z. **121**, 157–174 (1971).
6. Breuning, P.: Translationsebenen und Vektorraumbündel. Mitt. Math. Sem. Gießen, Heft 68, 1970.
7. Brouwer, L. E. J.: Die Theorie der endlichen kontinuierlichen Gruppen, unabhängig von den Axiomen von Lie. Math. Ann. **67**, 246–267 (1909).
8. Dugundji, J.: Topology. Boston: Allyn and Bacon 1966.
9. Freudenthal, H., Vries, H. de: Linear Lie groups. New York and London: Academic Press 1969.
10. Hofmann, K. H.: Topologische distributive Doppelloops. Math. Z. **71**, 36–68 (1959).
11. Montgomery, D.: Simply connected homogeneous spaces. Proc. Amer. Math. Soc. **1**, 467–469 (1950).
12. Montgomery, D., Zippin, L.: Topological transformation groups. New York: Interscience 1955.
13. Montgomery, D., Samelson, H.: Transformation groups of spheres. Ann. of Math. **44**, 454–470 (1943).
14. Mostow, G. D.: The extensibility of local Lie groups of transformations and groups on surfaces. Ann. of Math. **52**, 606–636 (1950).

15. Murnaghan, F. D.: The theory of group representations. New York: Dover 1963.
16. Pickert, G.: Projektive Ebenen. Berlin-Göttingen-Heidelberg: Springer 1955.
17. Plaumann, P., Strambach, K.: Zweidimensionale Quasikörper mit Zentrum. Arch. der Math. **21**, 455–465 (1970).
18. Poncet, J.: Groupes de Lie compact de transformations de l'espace euclidien et les spheres comme espaces homogenes. Commentarii Math. Helvet. **33**, 109–120 (1959).
19. Salzmann, H.: Kompakte zweidimensionale projektive Ebenen. Math. Ann. **145**, 401–428 (1962).
20. Salzmann, H.: Zur Klassifikation topologischer Ebenen. Math. Ann. **150**, 226–241 (1963).
21. Salzmann, H.: Topological planes. Advances Math. **2**, 1–60 (1967).
22. Salzmann, H.: Homomorphismen komplexer Ternärkörper. Math. Z. **112**, 23–25 (1969).
23. Salzmann, H.: Kompakte vier-dimensionale Ebenen. Arch. der Math. **20**, 551–555 (1969).
24. Salzmann, H.: Kollineationsgruppen kompakter vier-dimensionaler Ebenen. Math. Z. **117**, 112–124 (1970).
25. Salzmann, H.: Kollineationsgruppen kompakter vier-dimensionaler Ebenen II. Math. Z. **121**, 104–110 (1971).
26. Skornjakov, L. A.: Topologische projektive Ebenen. Trudy Moskov. Mat. Obšč. **3**, 347–373 (1954).
27. Tits, J.: Tabellen zu den einfachen Lie-Gruppen und ihren Darstellungen. Lecture Notes in Math. 40. Berlin-Heidelberg-New York: Springer 1967.
28. Tits, J.: Liesche Gruppen und Algebren. Vorlesung Bonn 1963/64.

Dr. D. Betten
Mathematisches Institut der Universität
D-7400 Tübingen
Hölderlinstraße 19
Bundesrepublik Deutschland

(Eingegangen am 13. April 1972)

Math. Z. 128, 153–167 (1972)

Varietal Homology and Parafree Groups

Urs Stammbach

Introduction

The notion of parafree groups has been introduced by Baumslag in [3, 4]. A definition of parafreeness (which is equivalent to Baumslag's definition by Proposition 4.3) is as follows. A group G in the variety $\mathfrak{V}$ is $\mathfrak{V}$-*parafree* if

(i) G is residually nilpotent;

(ii) there exists a $\mathfrak{V}$-free group F and a homomorphism $\vartheta\colon F \to G$ such that for all $n \geqq 1$ ϑ induces isomorphisms

$$\vartheta_n\colon F/F_n \xrightarrow{\sim} G/G_n.$$

Here F_n, G_n denote as usual the terms of the lower central series of F, G. In this paper we try to show that the homology theory associated with the variety $\mathfrak{V}$ may be used to great advantage in exhibiting properties of $\mathfrak{V}$-parafree groups. Our approach not only allows to simplify many proofs considerably, but it also leads to new insights and generalizations of theorems due to Baumslag.

Section 1 is devoted to the varietal homology. The homology theory associated with the variety $\mathfrak{V}$ is a special case of the categorical homology theory defined by André [1], Barr-Beck [2], Rinehart [12], Ulmer [16], a.o. Already in [15] it has been used successfully to obtain results in group theory. Its properties have recently been discussed by Leedham-Green [8, 9, 10]. In Section 1 we give a more or less selfcontained exposition with very simple proofs of the essential properties of $V(G, A)$, the first varietal homology group of G with trivial coefficients A. (Recall that $V(G, A)$ for $\mathfrak{V} = \mathfrak{Gr}$ is the ordinary second homology group with coefficients A.) For our applications the exact sequence (1.2) of our Theorem 1.1 associated with an extension in $\mathfrak{V}$ will be crucial.

Of course we restrict ourselves in Section 1 to those results which are used in the remaining parts of this paper; however we would like to remark that some of our arguments are valid in greater generality. We hope to come back to this point in a later paper.

In Section 2 we use the varietal homology to define (pseudo) stem extensions and (pseudo) stem covers in $\mathfrak{V}$. Following [6] we call the central extension $N \rightarrowtail G \xrightarrow{q} Q$ in $\mathfrak{V}$ a *stem extension* in $\mathfrak{V}$, if

$$\text{(i)}\quad q_*: G_{\mathrm{ab}} \xrightarrow{\sim} Q_{\mathrm{ab}}. \tag{0.1}$$

We call it a *stem cover* in $\mathfrak{V}$ if in addition

$$\text{(ii)}\quad q_*: V(G,\mathbb{Z}) \to V(Q,\mathbb{Z}) \quad \text{is the zero map}. \tag{0.2}$$

A not necessarily central extension $N \rightarrowtail G \twoheadrightarrow Q$ in $\mathfrak{V}$ satisfying (0.1) is called a *pseudo stem extension* in $\mathfrak{V}$; if it satisfies (0.1) and (0.2) it is called a *pseudo stem cover* in $\mathfrak{V}$. In Section 2 we list some elementary properties of (pseudo) stem extensions and covers.

Section 3 consists of the remark that a problem of Passi-Vermani [11] may be expressed naturally in terms of (pseudo) stem covers. Some results of Passi-Vermani are immediate consequences of that remark.

Section 4 is devoted entirely to parafree groups. We first prove that if G is $\mathfrak{V}$-parafree, then G is a pseudo stem cover in $\mathfrak{V}$ of all its quotients G/G_n, $n \geqq 2$ (Proposition 4.1). This leads to a new definition of $\mathfrak{V}$-parafreeness (Corollary 4.4): The group G is $\mathfrak{V}$-parafree if

(i) G is in $\mathfrak{V}$;

(ii) G is residually nilpotent;

(iii) G_{ab} is free in $\mathfrak{V} \cap \mathfrak{Ab}$ and G is a pseudo stem cover in $\mathfrak{V}$ of its quotients G/G_n, $n \geqq 2$.

Our Lemma 4.2 then is the key to most of the remaining propositions. Many of the results of Baumslag and generalizations may now be obtained in an almost standard way by methods similar to those already used in [15].

I would like to thank Karl Gruenberg for many helpful comments.

1. The Varietal Homology

We denote by $\mathfrak{Ab}$ the variety of all abelian groups, and by $\mathfrak{Gr}$ the variety of all groups.

For G in $\mathfrak{Gr}$ and A in $\mathfrak{Ab}$ the symbol $H_n(G,A)$, $n \geqq 0$ denotes the n-th homology group of G with coefficients A (trivial action). Recall that $H_1(G,A) = G_{\mathrm{ab}} \otimes A$, where $G_{\mathrm{ab}} = G/G'$ is the abelianized group of G. Also, $H_2(G,\mathbb{Z})$ is the Schur multiplicator.

Now let $\mathfrak{V}$ be a variety of groups. From [15] we recall (and generalize) the definition of the functor $V(-,A)$: $\mathfrak{V} \to \mathfrak{Ab}$ which, for $A = \mathbb{Z}_p$, was called S_p- in [15]. For G in $\mathfrak{V}$ the abelian group $V(G,A)$ is defined as follows. Let $q: F \twoheadrightarrow G$ be a $\mathfrak{V}$-free presentation of G, then

$$V(G,A) = \operatorname{coker}\big(q_*: H_2(F,A) \to H_2(G,A)\big). \tag{1.1}$$

It is easy to prove ([15]) that $V(G,A)$ does not depend upon the chosen $\mathfrak{V}$-free presentation of G, and that if $g: G \to H$ is a homomorphism in $\mathfrak{V}$, then $g_*: H_2(G,A) \to H_2(H,A)$ induces a (well-defined) map $g_*: V(G,A) \to V(H,A)$ making $V(-,A)$ into a (covariant) functor. It is even more trivial that for fixed G in $\mathfrak{V}$, $V(G,-)$: $\mathfrak{Ab} \to \mathfrak{Ab}$ is a (covariant) functor: in fact $V(-,-)$: $\mathfrak{V} \times \mathfrak{Ab} \to \mathfrak{Ab}$ is a bifunctor.

It is apparent from the definition that if G is $\mathfrak{V}$-free, $V(G,A)=0$. Also, if $\mathfrak{V}=\mathfrak{Gr}$ we have $V(G,A)=H_2(G,A)$. Indeed it may be shown (and we state it for the record) that $V(-,A)$ is the first homology group functor associated with the category $\mathfrak{V}$ (via the $\mathfrak{V}$-free group triple) provided A is in $\mathfrak{V} \cap \mathfrak{Ab}$ (see [2, 12, 8]).

We note that the proof of [15, Theorem 3.2] goes through in our more general setting; thus we may state

Theorem 1.1. *If* $N \rightarrowtail^{i} G \overset{q}{\twoheadrightarrow} Q$ *is an extension in* $\mathfrak{V}$, *and if* A *is in* $\mathfrak{Ab}$, *then the following sequence is exact and natural*

$$V(G,A) \xrightarrow{q_*} V(Q,A) \xrightarrow{\delta} N_{ab} \otimes_Q A \xrightarrow{i_*} G_{ab} \otimes A \xrightarrow{q_*} Q_{ab} \otimes A \to 0. \quad (1.2)$$

Note that N_{ab} is a Q-module by conjugation. Note also that if $\mathfrak{V}=\mathfrak{Gr}$ (1.2) is the well-known 5-term sequence in homology (see [7, Theorem VI, 8.1]).

An immediate consequence of (1.2) is that if $\mathfrak{V}=\mathfrak{Ab}$ and if Q is any abelian group, then $V(Q,A)=\mathrm{Tor}(Q,A)$.

Let G_j, $j=1,2$ be two groups in $\mathfrak{V}$. Denote by $G_1 *_V G_2$ the *varietal product* of G_1 and G_2, i.e. $G_1 *_V G_2 = G_1 * G_2 / V(G_1 * G_2)$, where for any group H the symbol $V(H)$ denotes the verbal subgroup of H corresponding to $\mathfrak{V}$. Note that the varietal product together with the obvious injections i_j: $G_j \to G_1 *_V G_2$, $j=1,2$ is the coproduct in the category $\mathfrak{V}$.

Proposition 1.2. $V(G_1 *_V G_2, A) \cong V(G_1,A) \oplus V(G_2,A)$.

Proof. Let q_i: $F_i \twoheadrightarrow G_i$, $i=1,2$ be $\mathfrak{V}$-free presentations. Consider the extensions

$$\begin{tikzcd} V(F_1 * F_2) \arrow[r, tail] \arrow[d, "q'"] & F_1 * F_2 \arrow[r, two heads] \arrow[d, "q"] & F_1 *_V F_2 \arrow[d, "q_V"] \\ V(G_1 * G_2) \arrow[r, tail] & G_1 * G_2 \arrow[r, two heads] & G_1 *_V G_2 \end{tikzcd}$$

and the map uniquely determined by q_i. For simplicity set $F=F_1 * F_2$, $F_V = F_1 *_V F_2$, $G=G_1 * G_2$, $G_V = G_1 *_V G_2$. The associated 5-term sequences in homology then read

$$\begin{tikzcd} H_2(F_1 * F_2, A) \arrow[r] \arrow[d, "q_*"] & H_2(F_1 *_V F_2, A) \arrow[r] \arrow[d, "(q_V)_*"] & V(F_1 * F_2)_{ab} \otimes_{F_V} A \arrow[r] \arrow[d, "(q')_*"] & 0 \\ H_2(G_1 * G_2, A) \arrow[r] & H_2(G_1 *_V G_2, A) \arrow[r] & V(G_1 * G_2)_{ab} \otimes_{G_V} A \arrow[r] & 0. \end{tikzcd}$$

We have zeros at the right hand end since clearly $F_{ab}=(F_V)_{ab}$, $G_{ab}=(G_V)_{ab}$. Now $F_V \twoheadrightarrow G_V$ is a $\mathfrak{V}$-free presentation, hence by definition $\operatorname{coker}(q_V)_* = V(G,A)$.

By [7, Theorem VI.14.2] we have

$$H_2(F,A)=H_2(F_1,A)\oplus H_2(F_2,A),$$
$$H_2(G,A)=H_2(G_1,A)\oplus H_2(G_2,A).$$

Hence $\operatorname{coker} q_* = V(G_1,A)\oplus V(G_2,A)$. It follows that the map α induced by the injections

$$\alpha\colon\ V(G_1,A)\oplus V(G_2,A)\to V(G,A) \tag{1.3}$$

is surjective, since trivially $(q')_*$ is surjective. Finally the injection $i_j\colon G_j \to G$ plainly has a left inverse $l_j\colon G\to G_j$. Thus since $V(-,A)$ is a functor, α is also injective, and the assertion is proved.

The next proposition generalizes a result of [10].

Proposition 1.3. *Let $\mathfrak{V}\subseteq\mathfrak{W}$, let G be in $\mathfrak{W}$ and let A be in $\mathfrak{Ab}$. Then there is a natural transformation of functors*

$$t(G,A)\colon\ W(G,A)\to V(G/VG,A).$$

Moreover $t(G,A)$ is surjective for all G in $\mathfrak{W}$ and all A in $\mathfrak{Ab}\cap\mathfrak{V}$.

Proof. Consider an absolutely free presentation $F \twoheadrightarrow G$; then $F/WF \twoheadrightarrow G$ is a $\mathfrak{W}$-free presentation of G, and $F/VF \twoheadrightarrow G$ is a $\mathfrak{V}$-free presentation of G/VG. We then consider the diagram

$$\begin{tikzcd}
H_2(F/WF,A) \arrow[r] \arrow[d] & H_2(G,A) \arrow[r] \arrow[d] & W(G,A) \arrow[r] \arrow[d, dotted, "t(G,A)"] & 0 \\
H_2(F/VF,A) \arrow[r] \arrow[d] & H_2(G/VG,A) \arrow[r] \arrow[d] & V(G/VG,A) \arrow[r] & 0 \\
(VF/WF)_{ab}\otimes_F A \arrow[r, "\alpha"] \arrow[d] & (VG)_{ab}\otimes_G A \arrow[d] & & \\
(F/WF)_{ab}\otimes A \arrow[d, "\beta"] & G_{ab}\otimes A \arrow[d, "\gamma"] & & \\
(F/VF)_{ab}\otimes A \arrow[d] & (G/VG)_{ab}\otimes A \arrow[d] & & \\
0 & 0 & &
\end{tikzcd} \tag{1.4}$$

Plainly the so defined $t(G,A)$ is a natural transformation. If A is in $\mathfrak{Ab}\cap\mathfrak{V}$, then β as well as γ are isomorphisms. Thus, since α is surjective, $t(G,A)$ is surjective also.

Corollary 1.4 (see [10]). *Let $\mathfrak{V}\subseteq\mathfrak{W}$, let G be in $\mathfrak{V}$ and let A be in $\mathfrak{Ab}\cap\mathfrak{V}$. Then if $F/VF \twoheadrightarrow G$ is a $\mathfrak{V}$-free presentation, there is an exact sequence*

$$W(F/VF,A)\to W(G,A)\xrightarrow{t(G,A)} V(G,A)\to 0. \tag{1.5}$$

The sequence (1.5) may be deduced from the sequence (2.1) of [10]. However we would like to emphasize that the proof given below is completely elementary, in particular it does not use spectral sequences.

Proof. This can be read off from (1.4). For, under the given hypotheses, β is an isomorphism. Thus $(VF/WF)_{ab}\otimes_F A = W(F/VF,A)$ by definition, and since $VG=\{e\}$, the sequence follows from (1.4) by applying the ker-coker lemma.

Proposition 1.5 ("Universal coefficient theorem"). *Let $\mathfrak{V}$ be of exponent zero, i.e. $\mathfrak{Ab}\subseteq\mathfrak{V}$. Then there is a split exact sequence*

$$V(G,\mathbb{Z})\otimes A \rightarrowtail V(G,A)\twoheadrightarrow \mathrm{Tor}(G_{ab},A). \tag{1.6}$$

Proof. This is immediate from the proof of [15; Lemma 3.3].

We close this section with a short word on cohomology. Of course we may define the first varietal cohomology group of G with coefficients A, $\tilde{V}(G,A)$ by proceeding dually to (1.1); thus

$$\tilde{V}(G,A)=\ker\big(q^*\colon\ H^2(G,A)\to H^2(F,A)\big). \tag{1.7}$$

By going through the arguments analogous to those establishing [7; Theorem VI, 10.3] it is then not hard to prove that if A is in $\mathfrak{Ab}\cap\mathfrak{V}$, then $\tilde{V}(G,A)$ classifies the extensions $A\rightarrowtail H\twoheadrightarrow G$ contained in $\mathfrak{V}$.

As easily as Proposition 1.5 one establishes the "universal coefficient theorem" for cohomology.

Proposition 1.6. *Let $\mathfrak{Ab}\subseteq\mathfrak{V}$. Then there is a split exact sequence*

$$\mathrm{Ext}(G_{ab},A)\rightarrowtail \tilde{V}(V,A)\twoheadrightarrow \mathrm{Hom}\big(V(G,\mathbb{Z}),A\big). \tag{1.8}$$

2. Pseudo Stem Extensions and Pseudo Stem Covers in $\mathfrak{V}$

In [6] a central extension of groups $N\rightarrowtail G\twoheadrightarrow Q$ is called a stem extension if and only if $N\subseteq G'$.

Proposition 2.1. *For an extension $N\overset{i}{\rightarrowtail} G\overset{q}{\twoheadrightarrow} Q$ (not necessarily central) the following statements are equivalent*

(i) *$N\subseteq G'$;*

(ii) *$i_*\colon N\to G_{ab}$ is the zeromap;*

(iii) *$q_*\colon G_{ab}\xrightarrow{\sim} Q_{ab}$;*

(iv) $\delta\colon V(Q,\mathbb{Z}) \xrightarrow{\sim} N/[G,N]$ *for all varieties* $\mathfrak{V}$ *containing* N, G, Q;

(v) *The associated central extension* $N/[G,N] \rightarrowtail G/[G,N] \twoheadrightarrow Q$ *is a stem extension.*

Proof. The equivalence of (i), (ii), (iii) and (v) is clear. It remains to prove (iii) $\Leftrightarrow$ (iv). For a variety $\mathfrak{V}$ containing N, G, Q Theorem 1.1 yields the exact sequence

$$V(G,\mathbb{Z}) \xrightarrow{q_*} V(Q,\mathbb{Z}) \xrightarrow{\delta} N/[G,N] \xrightarrow{i_*} G_{ab} \xrightarrow{q_*} Q_{ab} \to 0 \tag{2.1}$$

where $[G,N]$ denotes the (normal) subgroup generated by all $g\,n\,g^{-1}n^{-1}$ with $g\in G$ and $n\in N$, so that $N/[G,N] = N_{ab}\otimes_Q \mathbb{Z}$. It is then clear that (iii) is equivalent to (iv).

Definition. An extension satisfying one of the equivalent conditions of Proposition 2.1 is called a *pseudo stem extension*.

In order to generalize the concept of stem covers, we first state the following result which is immediate from the sequence (2.1).

Proposition 2.2. *For an extension* $N \overset{i}{\rightarrowtail} G \overset{q}{\twoheadrightarrow} Q$ *in* $\mathfrak{V}$ *the following statements are equivalent.*

(i') $\delta\colon V(Q,\mathbb{Z}) \xrightarrow{\sim} N/[G,N]$;

(ii') $q_*\colon G_{ab} \xrightarrow{\sim} Q_{ab}$, *and* $q_*\colon V(G,\mathbb{Z}) \to V(Q,\mathbb{Z})$ *is the zeromap.*

Definition. An extension in $\mathfrak{V}$ satisfying one of the equivalent conditions of Proposition 2.2 is called a *pseudo stem cover* in $\mathfrak{V}$. If in addition the extension is central, it is called a *stem cover* in $\mathfrak{V}$. (Sometimes we also say that G is a (pseudo) stem cover in $\mathfrak{V}$ of Q.)

We note that it is obvious that a stem cover in $\mathfrak{Gr}$ is just a stem cover in the sense of [6, p. 213].

Proposition 2.3. *Let* $N \rightarrowtail G \twoheadrightarrow Q$ *be an extension in* $\mathfrak{V}$. *Then it is a pseudo stem extension (pseudo stem cover) in* $\mathfrak{V}$ *if and only if the associated central extension*

$$N/[G,N] \rightarrowtail G/[G,N] \twoheadrightarrow Q$$

is a stem extension (stem cover) in $\mathfrak{V}$.

Proof. The naturality of the sequence (2.1) yields the commutative diagram

$$\begin{tikzcd}
V(G,\mathbb{Z}) \arrow[r] \arrow[d] & V(Q,Z) \arrow[r] \arrow[d, equal] & N/[G,N] \arrow[r] \arrow[d, equal] & G_{ab} \arrow[r] \arrow[d, equal] & Q_{ab} \arrow[r] \arrow[d, equal] & 0 \\
V(G/[G,N],\mathbb{Z}) \arrow[r] & V(Q,Z) \arrow[r] & N/[G,N] \arrow[r] & G_{ab} \arrow[r] & Q_{ab} \arrow[r] & 0
\end{tikzcd}$$

whence the assertion is obvious.

We continue with two examples:

a) Let G be a knot-group. Then $G' \rightarrowtail G \twoheadrightarrow C$ where C is infinite cyclic is obviously a pseudo stem cover in $\mathfrak{Gr}$. Since $H_2(C, \mathbb{Z}) = 0$, there are no non-trivial stem extensions of C, hence we immediately obtain $[G, G'] = G'$ (which is of course well-known).

b) A group G in $\mathfrak{V}$ with $V(G, \mathbb{Z}) = 0$ is a pseudo stem cover in $\mathfrak{V}$ of all its quotients by normal subgroups $N \subseteq G'$.

We will consider other examples in the Sections 3 and 4.

Proposition 2.4. Let $\mathfrak{Ab} \subseteq \mathfrak{V} \subseteq \mathfrak{W}$. *Let* $N \rightarrowtail G \xrightarrow{q} Q$ *be a pseudo stem cover in* $\mathfrak{W}$. *Then* $N/VG \cap N \rightarrowtail G/VG \xrightarrow{q} Q/VQ$ *is a pseudo stem cover in* $\mathfrak{V}$.

Proof. We have to show that $V(G/VG, \mathbb{Z}) \to V(Q/VQ, \mathbb{Z})$ is the zero-map. To that end we consider the commutative square.

$$\begin{array}{ccc} W(G, \mathbb{Z}) & \xrightarrow{q_*} & W(Q, \mathbb{Z}) \\ {\scriptstyle t(G,\mathbb{Z})}\downarrow & & \downarrow \\ V(G/VG, \mathbb{Z}) & \xrightarrow{q_*} & V(Q/VQ, \mathbb{Z}) \end{array}$$

Since $\mathfrak{Ab} \cap \mathfrak{V}$, the group $\mathbb{Z}$ is in $\mathfrak{V}$. Hence by Proposition 1.3 the map $t(G, \mathbb{Z})$ is surjective. Thus if the upper q_* is the zero map the lower q_* is.

We note that the "converse" of Proposition 2.3 is not true: We give an example of an extension which is a stem cover in $\mathfrak{V}$, but not a stem cover in $\mathfrak{W}(\supseteq \mathfrak{V})$. Let $\mathfrak{V} = \mathfrak{Ab}$ and let $\mathfrak{W} = \mathfrak{Gr}$. Since for any abelian group A, the group $\mathfrak{V}(A, \mathbb{Z}) = \mathrm{Tor}\,(A, \mathbb{Z}) = 0$, the extension $0 \rightarrowtail A \twoheadrightarrow A$ is a stem cover in $\mathfrak{V}$. But if $H_2(A, \mathbb{Z}) \neq 0$, it obviously is not a stem cover in $\mathfrak{W}$.

We infer from Proposition 1.3 that given a stem extension $N \rightarrowtail G \twoheadrightarrow Q$ in $\mathfrak{V}(\supseteq \mathfrak{Ab})$ then there exists a stem cover in $\mathfrak{V}$ which maps onto $N \rightarrowtail G \twoheadrightarrow Q$. For we choose first a stem cover $\tilde{N} \rightarrowtail \tilde{G} \twoheadrightarrow Q$ in $\mathfrak{Gr}$ mapping onto $N \rightarrowtail G \twoheadrightarrow Q$. (Such a stem cover exists by [6, p. 213].) Then

$$\tilde{N}/V\tilde{G} \rightarrowtail \tilde{G}/V\tilde{G} \to Q$$

in a stem cover in $\mathfrak{V}$ mapping onto $N \rightarrowtail G \twoheadrightarrow Q$. (Of course $V\tilde{G} \cap \tilde{N} = V\tilde{G}$!)

Proposition 2.5. *Let* $M \rightarrowtail H \twoheadrightarrow Q$ *be a pseudo stem extension in* $\mathfrak{V}$ *and let* $N \rightarrowtail G \twoheadrightarrow Q$ *be a pseudo stem cover in* $\mathfrak{V}$. *If* $f\colon H \to G$ *is a map of extensions, then* $M \rightarrowtail H \twoheadrightarrow Q$ *also is a pseudo stem cover in* $\mathfrak{V}$. *Moreover* f *induces isomorphisms* $f_2\colon H_{\mathrm{ab}} \xrightarrow{\sim} G_{\mathrm{ab}}$, $f_3\colon H/H_3 \xrightarrow{\sim} G/G_3$ *and epimorphisms* $f_n\colon H/H_n \twoheadrightarrow G/G_n$ *for all* $n \geqq 4$.

Here G_n, H_n denote as usual the terms of the lower central series.

Proof. Clearly $V(H,\mathbb{Z}) \xrightarrow{f_*} V(G,\mathbb{Z}) \to V(Q,\mathbb{Z})$ is the zeromap. Hence $M \rightarrowtail H \twoheadrightarrow Q$ is a pseudo stem cover in $\mathfrak{V}$. Plainly also, f_2 is an isomorphism. It is then clear that f_n is surjective for all $n \geqq 2$. It remains to prove that f_3 is injective. To that end consider the diagram

$$\begin{tikzcd} V(H,\mathbb{Z}) \arrow[r] \arrow[d, "f_*"] & V(H_{ab},\mathbb{Z}) \arrow[r] \arrow[d, "(f_2)_*", "\wr"'] & H_2/H_3 \arrow[r] \arrow[d, "\psi"] & 0 \\ V(G,\mathbb{Z}) \arrow[r] & V(G_{ab},\mathbb{Z}) \arrow[r] & G_2/G_3 \arrow[r] & 0. \end{tikzcd} \tag{2.2}$$

Since $\tau\colon V(G,\mathbb{Z}) \to V(G_{ab},\mathbb{Z})$ factors as

$$V(G,\mathbb{Z}) \xrightarrow{0} V(Q,\mathbb{Z}) \to V(Q_{ab},\mathbb{Z}) \xrightarrow{(f_2^{-1})_*} V(G_{ab},\mathbb{Z}).$$

it follows that $\tau = 0$, hence ψ in (2.2) is an isomorphism. It then follows that f_3 in the diagram

$$\begin{tikzcd} H_2/H_3 \arrow[r, tail] \arrow[d, "\psi"', "\wr"] & H/H_3 \arrow[r, two heads] \arrow[d, "f_3"] & H_{ab} \arrow[d, "\wr", "f_2"'] \\ G_2/G_3 \arrow[r, tail] & G/G_3 \arrow[r, two heads] & G_{ab} \end{tikzcd}$$

is an isomorphism, also.

We show that f_4 may fail to be an isomorphism. Let $\mathfrak{V} = \mathfrak{Gr}$. Take a non commutative absolutly free group F. Then F and $G = F/F_3$ are obviously pseudo stem covers of F_{ab}. But $F/F_4 \neq G/G_4 (= F/F_3)$.

If we call two pseudo stem covers in $\mathfrak{V}$ *of the same branch* if their associated central extensions are equivalent, then we may infer from Proposition 2.5 the following result.

Corollary 2.6. *If $M \rightarrowtail H \twoheadrightarrow Q$ and $N \rightarrowtail G \twoheadrightarrow Q$ are of the same branch, then $H/H_3 \cong G/G_3$.*

If $\mathfrak{V}$ is of exponent zero, it follows from a trivial extension of [5, (4.18) ff.] using Proposition 1.6 that there exists only one stem cover in $\mathfrak{V}$ if and only if $\operatorname{Ext}(Q_{ab}, V(Q,\mathbb{Z})) = 0$. Hence we obtain:

Corollary 2.7. *If $\operatorname{Ext}(Q_{ab}, V(Q,\mathbb{Z})) = 0$ and if $M \rightarrowtail H \twoheadrightarrow Q$ and $N \rightarrowtail G \twoheadrightarrow Q$ are pseudo stem covers in $\mathfrak{V}$, then $H/H_3 \cong G/G_3$.*

3. A Remark to a Paper of Passi and Vermani

In [11] Passi and Vermani investigate under what conditions the inflation homorphism $q^*\colon H^2(G/G_n, \mathbb{Q}/\mathbb{Z}) \to H^2(G, \mathbb{Q}/\mathbb{Z})$ is the zero homomorphism. Here G is a nilpotent group of class n, so that G_n is the last non trivial term of the lower central series. We shall show that this situation is most easily described in terms of (pseudo) stem covers. At the same time we will show that the problem of Passi and Vermani has an immediate generalization to varieties of exponent zero. We first prove the following Lemma.

Lemma 3.1. *Let* $N \overset{i}{\rightarrowtail} G \overset{q}{\twoheadrightarrow} Q$ *be an extension in a variety* $\mathfrak{V}$ *of exponent zero. Then* $q^*\colon \tilde{V}(Q, \mathbb{Q}/\mathbb{Z}) \to \tilde{V}(G, \mathbb{Q}/\mathbb{Z})$ *is the zeromap if and only if* $q_*\colon V(G, \mathbb{Z}) \to V(Q, \mathbb{Z})$ *is the zeromap.*

Proof. By Proposition 1.6 we have $\tilde{V}(G, \mathbb{Q}/\mathbb{Z}) = \mathrm{Hom}(V(G, \mathbb{Z}), \mathbb{Q}/\mathbb{Z})$ and similarly for Q. Since $\mathbb{Q}/\mathbb{Z}$ is a cogenerator in $\mathfrak{Ab}$ the assertion follows.

We now immediatly have:

Proposition 3.2. *Let* $N \overset{i}{\rightarrowtail} G \overset{q}{\twoheadrightarrow} Q$ *be a pseudo stem extension in the variety* $\mathfrak{V}$ *of exponent zero. Then* $q^*\colon \tilde{V}(Q, \mathbb{Q}/\mathbb{Z}) \to \tilde{V}(G, \mathbb{Q}/\mathbb{Z})$ *is the zero map if and only if* $N \rightarrowtail G \twoheadrightarrow Q$ *is a pseudo stem cover in* $\mathfrak{V}$.

This clearly covers the situation considered in [11], for if $N = G_n$, $n \geq 2$, then $G_n \rightarrowtail G \twoheadrightarrow G/G_n$ is a pseudo stem extension. Actually in the case considered in [11], G_n is central, so that we even have a stem extension.

It is now trivial to infer the various assertions about the particular form of the presentations of G if G has the property that

$$q^*\colon H^2(G/G_n, \mathbb{Q}/\mathbb{Z}) \to H^2(G, \mathbb{Q}/\mathbb{Z})$$

is the zeromap. For, if $N \rightarrowtail G \twoheadrightarrow Q$ is an extension in the variety $\mathfrak{V}$ of exponent zero, we consider a $\mathfrak{V}$-free presentation of $N \rightarrowtail G \twoheadrightarrow Q$, i.e. a diagram

$$\begin{tikzcd}
 & & N \arrow[d, tail] \\
R \arrow[r, tail] \arrow[d, tail] & F \arrow[r, two heads] \arrow[d, equal] & G \arrow[d, two heads] \\
S \arrow[r, tail] \arrow[d, two heads] & F \arrow[r, two heads] & Q \\
N & &
\end{tikzcd}$$

with F $\mathfrak{V}$-free. It is then clear from [15, Corollary 3.2.1] that

$$q_*\colon V(G, \mathbb{Z}) \to V(Q, \mathbb{Z})$$

is the zeromap if and only if

$$R \cap F' \subseteq [F, S].$$

Thus coming back to the extension $G_n \rightarrowtail G \twoheadrightarrow G/G_n$ of Passi and Vermani, $q_*: H_2(G, \mathbb{Z}) \to H_2(G/G_n, \mathbb{Z})$ is the zeromap if and only if

$$R \cap F' = [F, F_n R] = F_n[F, R]$$

since trivially $[F, F_n R] \subseteq R \cap F'$ if G is nilpotent of class n.

4. Parafree Groups

We first introduce some notation: Let p be a prime or $p=0$. If U, V are subgroups of G we shall use the symbol $U \#_p V$ to denote the subgroup of G generated by all $u v u^{-1} v^{-1} w^p$ with $u \in U$, v, $w \in V$. Using this notation we may define a series of normal subgroups of G by setting

$$G_1^{(p)} = G, \qquad G_{n+1}^{(p)} = G \#_p G_n^{(p)}, \qquad n \geqq 1; \qquad G_\omega^{(p)} = \bigcap_n G_n^{(p)}. \tag{4.1}$$

For $p=0$ this plainly is the lower central series, thus $G_n^{(0)} = G_n$. For p a prime it is the most rapidly descending central series whose successive quotients are vector spaces over $\mathbb{Z}_p$, the field of p elements.

In what follows the sequence (1.2) will often be used in the special case where $A = \mathbb{Z}_p$ ($\mathbb{Z}_0 = \mathbb{Z}$). This sequence then is

$$V(G, \mathbb{Z}_p) \xrightarrow{q_*} V(Q, \mathbb{Z}_p) \to N/G \#_p N \to G_{ab} \otimes \mathbb{Z}_p \xrightarrow{q_*} Q_{ab} \otimes \mathbb{Z}_p \to 0. \tag{4.2}$$

Next we recall the definition of $\mathfrak{V}$-parafree groups [3].

Two groups H, G are said to have the *same lower central sequence* if there exists isomorphisms $\vartheta_n: H/H_n \to G/G_n$, $n \geqq 1$ such that ϑ_n is induced by ϑ_{n+1}.

A group G is called $\mathfrak{V}$-*parafree* (of rank n) if the following is satisfied

(i) G is in $\mathfrak{V}$;

(ii) G is residually nilpotent;

(iii) G has the same lower central sequence as some $\mathfrak{V}$-free group (of rank n).

Baumslag to whom this definition is due has shown [3] that there exist many non $\mathfrak{V}$-free $\mathfrak{V}$-parafree groups. In order to tie up parafree groups with pseudo stem covers, we prove the following proposition.

Proposition 4.1. *A $\mathfrak{V}$-parafree group G is a pseudo stem cover in $\mathfrak{V}$ of its quotients by the terms G_n, $n \geqq 2$ of the lower central series.*

Proof. Given $n \geqq 2$ there exists $f: F \to G$ with F $\mathfrak{V}$-free inducing $\vartheta_{n+1}: F/F_{n+1} \to G/G_{n+1}$, and hence also $\vartheta_m: F/F_m \to G/G_m$, $1 \leqq m \leqq n$. The diagram

$$\begin{tikzcd}
F_n/F_{n+1} \arrow[r, tail] \arrow[d, "\psi_n"] & F/F_{n+1} \arrow[r, two heads] \arrow[d, "\vartheta_{n+1}"] & F/F_n \arrow[d, "\vartheta_n"] \\
G_n/G_{n+1} \arrow[r, tail] & G/G_{n+1} \arrow[r, two heads] & G/G_n
\end{tikzcd}$$

then shows that ψ_n is an isomorphism. Considering the diagram of exact sequences (4.2)

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & V(F/F_n, \mathbb{Z}) \arrow[r] \arrow[d, "(\vartheta_n)_*"] & F_n/F_{n+1} \arrow[r] \arrow[d, "\psi_n"] & F_{ab} \arrow[r] \arrow[d, "(\vartheta_2)_*"] & (F/F_n)_{ab} \arrow[r] \arrow[d, "(\vartheta_n)_*"] & 0 \\
V(G, \mathbb{Z}) \arrow[r] & V(G/G_n, \mathbb{Z}) \arrow[r] & G_n/G_{n+1} \arrow[r] & G_{ab} \arrow[r] & (G/G_n)_{ab} \arrow[r] & 0.
\end{tikzcd}$$

Since ϑ_n is an isomorphism, and $V(-, \mathbb{Z})$ is a functor $(\vartheta_n)_*$ is an isomorphism, also. The fact that ψ_n is an isomorphism yields that

$$q_*: V(G, \mathbb{Z}) \to V(G/G_n, \mathbb{Z})$$

is the zeromap. Finally $G_n \subseteq G_2$ for $n \geqq 2$, thus $G_n \rightarrowtail G \twoheadrightarrow G/G_n$ is a pseudo stem cover in $\mathfrak{V}$.

For many of our remaining results we shall need the following Lemma.

Lemma 4.2. *Let $f: H \to G$ be a homomorphism in $\mathfrak{V}$ with $f_*: H_{ab} \otimes \mathbb{Z}_p \xrightarrow{\sim} G_{ab} \otimes \mathbb{Z}_p$ for a fixed p. Suppose that $q_*: V(G, \mathbb{Z}_p) \to V(G/G_n^{(p)}, \mathbb{Z}_p)$ is the zeromap for all $n \geqq 1$. Then f induces isomorphisms*

$$f_n: H/H_n^{(p)} \xrightarrow{\sim} G/G_n^{(p)}$$

for all $n \geqq 1$ and a monomorphism $f_\omega: H/H_\omega^{(p)} \rightarrowtail G/G_\omega^{(p)}$.

Proof. The idea of the proof comes from [13, 14]; see also [15]. We proceed by induction on n. For $n = 1, 2$ the assertion is trivial or part of the hypothesis (notice that $G_{ab} \otimes \mathbb{Z}_p = G/G_2^{(p)}$). So suppose $f_n: H/H_n^{(p)} \xrightarrow{\sim} G/G_n^{(p)}$. Then we consider the diagram of exact sequences (4.2)

$$\begin{tikzcd}
V(H, \mathbb{Z}_p) \arrow[r] \arrow[d, "f_*"] & V(H/H_n^{(p)}, \mathbb{Z}_p) \arrow[r] \arrow[d, "(f_n)_*"] & H_n^{(p)}/H_{n+1}^{(p)} \arrow[r] \arrow[d, "\psi_n"] & 0 \\
V(G, \mathbb{Z}_p) \arrow[r, "q_*"] & V(G/G_n^{(p)}, \mathbb{Z}_p) \arrow[r] & G_n^{(p)}/G_{n+1}^{(p)} \arrow[r] & 0.
\end{tikzcd}$$

Since $(f_n)_*$ is an isomorphism and q_* is the zeromap it follows that ψ_n is an isomorphism. The diagram

$$\begin{tikzcd}
H_n^{(p)}/H_{n+1}^{(p)} \arrow[r, tail] \arrow[d, "\psi_n"] & H/H_{n+1}^{(p)} \arrow[r, two heads] \arrow[d, "f_{n+1}"] & H/H_n^{(p)} \arrow[d, "f_n"] \\
G_n^{(p)}/G_{n+1}^{(p)} \arrow[r, tail] & G/G_{n+1}^{(p)} \arrow[r, two heads] & G/G_n^{(p)}
\end{tikzcd}$$

then shows that, with f_n and ψ_n, the map f_{n+1} is an isomorphism. The remaining part of the assertion is then trivial.

Proposition 4.3. *Let $\mathfrak{B}$ be a variety of exponent zero. Let G be a pseudo stem cover in $\mathfrak{B}$ of its quotients G/G_n, $n \geqq 2$. If G_{ab} is free abelian, then there exists a free group F and a homomorphism $f: F \to G$ such that f induces isomorphisms*

$$f_n: F/F_n^{(p)} \xrightarrow{\sim} G/G_n^{(p)}, \qquad n \geqq 0, \ p \text{ prime or } p=0.$$

Proof. First let $p=0$. Choose elements $x_i \in G$, $i \in I$ such that their images in G_{ab} form a basis. Let F be the V-free group on (y_i), $i \in I$ and define $f: F \to G$ by $f(y_i) = x_i$. By Lemma 4.2 f induces isomorphisms $f_n: F/F_n \xrightarrow{\sim} G/G_n$, $n \geqq 1$. The same argument works in case p is a prime, provided we can show that $q'_*: V(G, \mathbb{Z}_p) \to V(G/G_n^{(p)}, \mathbb{Z}_p)$ is the zeromap for all $n \geqq 2$. But by definition $G_n^{(p)} \supseteq G_n$ so that q'_* factors as

$$V(G, \mathbb{Z}_p) \xrightarrow{q_*} V(G/G_n, \mathbb{Z}_p) \to V(G/G_n^{(p)}, \mathbb{Z}_p).$$

It thus is enough to show that q_* is the zeromap. Now since G_{ab} is free abelian we have $V(G, \mathbb{Z}_p) = V(G, \mathbb{Z}) \otimes \mathbb{Z}_p$, $V(G/G_n, \mathbb{Z}_p) = V(G/G_n, \mathbb{Z}) \otimes \mathbb{Z}_p$ by Proposition 1.5. Hence $q_* = 0$, since $G_n \rightarrowtail G \twoheadrightarrow G/G_n$ is a stem cover in $\mathfrak{B}$.

We note that the assertion for $p=0$ of Proposition 4.3 remains true for an arbitrary variety $\mathfrak{B}$ if we suppose G_{ab} free in $\mathfrak{B} \cap \mathfrak{Ab}$.

Corollary 4.4. *The group G is $\mathfrak{B}$-parafree if and only if the following is satisfied*

(i) *G is in $\mathfrak{B}$;*

(ii) *G is residually nilpotent;*

(iii) *G_{ab} is free in $\mathfrak{B} \cap \mathfrak{Ab}$ and G is a pseudo stem cover in $\mathfrak{B}$ of its quotients G/G_n, $n \geqq 2$.*

We remark that it is apparent from Corollary 4.4 and Proposition 4.3 that the maps ϑ_n in Baumslag's definition of parafree groups may all be obtained from *one* homomorphism $f: F \to G$.

Corollary 4.5. *Let G be a residually nilpotent group in $\mathfrak{B}$ with G_{ab} free in $\mathfrak{B} \cap \mathfrak{Ab}$ and $V(G, \mathbb{Z}) = 0$ then G is $\mathfrak{B}$-parafree.*

Corollary 4.6. *Let $\mathfrak{B}$ be a variety of exponent zero in which the $\mathfrak{B}$-free groups are residually nilpotent. If G is $\mathfrak{B}$-parafree, then G contains a $\mathfrak{B}$-free subgroup F such that the embedding $i: F \subseteq G$ induces isomorphisms.*

$$i_n: F/F_n^{(p)} \xrightarrow{\sim} G/G_n^{(p)}, \qquad n \geqq 0, \quad p \text{ prime or } p=0.$$

Proof. From Proposition 4.3 and Lemma 4.2 it follows that there exists $f: F \to G$ which induces isomorphisms f_n and a monomorphism $f_\omega: F/F_\omega \rightarrowtail G/G_\omega$. Since F is residually nilpotent, $F_\omega = \{e\}$, whence f_ω factors as $F \xrightarrow{f} G \to G/G_\omega$. Thus f is monomorphic; set $i=f$.

Corollary 4.7. *Let $\mathfrak{V}$ be a variety of exponent zero in which the free groups are residually finite p-groups for infinitely many primes p. Let G be $\mathfrak{V}$ parafree and let (x_i), $i \in I$, $x_i \in G$ be a set of elements, such that their images in G_{ab} are linearly independent. Then (x_i), $i \in I$ generate a $\mathfrak{V}$-free subgroup F of G.*

Proof. We first note that it is enough to prove the assertion for all finite subsets of (x_i), $i \in I$. Thus we may suppose that (x_i), $i \in I$ is finite. We then can find a prime p such that the images of x_i in $G_{\mathrm{ab}} \otimes \mathbb{Z}_p$ are linearly independent and that the free groups in $\mathfrak{V}$ are residually finite p-groups. We then may enlarge the set (x_i), $i \in I$ to a set (x_i), $i \in J$, $I \subseteq J$ such that the images in $G_{\mathrm{ab}} \otimes \mathbb{Z}_p$ form a $\mathbb{Z}_p$-basis of $G_{\mathrm{ab}} \otimes \mathbb{Z}_p$. We shall prove that (x_i), $i \in J$ generates a free subgroup. To that end consider the $\mathfrak{V}$-free group F on (y_i), $i \in J$ and the map $f\colon F \to G$ defined by $f(y_i) = x_i$. By Lemma 4.2 f is a monomorphism, whence our assertion.

Note that in Corollaries 4.6 and 4.7 we only used that G was a pseudo stem cover in $\mathfrak{V}$ of G/G_n, $n \geqq 2$ and that G_{ab} was free abelian: we did not make use of the fact that G was residually nilpotent.

Corollary 4.7 generalizes [4, Theorem 4.1]. We note two other immediate consequences of our Lemma 4.2.

Proposition 4.8. *Let $f\colon H \to G$ be a homomorphism in $\mathfrak{V}$ with G $\mathfrak{V}$-parafree and H residually nilpotent. Suppose $f_2\colon H_{\mathrm{ab}} \xrightarrow{\sim} G_{\mathrm{ab}}$, then H is $\mathfrak{V}$-parafree of the same rank as G and f is a monomorphism.*

Corollary 4.9. *Let $f\colon H \to G$ be an epimorphism of $\mathfrak{V}$-parafree groups of the same finite rank, then f is an isomorphism.*

Proof. We only have to prove that $f_2\colon H_{\mathrm{ab}} \xrightarrow{\sim} G_{\mathrm{ab}}$. But f_2 is trivially surjective: if H and G have the same finite rank, both H_{ab} and G_{ab} are finitely generated and free in $\mathfrak{V} \cap \mathfrak{Ab}$ of the same rank, hence f_2 is an isomorphism.

Note that Corollary 4.9 generalizes [4, Theorem 1.1].

Proposition 4.10. *Let G be a pseudo stem cover in $\mathfrak{V}$ of G/G_n, $n \geqq 2$ and let G_{ab} be free in $\mathfrak{Ab} \cap \mathfrak{V}$. Then G/G_ω is $\mathfrak{V}$-parafree.*

Proof. The assertion easily follows from the diagram

$$\begin{tikzcd}
V(G,\mathbb{Z}) \arrow[r, "q_*"] \arrow[d] & V(G/G_n,\mathbb{Z}) \arrow[r] \arrow[d, "\wr"] & G_n/G_{n+1} \arrow[r] \arrow[d, "\wr"] & 0 \\
V(G/G_\omega,\mathbb{Z}) \arrow[r, "q_*"] & V(G/G_n,\mathbb{Z}) \arrow[r] & G_n/G_{n+1} \arrow[r] & 0
\end{tikzcd}$$

If the upper q_* is the zeromap, the lower q_* is.

We conclude with two Propositions of a somewhat different kind.

Proposition 4.11. *Let $\mathfrak{V}$ be a variety of exponent zero. A $\mathfrak{V}$-parafree group G of rank $m \geqq 1$ has a $\mathfrak{V}$-parafree quotient of rank $m-1$.*

Proof. Let $a \in G$ such that the image of a in G_{ab} is a basis element. Consider the normal subgroup N of G generated by a. The extension $N \overset{i}{\rightarrowtail} G \overset{t}{\twoheadrightarrow} Q$ yields the sequence

$$V(G,\mathbb{Z}) \xrightarrow{t_*} V(Q,\mathbb{Z}) \rightarrow N/[G,N] \xrightarrow{i_*} G_{\mathrm{ab}} \xrightarrow{t_*} (G/N)_{\mathrm{ab}} \rightarrow 0.$$

Since N is generated by one element, $N/[G,N]$ is cyclic. Since rank $G_{\mathrm{ab}} = m$, rank $Q_{\mathrm{ab}} = m-1$ the map i_* is non trivial. Since G_{ab} is free abelian i_* is a monomorphism. It follows that t_*: $V(G,\mathbb{Z}) \rightarrow V(Q,\mathbb{Z})$ is epimorphic. Now consider the square ($n \geqq 2$)

$$\begin{array}{ccc} V(G,\mathbb{Z}) & \xrightarrow{t_*} & V(Q,\mathbb{Z}) \\ {\scriptstyle q_*}\downarrow & & \downarrow{\scriptstyle q'_*} \\ V(G/G_n,\mathbb{Z}) & \longrightarrow & V(G/G_n N,\mathbb{Z}). \end{array}$$

q_* is the zeromap, hence the diagonal is the zeromap. Since t_* is epimorphic, q'_* has to be the zeromap, also. But of course $Q/Q_n = G/G_n N$.

We note that [4, Theorem 5.2 and Theorem 5.3] show that for $\mathfrak{V} = \mathfrak{Gr}$ and $\mathfrak{V}$ the variety of metabelian groups Propostion 4.11 can be strengthened considerably. However we do not see how Baumslag's proofs of these theorems can be simplified using our homological machinery.

Proposition 4.12. *Every (non-free) $\mathfrak{V}$-parafree group of rank m is a quotient of a (non free) $\mathfrak{V}$-parafree group of rank $m+1$.*

Proof. Let G be $\mathfrak{V}$-parafree of rank m. Consider $H = G *_V C$ where C is the $\mathfrak{V}$-free group on one generator. Clearly H_{ab} is free in $\mathfrak{Ab} \cap \mathfrak{V}$ of rank $m+1$. We next prove that q'_*: $V(H,\mathbb{Z}) \rightarrow V(H/H_n,\mathbb{Z})$, $n \geqq 2$ is the zeromap. Since $G/G_n \subseteq H/H_n$ we can consider the following square

$$\begin{array}{ccc} V(H,\mathbb{Z}) & \dashrightarrow^{q'_*} & V(H/H_n,\mathbb{Z}) \\ {\scriptstyle\alpha}\uparrow & & \uparrow \\ V(G,\mathbb{Z}) \oplus V(C,\mathbb{Z}) & & \\ {\scriptstyle\beta}\downarrow & & \Big| \\ V(G,\mathbb{Z}) & \xrightarrow{q_*} & V(G/G_n,\mathbb{Z}) \end{array} .$$

where α is the isomorphism of Proposition 1.2. The map β is an isomorphism since $V(C,\mathbb{Z}) = 0$. Hence the fact that q_* is the zeromap implies that q'_* is the zeromap. Finally since H_{ab} is free in $\mathfrak{Ab} \cap \mathfrak{B}$ Proposition 4.10 implies that H/H_ω is $\mathfrak{B}$-parafree. Thus G is a quotient of a $\mathfrak{B}$-parafree group H/H_ω of rank $m+1$.

References

1. André, M.: Méthode simpliciale en algèbre homologique et algèbre commutative, Lecture Notes in Mathematics, vol. 32. Berlin-Heidelberg-New York: Springer 1967.
2. Barr, M., Beck, J.: Homology and standard constructions, Lecture Notes in Mathematics, vol. 80. Berlin-Heidelberg-New York: Springer 1969.
3. Baumslag, G.: Groups with the same lower central sequence as a relatively free group. I. The groups. Trans. Amer. Math. Soc. **129**, 308–321 (1967).
4. Baumslag, G.: Groups with the same lower central sequence as a relatively free group. II. Properties. Trans. Amer. Math. Soc. **142**, 507–538 (1969).
5. Eckmann, B., Hilton, P., Stammbach, U.: On the homology theory of central group extensions I. The commutator map and stem extensions. Commentarii Math. Helvet. **47**, 102–122 (1972).
6. Gruenberg, K.W.: Cohomological topics in group theory. Lecture Notes in Mathematics, vol. 143. Berlin-Heidelberg-New York: Springer 1970.
7. Hilton, P., Stammbach, U.: A course in homological algebra. Graduate Texts in Mathematics, vol. 4. Berlin-Heidelberg-New York: Springer 1971.
8. Leedham-Green, C.: Homology in varieties of groups I. Trans. Amer. Math. Soc. **162**, 1–14 (1971).
9. Leedham-Green, C.: Homology in varieties of groups II. Trans. Amer. Math. Soc. **162**, 15–26 (1971).
10. Leedham-Green, C.: Homology in varieties of groups III. Trans. Amer. Math. Soc. **162**, 27–34 (1971).
11. Passi, I.B.S., Vermani, L.R.: The inflation homomorphism. J. London Math. Soc. (to appear).
12. Rinehart, G.S.: Satellites and cohomology. J. Algebra **12**, 295–329 (1969).
13. Stallings, J.: Homology and central series of groups. J. Algebra **2**, 170–181 (1965).
14. Stammbach, U.: Anwendungen der Homologietheorie der Gruppen auf Zentralreihen und auf Invarianten von Präsentierungen. Math. Z. **94**, 157–177 (1966).
15. Stammbach, U.: Homological methods in group varieties. Commentarii Math. Helvet. **45**, 287–298 (1970).
16. Ulmer, F.: Kan extensions, cotriples and André (co)homology. Lecture Notes in Mathematics, vol. 92. Berlin-Heidelberg-New York: Springer 1969.

Prof. U. Stammbach
Eidgenössische Technische Hochschule
CH-8006 Zürich, Switzerland

(Received April 25, 1972)

Math. Z. 128, 169–185 (1972)

Globale Ringe und Moduln

Klaus Langmann

Einleitung

Diese Arbeit ist die Fortsetzung der Artikel „Globale Moduln" (s. [10]) und „Konstruktionen globaler Moduln und Anwendungen" (s. [9]); die Numerierung dieser Sätze schließt deshalb auch an die Numerierung der Sätze dieser Arbeiten an. Wenn k ein vollständig bewerteter Körper, Z ein k-analytischer Raum und $G \subset Z$ eine beliebige Teilmenge ist, können wir Unterringe R des „vollen" Ringes $R_Z(G) := \Gamma(G, \mathcal{O}) \equiv \lim \Gamma(Z_i, \mathcal{O})$ (wobei sich der Limes über alle k-analytischen Räume Z_i mit $G \subset Z_i \subset Z$ erstreckt) betrachten. Ist speziell Z eine reindimensionale Mannigfaltigkeit, so ist $R_Z(G)$ nur noch von $n := \dim Z$ abhängig und wir schreiben statt $R_Z(G)$ einfach nur $R_n(G)$. Ein analytischer R-Modul ist ein Modul, zu dem eine in einer Umgebung von G kohärente Garbe $\mathcal{M}$ existiert mit $M \subset \Gamma(G, \mathcal{M})$. Ist $N \subset M$ ein R-Untermodul, so erfülle $N \subset M$ die Eigenschaft (I) bzw. (III) bezüglich $G \subset Z$, wenn es irgendwelche bzw. sogar endlich viele Punkte $z_i \in G$ gibt mit

$$N = \{m \in M \mid m \in N \mathcal{O}_{z_i}^Z \ \forall i\}.$$

Der Modul M erfüllte die Eigenschaft I bzw. III bezüglich $G \subset Z$, wenn alle endlichen Untermoduln $N \subset M$ die Eigenschaft (I) bzw. (III) erfüllen; M erfülle die Eigenschaft $\bar{\mathrm{I}}$, wenn sogar alle Untermoduln $N \subset M$ die Eigenschaft (I) erfüllen. Eine große Klasse von Moduln erfüllt diese Eigenschaften (s. [9]) und [10]); wir wollen hier erste Folgerungen daraus ziehen und insbesondere Anwendungen in der globalen Funktionentheorie geben. Bemerkenswert ist dabei, daß beim Beweis vieler Sätze, die sich nur auf die „üblichen" globalen Ringe $R_Z(G)$ beziehen, echt die Theorie der Unterringe $R \subset R_Z(G)$ gebraucht wurde (z.B. bei Folgerung 4.4.1, bei 4.13 und bei Folgerung 4.15).

Zu Beginn der Arbeit werden Flachkeitskriterien gegeben: Nach 4.1 ist ein Ring R_2 mit $\bar{\mathrm{I}}$ bezüglich G_2 flach über einem Unterring R_1 mit I bezüglich G_1. Ist R_2 gleich dem „vollen" Ring $R_Z(G_2)$ (den wir der Einfachheit halber meistens mit $R(G_2)$ abkürzen), so erfüllt R_2 die stärkere Eigenschaft $\bar{\mathrm{I}}$ nur dann, wenn G_2 kompakt ist (wobei wir im Nichtarchimedischen eine Menge $G_2 \subset Z$ kompakt nennen, wenn sie eine Umgebungsbasis von affinoiden Bereichen hat). Deshalb interessieren Satz 4.2 bzw. 4.3, in denen gezeigt wird, wann bei beliebigem G_2 der Ring $R(G_2)$ flach über R_1 ist. In 4.4 zeigen wir weiter, daß ein Oberring R_2, der die Eigenschaft $\bar{\mathrm{I}}$ bezüglich $G_2 \subset Z \oplus k$ erfüllt (der also aus holomorphen Funktionen mit einer zusätzlichen Veränderlichen besteht), auch flach über R_1 ist. Insbesondere ist also für eine kompakte Teilmenge $G \subset k^n$ der Ring $R_{n+1}(G)$ flach über $R_n(G)$. Satz 4.5 besagt schließlich, daß auch ein Oberring R_2 mit $\bar{\mathrm{I}}$, der aus K-analytischen Funktionen auf G besteht (wobei $K \supset k$ eine komplette Körpererweiterung ist), flach ist über dem Ring R_1 (der ja nur aus k-analytischen Funktionen bestand).

Sei nun M ein R_1-Modul mit I bezüglich G, $R_2 \subset R(G)$ sei eine Ringerweiterung von R_1. Satz 4.6 zeigt, daß bei einer Dichtheitsvoraussetzung für $R_1 \subset R_2$ aus MR_2 freier R_2-Modul (was mit $M \otimes R_2$ freier R_2-Modul äquivalent ist) schon M freier R_1-Modul folgt. Dies ist schon insofern bemerkenswert, als in der kommutativen Algebra auch bei treuflachen Ringerweiterungen $R_2 \supset R_1$ nicht aus $M \otimes R_2$ frei schon M frei folgt. In Satz 4.13 folgern wir aus 4.6, daß im komplexen für projektive R-Moduln M der lokalisierte Modul

$$M_K := \left\{ \frac{m}{r} \,\middle|\, m \in M, r \in R, r(z) \neq 0 \ \ \forall z \in K \right\}$$

frei über dem entsprechend lokalisierten Ring R_K ist, wenn $K \subset G$ kompakt in einer zusammenziehbaren Menge enthalten ist. Ähnlich ergibt sich Satz 4.13′: Endliche projektive Moduln M über Ringen R mit I bezüglich zusammenziehbaren kompakten Mengen sind schon frei.

Weiter brauchen wir 4.6 zum Beweis von Faktorialitätskriterien: Satz 4.7 besagt u.a., daß Unterringe von faktoriellen bzw. fastfaktoriellen Ringen mit I wieder diese Eigenschaften haben. Satz 4.11 zeigt, daß im Reellen ein Ring R mit I bezüglich kompaktem einfach zusammenhängendem G faktoriell ist; nach 4.9 ist er bei nur kompaktem G wenigstens noch fastfaktoriell. In 4.12 wird bewiesen, daß im komplexen bei kompakten Mengen K mit $H^2(K, \mathbb{Z}) = 0$ ein Unterring R mit I bezüglich K faktoriell ist; ist umgekehrt der volle Ring $R(K)$ faktoriell, so ist auch $H^2(K, \mathbb{Z}) = 0$ (Folgerung 4.12′). Nach 4.10 schließlich sind im Archimedischen bei einfach zusammenhängendem G für einen Ring R mit I bezüglich G die Eigenschaften „faktoriell" und „fastfaktoriell" äquivalent.

Zur Motivation der Sätze 4.14 und 4.15, die wieder für alle Körper k gelten, wollen wir folgendes viel spezielleres Problem behandeln: Wenn für $G \subset k^n$ der Ring $R_n(G)$ gewisse algebraische Eigenschaften (wie etwa noethersch oder faktoriell) erfüllt, gelten dann auch in dem viel größeren Ring $R_{n+1}(G)$ diese Eigenschaften? Und gilt die Umkehrung? Ist G speziell eine einelementige Menge $\{z\}$, so bedeutet dies, von der Eigenschaft noethersch bzw. faktoriell des Ringes $\mathcal{O}_z^n$ auf diejenige des Ringes

$\mathcal{O}_z^{n+1}$ zu schließen und umgekehrt. Dieses Problem ist mit der Weierstraßformel gelöst. Das allgemeine Problem beweisen wir hier mit einer ganz anderen rein algebraischen Methode und zeigen, daß $R_{n+1}(G)$ genau dann noethersch ist, wenn $R_n(G)$ noethersch ist und daß bei regulärem G zusätzlich $R_{n+1}(G)$ genau dann faktoriell bzw. fastfaktoriell ist, wenn dies auf $R_n(G)$ zutrifft (Folgerung 4.15). Dieser Satz ist das analytische Analogon zu dem bekannten algebraischen Satz (s. [15]), daß bei regulären noetherschen Ringen R die Eigenschaften faktoriell und fastfaktoriell bei Potenzreihenbildung $R[\![X]\!]$ erhalten bleiben. Folgerung 4.14 zeigt, daß wir sogar von der Gültigkeit einer dieser drei Eigenschaften im Ring $R_{n+1}(G')$ für irgendein G' mit $G \subset G' \subset G \oplus k$ auch auf die entsprechende Eigenschaft im Ring $R_n(G)$ runterschließen können; dabei ist natürlich hier die sowieso schwierigere umgekehrte Schlußweise vom kleineren Ring $R_n(G)$ auf den größeren Ring $R_{n+1}(G')$ nicht erlaubt, da G' und damit auch $R_{n+1}(G')$ selbst bei gutem G beliebig schlecht sein können. Satz 4.15 bzw. Satz 4.14 verallgemeinern noch diese Ergebnisse auch auf Unterringe. Mit Satz 4.15 kann man übrigens auf rein algebraische Weise zeigen, daß ein Ring mit I bezüglich einer kompakten Menge G, deren Rand reellanalytisch ist, schon noethersch ist.

Zum Schluß zeigen wir in 4.16 in Verallgemeinerung eines bekannten Satzes, daß aus der Isomorphie zweier Ringe R_i mit I bezüglich Mengen G_i (die keine k-analytischen Räume zu sein brauchen) auch schon $G_1 \cong G_2$ folgt; in Folgerung 4.16 sehen wir dann, daß dies schon erfüllt ist, wenn es nur einen treuflachen Morphismus $\varphi: R_1 \to R_2$ gibt, so daß $\varphi(R_1)$ den „Polynomring" in R_2 enthält.

Zum Beweis dieser Sätze brauchen wir eine Reihe von Hilfssätzen über globale Moduln, die hier ohne Beweis aufgeführt werden sollen und deren Beweis unter den angegebenen Stellen zu finden ist:

[A] Sei M ein R-Modul mit I bezüglich G. Ist $K \equiv \{z\} \subset G$, so ist der lokalisierte Modul $M_z := M_K$ ein noetherscher R_z-Modul ([10], 2.8).

[B] Ist $R \subset R(G)$ dicht und erfüllt der R-Modul M die Eigenschaft $\bar{\mathrm{I}}$ bezüglich G, so erfüllt für $K = \hat{K}_G$ der R_K-Modul M_K die Eigenschaft $\bar{\mathrm{I}}$ bezüglich K ([9], 3.9).

[C] Erfüllt R die Eigenschaft I, so gilt für die Komplettierungen $\hat{R}_z = \hat{\mathcal{O}}_z$ ([9], 3.4). Ist $f \in \mathcal{O}_z$, $J \subset \mathcal{O}_z$ und $f \in J\hat{\mathcal{O}}_z$, so ist schon $f \in J$ ([10], 2.7.3).

[D] Die Ringe R mit $\bar{\mathrm{I}}$ bezüglich G sind grade die Ringe R mit I, in denen jedes Ideal eine Nullstelle in G hat (in denen also Max Spec $R = G$ ist) ([10], 2.9). Die Ringe R mit III sind grade die noetherschen Ringe mit I ([10], 2.10).

[E] R erfülle I bezüglich $G \subset Z$. Dann erfüllt $(R[X])_z$ die Eigenschaft $\bar{\mathrm{I}}$ bezüglich $\{z\} \subset Z \oplus k$ für jedes $z \in G \oplus k$ ([9], 3.10).

[F] R erfülle I bezüglich $G \subset k^n$. $K \supset k$ sei eine komplette Körpererweiterung. Dann erfüllt für jedes $z \in G$ auch $(R[K])_z$ die Eigenschaft $\bar{\mathrm{I}}$ bezüglich $\{z\} \subset K^m$ ([9], 3.33).

[G] R erfülle $\bar{\mathrm{I}}$ bezüglich G. Dann erfüllt auch für jede Zahl n der R-Modul R^n die Eigenschaft $\bar{\mathrm{I}}$ bezüglich G ([9], 3.5).

[H] Sei R ein Ring mit I bezüglich $G \subset Z$. Dann gibt es zu jedem $z \in Z - G$ ein $f \in R$, das sich nicht holomorph nach z fortsetzen läßt ([10], 2.20).

[I] Sei G kompakt. Es sei $k = \mathbb{C}$ und G habe eine Umgebungsbasis von Steinschen Mannigfaltigkeiten X mit $H^2(X, \mathbb{Z}) = 0$. Der Durchschnitt zweier Hauptideale in $R(G)$ ist wieder ein Hauptideal. Ist ferner $J \subset R(G)$ ein endlich erzeugtes lokalfreies Ideal, so ist schon J Hauptideal ([8], 68).

[K] Ist $R(G)$ faktoriell, so ist jede Cousin-II-Verteilung in einer Umgebung von G durch ein $f \in R(G)$ lösbar ([8], 69).

[L] Sei R ein Ring mit I bezüglich kompaktem G. Dann ist der Durchschnitt zweier endlicher Ideale wieder ein endliches Ideal ([8], 66).

Anwendungen in der globalen Funktionentheorie

Nachdem in [9] und [10] schon einige Anwendungen der Theorie der Moduln mit I gegeben wurden, sollen jetzt weitere Folgerungen gezogen werden.

Satz 4.1. *Sei $R_2 \supset R_1$. R_2 sei ein Ring mit* $\bar{\mathrm{I}}$ *bezüglich G_2, R_1 ein Ring mit* I *bezüglich $G_1 \supset G_2$. Dann ist R_2 flach über R_1.*

Beweis. $M := \prod_{z \in G_2} (R_2)_z$ ist treuflach über R_2 nach [D]. Wegen $\widehat{(R_1)}_z = \widehat{(R_2)}_z$ (s. [G]) und $(R_i)_z$ noethersch (s. [A]) folgt, daß auch $(R_2)_z$ flach über $(R_1)_z$ ist; damit ist M auch flach über R_1. Es folgt R_2 flach über R_1.

Bemerkung 4.1. Es genügt natürlich für R_1 zu fordern, daß $(R_1)_z$ die Eigenschaft I bezüglich $\{z\}$ für alle $z \in G_1$ erfüllt.

Insbesondere folgt aus 4.1, daß für eine kompakte Steinsche Menge G_2 und für Steinsches $G_1 \supset G_2$ stets $R(G_2)$ flach über $R(G_1)$ ist. Diese Aussage kann noch verschärft werden:

Satz 4.2. *Sei die Steinsche Menge G_2 relativ kompakt in $\mathring{G}_1$. Ist R_1 ein Ring mit* I *bezüglich G_1, so ist $R(G_2)$ flach über R_1 und für $f_i \in R_1$ ist stets der Relationenmodul*

$$\operatorname{Rel}(f_1, \ldots, f_n | R(G_2)) := \{(r_1, \ldots, r_n) \in (R(G_2))^n | \sum r_i f_i = 0\}$$

ein endlicher $R(G_2)$-Modul.

Beweis. Wir überdecken G_2 durch endlich viele, in G_1 gelegene kompakte Kugeln K_i. Nach 4.1 ist $R(K_i)$ flach über R_1; somit folgt mittels der Charakterisierung von flach durch die Relationenmoduln [11]: Es ist

$$\operatorname{Rel}(f_1, \ldots, f_n | R(K_i)) = (\operatorname{Rel}(f_1, \ldots, f_n | R_1))\, R(K_i)$$

für alle Tupel $(f_1, \ldots, f_n) \in R_1^n$. Da $R(K_i)$ noethersch ist [3] und die K_i eine endliche Überdeckung bilden, gibt es somit endlich viele

$$m_1, \ldots, m_q \in M := \operatorname{Rel}(f_1, \ldots, f_n | R_1),$$

so daß

$$\operatorname{Rel}(f_1, \ldots, f_n | R(K_i)) = (m_1, \ldots, m_q)\, R(K_i) \quad \forall i \text{ ist.}$$

Ist nun $m_0 \in \operatorname{Rel}(f_1, \ldots, f_n | R(G_2))$, so ist insbesondere für jedes $z \in G_2$ auch $m_0 \in \operatorname{Rel}(f_1, \ldots, f_n | \mathcal{O}_z)$. Unter abermaliger Anwendung von 4.1 folgt, daß $\mathcal{O}_z$ flach über $R(K_i)$ für jedes $K_i \supset \{z\}$ ist; somit ist

$$\operatorname{Rel}(f_1, \ldots, f_n | \mathcal{O}_z) = \operatorname{Rel}(f_1, \ldots, f_n | R(K_i))\, \mathcal{O}_z.$$

Also ist $m_0 \in (m_1, \ldots, m_q)\, \mathcal{O}_z$. Da dies für jedes $z \in G_2$ gilt, folgt daraus $m_0 \in (m_1, \ldots, m_q)\, R(G_2)$, womit die zweite Behauptung bewiesen ist. Weiter haben wir dabei

$$\operatorname{Rel}(f_1, \ldots, f_n | R(G_2)) \subset \operatorname{Rel}(f_1, \ldots, f_n | R_1)\, R(G_2)$$

gezeigt; daraus folgt aber, daß $R(G_2)$ flach über R_1 ist.

Folgerung 4.2. *Seien G_i Steinsch, $G_2 \subset\subset \mathring{G}_1$. Dann ist $R(G_2)$ flacher $R(G_1)$-Modul.*

Für den vollen Ring $R_1 = R(G_1)$ hätten wir 4.2 auch direkt mittels Theorem A und Theorem B gewinnen können: Denn

$$\mathscr{F}_z := \operatorname{Rel}(f_1, \ldots, f_n | \mathcal{O}_z)$$

bildet eine kohärente Garbe, die durch endlich viele globale Schnitte $\Gamma(G_1, \mathscr{F})$ über G_2 erzeugt wird. Ähnlich wie 4.2 zeigen wir:

Satz 4.3. *Sei G_2 Steinsch. R_1 sei ein Ring mit* III *bezüglich $G_1 \supset G_2$. Dann ist $R(G_2)$ flach über R_1.*

Beweis. $M := \operatorname{Rel}(f_1, \ldots, f_n | R_1)$ ist wegen [D] ein endlicher R_1-Modul, wenn $f_i \in R_1$ ist. Weiter folgt aus $m_0 \in \operatorname{Rel}(f_1, \ldots, f_n | R(G_2))$, daß für jedes $z \in G_2$ auch $m_0 \in \operatorname{Rel}(f_1, \ldots, f_n | \mathcal{O}_z)$ ist; dies bedeutet, da nach 4.1 auch $\mathcal{O}_z$ flach über R_1 ist, daß $m_0 \in M\mathcal{O}_z\ \forall z \in G_2$ ist. Damit folgt wegen $MR(G_2)$ endlicher $R(G_2)$-Modul schon $m_0 \in MR(G_2)$, womit die Behauptung folgt.

Zusatz 4.3. Wir brauchen in 4.3 für R_1 nur vorauszusetzen: R_1 ist noethersch und $(R_1)_z$ erfüllt I bezüglich $\{z\}$ für alle $z \in G_1$.

Insbesondere folgt aus dem Zusatz, daß bei Steinschem G der Ring $R(G)$ flach über dem Polynomring ist. Für algebraisch abgeschlossene k

folgt wegen der Gültigkeit des algebraischen Hilbertschen Nullstellensatzes, daß $R(k^n)$ sogar treuflach über dem Polynomring ist, was für $k=\mathbb{R}$ sicher falsch ist. Für alle Körper ist in Satz 4.3 bei $G_1=G_2$ aber trivialerweise stets $R(G_2)$ treuflach über R_1.

Wie zu Anfang erwähnt wurde, kann dieselbe Teilmenge $G\subset k^m$ auch funktionentheoretisch sinnvoll als Teilmenge $G\subset k^{m+1}$ betrachtet werden. Über den Zusammenhang der dabei auftretenden verschiedenen Ringe gibt der folgende Satz Auskunft:

Satz 4.4. *Sei R_1 ein Ring mit* I *bezüglich $G_1\subset Z$. Dann ist jeder Ring $R_2\supset R_1$, der* $\bar{\mathrm{I}}$ *bezüglich $G_2\subset Z\oplus k$ erfüllt, flach über R_1.*

Beweis. Zunächst ist $G_2\subset G_1\oplus k$, da sämtliche Funktionen aus R_1 nach Voraussetzung auch in G_2 holomorph sind und wegen [H] aber $G_1\oplus k$ der größte Holomorphiebereich für R_1 ist. Deshalb erfüllt nach [E] der Ring $(R_1[X])_z$ die Eigenschaft I bezüglich $\{z\}$ für jedes $z\in G_2$. Wegen Bemerkung 4.1 ist R_2 flach über $R_1[X]$. Da Polynomringbildung eine flache Operation ist, folgt die Behauptung.

Folgerung 4.4.1. *Für jedes kompakte G' mit $G\subset G'\subset G\oplus k$ ist $R_{Z\oplus k}(G')$ treuflach über $R_Z(G)$.*

Folgerung 4.4.2. *Ist $G\subset k^m$ kompakt, so ist $R_{m+1}(G)$ flach über $R_m(G)$.*

Ganz speziell folgt daraus das bekannte Ergebnis, daß $\mathcal{O}_z^{m+1}$ treuflach über $\mathcal{O}_z^m$ ist. Weiter können wir entsprechend 4.4 dieselbe Menge $G\subset k^m$ auch als Teilmenge des K^m bei einem größeren Körper $K\supset k$ auffassen. Dann gilt:

Satz 4.5. *Sei K eine Körpererweiterung von k und sei R_1 ein Ring von k-analytischen Funktionen mit* I *bezüglich $G\subset k^m$. Dann ist jeder Ring $R_2\supset R_1$ von K-analytischen Funktionen mit* $\bar{\mathrm{I}}$ *bezüglich $G\subset K^m$ eine flache Ringerweiterung.*

Beweis. Nach [F] erfüllt für jedes $z\in G$ auch $(R_1[K])_z$ I bezüglich $\{z\}\subset K^m$. Nach Bemerkung 4.1 folgt dann, daß R_2 flach über $R_1[K]$ ist. $R_1[K]$ ist als freier R_1-Modul aber auch flach über R_1.

Wir wollen nun weiter die Situation $R_1\subset R_2$ betrachten. In [9] haben wir gesehen, daß sich die Eigenschaft „noethersch" von R_2 auf R_1 überträgt. Wichtig für das Weitere ist, daß auch die Eigenschaft „faktoriell" meistens erhalten bleibt. Dazu beweisen wir einen allgemeineren Satz. Wir wollen vorher jedoch abkürzend definieren:

Definition 4.6. Ein Ring $R_1\subset R_2$ heißt *dicht* in R_2, wenn es zu jeder Summe $\sum_1^k a_i b_i$ mit $a_i, b_i\in R_2$, die in R_2^* liegt, irgendwelche $c_i\in R_1$ gibt, so daß auch $\sum_1^k a_i c_i\in R_2^*$ ist.

Satz 4.6. *Sei M ein R_1-Modul mit* I *bezüglich G; R_1 sei in einem beliebigen Oberring $R_2 \subset R(G)$ dicht. Ist MR_2 ein endlicher freier R_2-Modul, so ist auch M schon ein endlicher freier R_2-Modul vom gleichen Rang n.*

Zusatz 4.6. Erfüllt M nicht I, so folgt immer die Existenz eines freien Untermoduls $N \subset M$ mit $NR_2 = MR_2$.

Beweis. Es sei $MR_2 = (f_1, \ldots, f_n) R_2$. Jedes f_i hat die Gestalt

$$f_i = \sum_1^k g_v r_v^i \qquad \text{mit } g_v \in M, r_v^i \in R_2.$$

Es folgt

$$g_v = \sum_1^n f_u R_u^v \qquad \text{mit } R_u^v \in R_2.$$

Also ist

$$f_i = \sum_{u=1}^n f_u \sum_{v=1}^k r_v^i R_u^v.$$

Da f_i eine Basis von MR_2 sein sollte, ist $\sum_{v=1}^k r_v^i R_u^v = \delta_u^i$ und somit $\det\left(\sum_{v=1}^k r_v^i R_u^v\right)_{i,u}$ Einheit in R_2. Wir approximieren dann r_v^i durch $s_v^i \in R_1$, so daß auch $\det\left(\sum_{v=1}^k s_v^i R_u^v\right)$ Einheit in R_2 ist (dies geschieht mittels der Voraussetzung sukzessive: Erst werden alle r_v^1 durch s_v^1 ersetzt, so daß wir wieder eine Einheit f_1^* bekommen; in dieser Funktion f_1^* werden alle r_v^2 durch s_v^2 ersetzt usw.). Es ist also die Abbildung $MR_2 \to MR_2$: $f_i \to \sum_{u=1}^n f_u \sum_{v=1}^k s_v^i R_u^v$ surjektiv; d.h. es gibt $a_{ij} \in R_2$ mit

$$\begin{aligned} f_j &= \sum_{i=1}^n a_{ij} \sum_{u=1}^n f_u \sum_{v=1}^k s_v^i R_u^v \\ &= \sum_{i=1}^n a_{ij} \sum_{v=1}^k s_v^i \sum_{u=1}^n f_u R_u^v = \sum_{i=1}^n a_{ij} \sum_{v=1}^k s_v^i g_v. \end{aligned}$$

Es ist aber $\sum_{v=1}^k s_v^i g_v =: h_i \in M$ und $MR_2 = (h_1, \ldots, h_n) R_2$; somit folgt $M \subset (h_1, \ldots, h_n) \mathcal{O}_z \; \forall z \in G$ und damit $M = (h_1, \ldots, h_n) R_1$. Da $h_1, \ldots, h_n$ ein Erzeugendensystem von MR_2 ist, ist es auch eine Basis von MR_2 und damit auch eine Basis von M.

Satz 4.6. läßt sich bei gleichem Beweis auch algebraisieren und gilt für beliebige treuflache Ringerweiterungen von R_1.

Folgerung 4.6.1. *Sei G Steinsch, $R \subset R(G)$ ein dichter Ring, M ein R-Modul mit* I *bezüglich G und $\mathcal{M}$ die zu M assoziierte kohärente Garbe. Wird $\mathcal{M}$ über $K = \hat{K}_G \subset G$ durch n freie Schnitte $s_1, \ldots, s_n \in \Gamma(K, \mathcal{M})$ erzeugt (d.h. gibt es $s_1, \ldots, s_n \in \Gamma(K, \mathcal{M})$, so daß $s_1, \ldots, s_n$ eine Basis von $\mathcal{M}_z$*

in jedem Punkt von K ist) so gibt es auch schon n freie globale Schnitte $m_1, \ldots, m_n \in \Gamma(G, \mathcal{M})$, *die die Garbe* $\mathcal{M}$ *über K erzeugen und die zudem noch in M liegen.*

Beweis. Nach [B] erfüllt der R_K-Modul M_K auch I bezüglich K. Nach Voraussetzung ist $\Gamma(K, \mathcal{M}) = (M_K) R(K)$ ein freier $R(K)$-Modul mit der Basis $(s_1, \ldots, s_n)$. Ferner ist R_K dicht in $R(K)$, da R dicht in $R(G)$ ist und nach [6] auch $R(G)$ dicht in $R(K)$ ist. Somit ist nach 4.6 auch M_K ein freier R_K-Modul. Es gibt also $m_1, \ldots, m_n \in M$ und $s \in R$, $s(z) \neq 0$ $\forall z \in K$, so daß $\frac{m_1}{s}, \ldots, \frac{m_n}{s}$ eine Basis von M_K ist. Dann erfüllen die m_i das Verlangte.

Die Garbe $\mathcal{M}$ wird übrigens genau dann über K durch n freie Schnitte erzeugt, wenn $\mathcal{M}_z$ freier $\mathcal{O}_z$-Modul vom Rang n ist und $\mathcal{M}$ über K durch n Schnitte erzeugt wird.

Folgerung 4.6.2. *Sei G Steinsch,* $\mathcal{M}$ *eine kohärente Garbe auf G. Wird* $\mathcal{M}$ *über* $K = \hat{K}_G$ *durch n freie Schnitte aus* $\Gamma(K, \mathcal{M})$ *erzeugt, so gibt es auch schon n globale freie Schnitte aus* $\Gamma(G, \mathcal{M})$, *die* $\mathcal{M}$ *über K erzeugen.*

Folgerung 4.6.3. *Ist M ein projektiver R-Modul mit* I *bezüglich G vom Rang n und wird die zu M assoziierte kohärente Garbe* $\mathcal{M}$ *über* $K \subset G$ *durch n Schnitte erzeugt, so gibt es n globale Schnitte aus M, die einen freien Untermodul von M aufspannen und die* $\mathcal{M}$ *über K erzeugen.*

Insbesondere folgt noch mit 4.6 bei $k = \mathbb{R}$ und kompaktem G wegen der Gültigkeit des Weierstraßschen Approximationssatzes: Ist $NR(G)$ frei, so ist auch schon N frei. Dieser Satz läßt sich sogar verschärfen: Ist N ein R-Modul mit I und $NC^\infty(G)$ freier $C^\infty(G)$-Modul, so ist schon N freier R-Modul. (Denn nach Weierstraß ist R wieder dicht in $C^\infty(G)$; also gibt es wie in 4.6 einen freien Untermodul $N' \subset N$ mit $NC^\infty(G) = N' C^\infty(G)$. Daraus folgt, da der Ring $C^\infty(z)$ der unendlich oft differenzierbaren Funktionskeime in z treuflach über $\mathcal{O}_z$ ist [11], daß schon $N \subset N' \mathcal{O}_z$ $\forall z \in G$ und damit $N = N'$ ist.)

Wir wollen nun 4.6 weiter anwenden. Dazu definieren wir: Ein Ring R heiße fastfaktoriell, wenn es erstens zu $f, g \in R$ ein q gibt, so daß $f^q R \cap g^q R$ ein Hauptideal ist und zweitens R der aufsteigenden Kettenbedingung für Hauptideale genügt. Ist R noethersch, so ist dies die übliche Definition (s. [15]); außerdem ist in *jedem* Fall ein faktorieller Ring bei dieser Terminologie festfaktoriell.

Satz 4.7. *Sei G kompakt,* $R_1 \subset R_2$ *mögen* I *erfüllen und es sei* R_1 *bezüglich der Maximumsnorm dicht in* R_2.

Ist für $f, g \in R_1$ *der Durchschnitt* $f R_2 \cap g R_2$ *ein Hauptideal in* R_2, *so ist auch* $f R_1 \cap g R_1$ *ein Hauptideal in* R_1.

Ist für $J \subset R_1$ das Erweiterungsideal JR_2 ein Hauptideal, so ist bereits J Hauptideal.

Ist insbesondere R_2 faktoriell, so ist R_1 faktoriell; ist R_2 fastfaktoriell, so ist R_1 fastfaktoriell.

Beweis. Sei $J:=fR_1 \cap gR_1$. Wegen 4.1 ist $JR_2 = fR_2 \cap gR_2$; somit ist also JR_2 frei. Damit ist nach 4.6 auch J Hauptideal. Entsprechend folgt die zweite Behauptung. Die dritte Behauptung ergibt sich so: Wäre $f_i R_1$ eine echt aufsteigende Hauptidealkette in R_1, so wäre wegen der Eigenschaft I für R_1 auch $f_i R_2$ eine echt aufsteigende Hauptidealkette in R_2, was aber unmöglich ist. Somit ist R_1 faktoriell, wenn R_2 faktoriell ist. Ist R_2 nur fastfaktoriell, so gibt es ein $q \in \mathbb{N}$, so daß $f^q R_2 \cap g^q R_2$ ein Hauptideal in R_2 ist; dann ist nach Obigem auch $f^q R_1 \cap g^q R_1$ Hauptideal in R_1.

Im Reellen ist die Voraussetzung „R_1 dicht in R_2" natürlich immer erfüllt, da der Weierstraßsche Approximationssatz gilt.

Für nichtarchimedische Körper folgt sofort aus 4.7:

Satz 4.8. *Sei $R \subset T_n$ ein Unterring mit* I *bezüglich E_n. Dann ist R faktoriell und noethersch und erfüllt sogar* III.

Beispiel 4.8. Sei R der Ring der im „abgeschlossenen Polyzylinder holomorphen Funktionen":

$$R = \{\sum_i a_i z^i \mid \exists\, r>1 \text{ mit } |a_i|\, r^i \to 0\}.$$

Für algebraisch abgeschlossene k ist R ein faktorieller noetherscher Ring mit III bezüglich dem Einheitszylinder.

Beweis. Es genügt zu zeigen, daß R I erfüllt. Sei also

$$f_0 \in (f_1, \ldots, f_n)\, \mathcal{O}_z \qquad \forall\, z \in E_n.$$

Dann ist $f_i \in T_n(r)$ für ein $r>1$ ($T_n(r)$ bedeutet der Tatesche Ring auf dem Polyzylinder mit dem Radius r). Nach [B] folgt dann $f_0\, s \in (f_1, \ldots, f_n)\, T_n(r)$ für ein $s \in T_n(r)$, $s(z) \neq 0$ $\forall\, z \in E_n$. Nun hat s auch keine Nullstelle in einem Polyzylinder $P_n(r')$ mit $1<r' \leqq r$ (denn es gilt allgemein

$$s(E_n) = \{a \in k \mid |a - s(0)| \leqq \|s - s(0)\|\};$$

woraus bei $s = \sum a_i X^i$ wegen $0 \notin s(E_n)$ folgt, daß $|s(0)| > |a_i|$ $\forall\, |i|>0$ ist, daß somit auch $|s(0)| > |\sum_{|i|>0} a_i (r')^i|$ für ein $r'>1$ ist). Dann ist $\frac{1}{s} \in T_n(r')$, somit $\frac{1}{s} \in R$.

Es ist klar, daß wir Faktorialitätseigenschaften für Ringe R mit I nur dann erwarten können, wenn $\mathcal{O}_z$ selbst faktoriell ist (einfach deshalb, weil $\mathcal{O}_z$ selbst ein Ring mit I bezüglich z ist). Somit nehmen wir jetzt an, daß der k-analytische Raum Z nur rationale Singularitäten habe; d.h. $\mathcal{O}_z^Z$ ist $\forall z \in G$ faktoriell. In diesem Fall gilt auch, daß R_z $\forall z \in G$ faktoriell ist. (Denn $\mathcal{O}_z$ ist treuflach über R_z nach [C] und [A]. Somit ist

$$(f R_z \cap g R_z)\, \mathcal{O}_z = f \mathcal{O}_z \cap g\, \mathcal{O}_z$$

ein Hauptideal in $\mathcal{O}_z$. Ist $f_1, \ldots, f_k$ ein minimales Erzeugendensystem von $f R_z \cap g R_z$, so ist es wegen $\mathcal{O}_z$ treuflach über R_z auch ein minimales Erzeugendensystem von $(f R_z \cap g R_z)\, \mathcal{O}_z$; somit ist $k=1$. – Bei diesem Beweis wurde also im Gegensatz zu [8] nicht benutzt, daß jeder reguläre lokale Ring A faktoriell ist; diesen rein algebraischen Satz kann man dann umgekehrt im charakteristikgleichen Fall leicht ähnlich wie hier zeigen, wenn man R_z durch A und $\mathcal{O}_z$ durch den nach Weierstraß faktoriellen Ring $\hat{A} = A/\mathfrak{m}[\![z_1, \ldots, z_n]\!]$ ersetzt.)

Satz 4.9. *Sei $k=\mathbb{R}$. R sei ein Ring mit* I *bezüglich kompaktem G. Sind $f_i \in R$, so ist $\bigcap_1^n f_i^2 R$ und $\left(\bigcap_1^n f_i R\right)^2$ ein Hauptideal in R.*

Weiter ist für jedes endlich erzeugte Primideal p der Höhe 1 *in R schon p^2 ein Hauptideal. Allgemeiner ist für jedes endlich erzeugte lokalfreie Ideal J schon J^2 ein Hauptideal.*

Genügt also R der aufsteigenden Kettenbedingung für Hauptideale, so ist R fastfaktoriell.

Beweis. Sei $J := f^2 R \cap g^2 R$. Da R_z faktoriell ist, gibt es zu jedem $z \in G$ ein $h_z \in R$ mit $J R_z = (h_z)^2 R_z$. O.B.d.A. ist $(h_z)^2 \in J$. Nach [L] ist ferner J endlich erzeugt. Es gibt dann zu $z_i \in G$ eine Umgebung $U(z_i)$ mit $J R_z = (h_{z_i})^2 R_z$ $\forall z \in U(z_i)$. Es mögen nun $U(z_1), \ldots, U(z_q)$ schon G überdecken. Dann sei $h := \sum_1^q (h_{z_i})^2$. Wegen $(h_{z_j})^2 \in J$ folgt für $z \in U(z_i)$, daß $(h_{z_j})^2 = (h_{z_i})^2 s_{ij}$ ist, daß also $h = (h_{z_i})^2 \sum s_{ij}$ ist. Wegen $s_{ii}=1$ und $s_{ij} \geqq 0$ ist damit $h\, \mathcal{O}_z = (h_{z_i})^2\, \mathcal{O}_z$. Somit ist $h\, \mathcal{O}_z = J\, \mathcal{O}_z$ $\forall z \in G$ und deshalb $J = hR$. Die anderen Aussagen beweisen wir entsprechend.

Satz 4.10. *Sei G einfach zusammenhängend. Ist $R \subset R(G)$ dicht und erfüllt R die Eigenschaft* I, *so ist R fastfaktoriell genau dann, wenn R faktoriell ist.*

Beweis. Es sei $f^q R \cap g^q R = hR$. Da R_z faktoriell $\forall z \in G$ ist, ist aus h lokal die q-te Wurzel zu ziehen; somit gibt es ein $k \in R(G)$, so daß $k^q = h$ ist. Aus $k^q R_z = f^q R_z \cap g^q R_z$ folgt aber wegen 4.1 auch

$$k^q (R(G))_z = f^q (R(G))_z \cap g^q (R(G))_z;$$

dann folgt wegen der Faktorialität von $(R(G))_z$, daß

$$k (R(G))_z = f (R(G))_z \cap g (R(G))_z$$

ist. Somit ist $kR(G)=fR(G)\cap gR(G)$; nach 4.7 folgt dann, daß auch schon $fR\cap gR$ ein Hauptideal ist.

Aus Satz 4.9 und Satz 4.10 bzw. Beweis 4.10 folgt leicht unter Benutzung des Weierstraßschen Approximationssatzes:

Satz 4.11. *Sei $k=\mathbb{R}$ und G kompakt einfach zusammenhängend. Ist R ein Ring mit I bezüglich G, so ist jedes endliche lokal freie Ideal und jedes endliche Primideal der Höhe 1 ein Hauptideal und in R ist der Durchschnitt zweier Hauptideale wieder ein Hauptideal.*

Erfüllt R zusätzlich die Kettenbedingung für Hauptideale, so ist R faktoriell.

Bemerkenswert ist hierbei, daß es auch bei einfachzusammenhängendem kompaktem G Primideale der Höhe 1 in $R(G)$ geben kann, die keine Hauptideale sind und somit nicht endlich erzeugt sein können. In diesem Fall genügt dann nach 4.11 der Ring $R(G)$ nicht der Kettenbedingung für Hauptideale.

Für $k=\mathbb{C}$ gilt schließlich:

Satz 4.12. *Sei $k=\mathbb{C}$. Das kompakte G besitze eine Umgebungsbasis von Steinschen Räumen M_i mit $H^2(M_i,\mathbb{Z})=0$. Ist R ein dichter Unterring von $R(G)$ mit I bezüglich G, so ist in R der Durchschnitt zweier Hauptideale ein Hauptideal und jedes endliche Primideal der Höhe 1 ein Hauptideal.*

Beweis. [J] und 4.7 ergeben die Behauptungen.

Satz 4.12'. *Sei $k=\mathbb{C}$, G kompakt Steinsch und $H^2(G,\mathbb{Z})=0$. Ist R ein dichter Unterring von $R(G)$ mit I, so ist in R der Durchschnitt zweier Hauptideale ein Hauptideal und jedes endliche Primideal der Höhe 1 ein Hauptideal.*

Beweis. Wie im Fall, daß G offen ist (s. etwa [6]), folgt $H^1(G,\mathcal{O}^*)\cong H^2(G,\mathbb{Z})$, da auch für kompakte Steinsche Mengen G bei $i>0$ stets $H^i(G,\mathcal{O})=0$ ist. Somit folgt aus unserer Voraussetzung, daß jedes Cousin-Problem auf G lösbar ist. Dann folgt wie in 4.12 die Behauptung.

Folgerung 4.12'. *Sei $k=\mathbb{C}$, G kompakt Steinsch. Genau dann ist $R(G)$ faktoriell, wenn $R(G)$ der Kettenbedingung für Hauptideale genügt und $H^2(G,\mathbb{Z})=0$ ist.*

Beweis. Die Umkehrung wurde gerade gezeigt. Ist nun $R(G)$ faktoriell, so ist nach [K] jedes Cousin-II-Problem in einer Umgebung von G lösbar. Wenn wir ein Element aus $H^1(G,\mathcal{O}^*)$ haben, so entspricht das einer Verteilung (U_{ij},g_{ij}) mit $g_{ij}\in R(U_{ij})$ und den bekannten Relationen dazwischen. Diese läßt sich auf einer offenen Steinschen Mannigfaltigkeit $G'\supset G$ forsetzen und ist nach [6] dort zu einer Cousin-II-Verteilung äquivalent. Da diese auf G nach Obigem lösbar ist, ist $H^1(G,\mathcal{O}^*)=0$ und somit $H^2(G,\mathbb{Z})=0$.

4.12 bzw. 4.12′ bleiben für $k=\mathbb{R}$ bei gleichem Beweis richtig, wenn wir $H^2(G,\mathbb{Z})$ durch $H^2(G,\mathbb{Z}_2)$ ersetzen, wobei $\mathbb{Z}_2=\mathbb{Z}/2\mathbb{Z}$ ist (s. [1]).

Beispiel 4.12. Sei $G=\prod G_i$ mit einfach zusammenhängenden kompakten $G_i\subset\mathbb{C}$. Ist R ein Ring mit I, so ist in R der Durchschnitt zweier Hauptideale ein Hauptideal.

Wenn wir tiefere Sätze aus der globalen Funktionentheorie heranziehen, können wir mit Satz 4.6 auch weitergehende Anwendungen machen:

Satz 4.13. *Sei $k=\mathbb{C}$. $R\subset R(G)$ sei ein dichter Unterring mit* I *bezüglich Steinschem G und M sei ein endlicher projektiver R-Modul. Für jede Teilmenge $K\subset G$, bei der $\hat{K}_G$ kompakt in einem zusammenziehbaren Steinschen Raum $G'\subset G$ enthalten ist, ist der lokalisierte R_K-Modul M_K schon frei.*

Beweis. Für $K=\hat{K}_G$ erfüllt nach [B] der Ring R_K die Eigenschaft $\bar{\mathrm{I}}$ bezüglich K. Also erfüllt nach [G] jeder endliche freie R_K-Modul I bezüglich K und somit erfüllt auch jeder endliche projektive R_K-Modul M_K als Untermodul eines endlichen freien die Eigenschaft I.

Weiter ist nach [6] der Ring $R(G)$ dicht in $R(K)$ bezüglich der K-Maximumsnorm; dann ist auch R dicht in $R(K)$ und damit erst recht R_K dicht in $(R(G'))_K$.

Nun ist M R-projektiv, d.h. MR_z ist frei für alle $z\in G$. Dann ist auch $M\mathcal{O}_z$ frei für alle $z\in G'$; somit ist nach [2] $MR(G')$ projektiver $R(G')$-Modul, also nach [4] auch freier $R(G')$-Modul. Dann ist $M(R(G'))_K$ freier $(R(G'))_K$-Modul; nach 4.6 folgt dann, daß M_K freier R_K-Modul ist. Damit ist die Behauptung für $K=\hat{K}_G$ gezeigt; für $K\subset\hat{K}_G$ folgt sie dann trivialerweise, da nach noch mehr Funtionen lokalisiert wird.

Beispiel 4.13. Ist M ein projektiver Modul über einem Ring R mit I bezüglich $\mathbb{C}^n$ (also z.B. über $R=\mathbb{C}[z_1,\ldots,z_n]$), so ist für jedes kompakte $K\subset\mathbb{C}^n$ schon M_K freier R_K-Modul.

Eine ähnliche Aussage liefert der folgende Satz:

Satz 4.13′. *Sei $k=\mathbb{C}$. Die kompakte Menge G habe eine Umgebungsbasis von zusammenziehbaren Steinschen Mannigfaltigkeiten. $R\subset R(G)$ sei ein dichter Unterring mit* I. *Dann ist jeder endliche projektive R-Modul M schon frei.*

Beweis. Sei $M\oplus Q=R^n$. Wie in 4.13 zeigen wir, daß M die Eigenschaft I besitzt. Es gibt sodann eine Umgebung $U\supset G$ und in U holomorphe Funktionen m_{ij} und q_{ij}, so daß $M=\sum_j(m_{1j},\ldots,m_{nj})R$ und $Q=\sum_j(q_{1j},\ldots,q_{nj})R$ ist; somit ist für eine geeignete zusammenziehbare

Mannigfaltigkeit $U' \subset U$, $U' \supset G$ auch

$$\sum_j (m_{1j}, \ldots, m_{nj}) R(U') \oplus \sum_j (q_{1j}, \ldots, q_{nj}) R(U') = (R(U'))^n.$$

Es ist damit $\sum_j (m_{1j}, \ldots, m_{nj}) R(U')$ projektiv und deshalb nach [4] frei; dann ist, da $R(G)$ nach 4.1 flach über $R(U')$ ist, auch

$$MR(G) = \sum_j (m_{1j}, \ldots, m_{nj}) R(G)$$

freier $R(G)$-Modul. Somit folgt nach 4.6, daß M freier R-Modul ist.

Beispiel 4.13'. Ist $G = \prod G_i$, $G_i \subset \mathbb{C}$ zusammenziehbar kompakt, so sind für jeden Ring R mit I die Eigenschaften frei und projektiv für endliche Moduln äquivalent.

Wir können für die gerade gewonnenen Ergebnisse auch mehr geometrische Interpretationen geben: Ist bei $k = \mathbb{C}$ eine über G gegebene kohärente Garbe $\mathcal{M}$ über dem zusammenziehbaren $K = \hat{K}_G$ lokal frei vom Rang n, so gibt es n globale freie Schnitte aus $\Gamma(G, \mathcal{M})$, die $\mathcal{M}$ über K erzeugen (dies folgt aus 4.13 mit 4.6.2). Für dichte Unterringe $R \subset R(G)$ gibt es weiter zum projektiven R-Modul M vom Rang n und zu der zusammenziehbaren kompakten Menge $K \subset G$ einen freien Untermodul $N \subset M$ vom Rang n mit $N\mathcal{O}_z = M\mathcal{O}_z$ $\forall z \in K$.

Wir haben in den vorherigen Sätzen von Eigenschaften des Ringes R_2 auf die entsprechenden Eigenschaften des Ringes R_1 geschlossen, wenn $R_1 \subset R_2$ Ringe mit I bezüglich G waren. Es ist nun noch interessant, folgende Fragestellung zu untersuchen: R_1 erfülle I bezüglich $G \subset k^m$ und $R_2 \supset R_1$ sei faktoriell und erfülle I bezüglich $G \subset k^{m+1}$. Ist dann auch R_1 faktoriell? Der nachfolgende Satz gibt dazu eine allgemeinere Aussage:

Satz 4.14. *Sei R_2 ein noetherscher bzw. fastfaktorieller bzw. faktorieller Ring mit $\bar{\mathrm{I}}$ bezüglich $G \subset Z \oplus k$. Ist $R_1 \subset R_2$ ein Ring mit I bezüglich $G \cap Z \subset Z$ und läßt sich jedes $f_2 \in R_2$ auf $G \cap Z$ durch ein $f_1 \in R_1$ approximieren, so ist auch R_1 noethersch bzw. faktoriell bzw. fastfaktoriell.*

Beweis. Zunächst gibt es in R_1 keine aufsteigenden Ketten von Hauptidealen bzw. von endlich erzeugten Idealen J_i, wenn es in R_2 keine solchen Ketten gibt: Denn es müßte $J_i R_2 = J_{i_0} R_2$ für alle $i \geqq i_0$ sein, woraus $J_i \subset J_{i_0} \mathcal{O}_z^{Z \oplus k}$ $\forall z \in G$ und damit auch $J_i \subset J_{i_0} \mathcal{O}_z^Z$ $\forall z \in G^* := G \cap Z$ folgt; somit ergibt sich $J_i \subset J_{i_0}$.

Ist nun $J := f^q R_1 \cap g^q R_1$ für gewisse $f, g \in R_1$, so ist wegen 4.4 auch $JR_2 = f^q R_2 \cap g^q R_2$; somit ist für $q=1$ bzw. für großes q auch $JR_2 \subset R_2$ und damit erst recht $J(R_2)_{G^*} \subset (R_2)_{G^*}$ ein Hauptideal. Nun ist R_1 dicht in $(R_2)_{G^*}$ im Sinne der Definition 4.6 (nicht aber R_1 dicht in dem kleineren Ring R_2!), da jedes $f_2 \in R_2$ nach Voraussetzung auf G^* beliebig gut durch ein $f_1 \in R_1$ approximiert werden kann und die Einheiten in $(R_2)_{G^*}$ grade die Funktionen ohne Nullstelle in G^* sind. Somit folgt nach Zusatz 4.6, daß es ein $h \in J$ gibt mit $h(R_2)_{G^*} = J(R_2)_{G^*}$. Damit folgt $J \subset h\,\mathcal{O}_z^{Z \oplus k}\ \forall z \in G^*$, also schon $J \subset h\,\mathcal{O}_z^Z\ \forall z \in G^*$ und damit $J = hR$.

Folgerung 4.14. *Sei $G \subset k^m$ und G' irgendeine Menge mit $G \subset G' \subset G \oplus k$. Ist $R_{m+1}(G')$ noethersch bzw. faktoriell bzw. fastfaktoriell, so ist auch $R_m(G)$ noethersch bzw. faktoriell bzw. fastfaktoriell.*

Hieran anschließend behandeln wir das umgekehrte Problem; dabei ist jedoch trivialerweise für beliebige Mengen G' mit $G \subset G' \subset G \oplus k$ keine Aussage zu erwarten. Wir nennen nun G regulär, wenn $\mathcal{O}_z^Z$ regulär $\forall z \in G$ ist. Dann gilt:

Satz 4.15. *Sei R_1 ein Ring mit* I *bezüglich $G \subset Z$. $R_2 \supset R_1$ sei ein Ring mit* I *bezüglich $G \subset Z \oplus k$.* Wenn *$f \in R_2$ ist, sei $f^* \in R_1$, wobei f^* durch $f^*(z) := f(z, 0)$ definiert ist.*

Genau dann ist R_1 noethersch, wenn R_2 noethersch ist. Ist G regulär, so ist genau dann zusätzlich R_1 faktoriell bzw. fastfaktoriell, wenn R_2 zusätzlich faktoriell bzw. fastfaktoriell ist.

Beweis. Eine Richtung dieser Aussagen folgt mit 4.14, da trivialerweise jede Funktion aus R_2 sich auf G durch eine Funktion aus R_1 approximieren läßt.

Zur Umkehrung zeigen wir zunächst, daß kanonisch $R_2 \subset A := R_1[\![z']\!]$ ist: Denn ist $f \in R_2$, so haben wir für jedes feste $z_0 \in G$ eine Entwicklung

$$f(z, z') = \sum f_n^{(z_0)}(z) \cdot (z')^n \quad \text{für } (z, z') \in (G \oplus k) \cap U(z_0).$$

Man erkennt, daß die Funktionen $f_n^{(z_0)}(z)$ in ganz G holomorph sind und nicht von z_0 abhängen. Sei weiter schon $f_n \in R_1$ für alle $n < m$. Dann ist

$$(z')^m \cdot \sum_{n \geq m}^{\infty} f_n(z) \cdot (z')^{n-m} = f(z, z') - \sum_{0}^{m-1} f_n(z) \cdot (z')^n =: g_m(z, z'),$$

woraus $(z')^m \in g_m\,\mathcal{O}_z^{Z \oplus k}\ \forall z \in G$ folgt. Wegen $f \in R_2, f_n \in R_1 \subset R_2$ für $n < m$ und $z' \in R_2$ ist dann auch $g_m \in R_2$; da R_2 die Eigenschaft I bezüglich G erfüllte, folgt $(z')^m \in g_m R_2$, und somit ist

$$g'_m(z, z') := \sum_{n \geq m} f_n(z) \cdot (z')^{n-m} \quad \text{aus } R_2.$$

Nach Voraussetzung ist dann auch $f_n(z) = g'_m(z, 0)$ aus R_1. Damit sind alle f_n aus R_1, woraus $R_2 \subset A$ folgt.

Da R_1 noethersch ist, ist auch A noethersch. Ist nun $J \subset R_2$ ein Ideal, so gibt es endlich viele Elemente $f_i \in J$ mit $(f_1, \ldots, f_n)\,A = JA$. Für jedes $f \in J$ ist damit $f \in (f_1, \ldots, f_n)\,\hat{\mathcal{O}}_z^{Z \oplus k}\ \forall z \in G$ erfüllt, woraus mit [C] schon $f \in (f_1, \ldots, f_n)\,\mathcal{O}_z^{Z \oplus k}\ \forall z \in G$ und wegen der Eigenschaft I für R_2 auch $f \in (f_1, \ldots, f_n)\,R_2$ folgt.

Sei nun G regulär. Es ist also $\mathcal{O}_z$ regulär $\forall z\in G$. Wegen [A] und [C] ist R_z regulär $\forall z\in G$. Da R_1 noethersch ist und Max Spec $R_1=G$ ist, ist dann auch R_1 ein regulärer Ring. Nun besagt ein Satz aus [15], daß bei einem faktoriellen bzw. fastfaktoriellen regulären noetherschen Ring R_1 auch $R_1[\![z']\!]\equiv A$ faktoriell bzw. fastfaktoriell ist.

Es sei jetzt $f,g\in R_2$ und $J:=f^q R_2\cap g^q R_2$. Für $z\in G$ gilt für die Lokalisationen $\hat{A}_z=\hat{\mathcal{O}}_z^{Z\oplus k}=\widehat{(R_2)_z}$, womit A_z flach über $(R_2)_z$ folgt. Damit ist auch $JA_z=f^q A_z\cap g^q A_z$. Nun ist Max Spec $A=G$: Denn ist $\sum f_i A$ ein Ideal in A, wobei die f_i die Gestalt $f_i=\sum_j f_{ij}(z')^j$ mit $f_{ij}\in R_1$ haben mögen, so müssen die f_{i0} mindestens eine gemeinsame Nullstelle $z_0\in G$ haben, da andernfalls wegen der Eigenschaft I für R_1 Funktionen $g_i\in R_1$ existierten mit $\sum g_i f_{i0}=1$ und dann auch $\sum g_i f_i$ Einheit in A wäre.

Wegen Max Speec $A=G$ und $JA_z=f^q A_z\cap g^q A_z$ $\forall z\in G$ folgt dann $JA=f^q A\cap g^q A$; da A faktoriell bzw. fastfaktoriell ist, folgt für $q=1$ bzw. für großes q, daß JA Hauptideal ist. Nun liegt R_2 dicht in A im Sinne von 4.6: Denn ist $\sum_1^k a_i b_i\in A^*$ mit $a_i, b_i\in A$, $a_i=\sum_j a_{ij}(z')^m$, so ist auch $\sum_1^k a_{i0} b_i\in A^*$ und $a_{i0}\in R_1\subset R_2$. Dann folgt mit Zusatz 4.6, daß es schon ein $h\in J$ gibt mit $JA=hA$. In unserem Fall war zwar nicht – wie in 4.6 gefordert wurde – der Ring A ein Ring von auf G analytischen Funktionen; doch das macht hier nichts. Ist nun $j\in J$, so ist $j\in h\hat{\mathcal{O}}_z^{Z\oplus k}$ $\forall z\in G$, woraus nach [C] schon $j\in h\mathcal{O}_z^{Z\oplus k}$ $\forall z\in G$ folgt. Wegen der Eigenschaft I für R_2 ist $j\in hR_2$ und damit $J=hR_2$.

Folgerung 4.15. *Sei $G\subset k^m$. Genau dann ist $R_m(G)$ noethersch bzw. zusätzlich faktoriell bzw. fastfaktoriell, wenn $R_{m+1}(G)$ noethersch bzw. zuzätzlich faktoriell bzw. fastfaktoriell ist.*

Insbesondere folgt aus 4.15, daß für kompakte $G\subset k^m$, die in einer eindimensionalen Graden enthalten sind, stets $R_m(G)$ noethersch und faktoriell ist (da nach [8], Satz 29 ja $R_1(G)$ Hauptidealring ist).

Wir wollen nun ein Analogon zu einem bekannten Theorem von Forster-Igusa-Remmert hier für Unterringe und beliebige Mengen G_i (die nicht notwendig k-analytische Räume zu sein brauchen, die aber in einem k^{n_i} einbettbar sind) geben:

Satz 4.16. *Seien R_i Ringe mit I bezüglich G_i. Ist $\varphi\colon R_1\to R_2$ ein k-Isomorphismus, so ist auch $\tilde{\varphi}\colon G_2\to G_1$, $x\to(g_1(x),\dots,g_{n_1}(x))$ ein Isomorphismus für gewisse Funktionen $g_i\in R_2$.*

Beweis. Wir können nach Definition der analytischen Ringe (s. [10]) die Menge G_1 derart als Teilmenge eines k^{n_1} auffassen, so daß R_1 die auf G_1 eingeschränkten Koordinatenfunktionen z_i des k^{n_1} enthält. Sei $g_i:=\varphi(z_i)$. Wir behaupten nun, daß für das in der Behauptung definierte

$\tilde{\varphi}$ schon $\tilde{\varphi}(G_2) \subset G_1$ ist: Denn sonst gäbe es $x \in G_2$ mit $(a_1, \ldots, a_{n_1}) \notin G_1$, wobei $a_i := g_i(x)$ ist. Dann wäre wegen der Eigenschaft I für R_1 schon $(z_i - a_i) R_1$ das Einheitsideal in R_1, woraus auch

$$(g_i - a_i) R_2 = \varphi((z_i - a_i) R_1) = \varphi(R_1) = R_2$$

folgte. Nun sollte aber nach Voraussetzung $x \in G_2$ eine gemeinsame Nullstelle von $g_i - a_i$ sein, was dem widerspricht.

Als nächstes zeigen wir, daß $\tilde{\varphi}$ injektiv ist: Wäre $g_i(x) = g_i(y) =: a_i$ für alle i mit $x \neq y$, $x, y \in G_2$, so hätte $(g_i - a_i)$ die beiden Nullstellen x und y. Nach Obigem ist $a := (a_1, \ldots, a_{n_1}) \in G_1$, somit ist $(z_i - a_i) R_1$ ein maximales Ideal in R_1 und damit $\varphi((z_i - a_i) R_1) = (g_i - a_i) R_2$ ein maximales Ideal in R_2; das kann dann aber keine zwei verschiedenen Nullstellen haben, da R_2 ein Ring mit I bezüglich G_2 ist.

Schließlich ist auch $\tilde{\varphi}$ surjektiv: Sei $a = (a_1, \ldots, a_{n_1}) \in G_1$. Wenn es kein $x \in G_2$ gäbe mit $\tilde{\varphi}(x) = a$, also mit $g_i(x) = a_i$, so wäre $(g_i - a_i) R_2$ das Einheitsideal in R_2 wegen der Eigenschaft I. Aber $(z_i - a_i) R_1$ ist wegen $a \in G_1$ ein echtes Ideal in R_1, woraus wegen $\varphi((z_i - a_i) R_1) = (g_i - a_i) R_2$ der Widerspruch folgt.

Wir zeigen nun, daß $\tilde{\varphi}$ auch einen Isomorphismus der Strukturgarben induziert. Zunächst induziert $\tilde{\varphi}$ eine kanonische Abbildung $R_1 \xrightarrow{\tilde{\tilde{\varphi}}} R(G_2)$ durch $r \mapsto r \circ \tilde{\varphi}$. Dabei werden die Funktionen $z_i - a_i$ für $a_i \in k$ nach $\varphi(z_i) - a_i$ abgebildet; $\tilde{\tilde{\varphi}}$ stimmt also mit φ auf dem Polynomring in R_1 überein. Sei jetzt $z'' \in G_2$, $z' \equiv (z'_1, \ldots, z'_{n_1}) := \tilde{\varphi}(z'')$ und seien $\mathfrak{m}'$ bzw. $\mathfrak{m}''$ die Ideale in R_1 bzw. R_2 aller in z' bzw. z'' verschwindenden Funktionen. Es ergibt sich $\mathfrak{m}'' = \varphi(\mathfrak{m}')$ und somit gibt es einen durch φ induzierten Isomorphismus $\varphi_{(z', z'')}$: $(R_1)_{z'} \to (R_2)_{z''}$. Nach [C] erhalten wir damit einen Isomorphismus $\hat{\varphi}_{(z', z'')}$: $\hat{\mathcal{O}}_{z'} \to \hat{\mathcal{O}}_{z''}$. Dann stimmt $\hat{\varphi}_{(z', z'')}$ mit dem durch $\tilde{\varphi}$ induzierten Homomorphismus $\hat{\tilde{\tilde{\varphi}}}_{(z', z'')}$: $\hat{\mathcal{O}}_{z'} \to \hat{\mathcal{O}}_{z''}$, $o \mapsto o \circ \tilde{\varphi}$ auf dem durch $(z_i - z'_i)$ erzeugten Polynomring in $\hat{\mathcal{O}}_{z'}$ überein; es folgt $\hat{\varphi}_{(z', z'')} = \hat{\tilde{\tilde{\varphi}}}_{(z', z'')}$ und somit ist $\hat{\tilde{\tilde{\varphi}}}_{(z', z'')}$ ein Isomorphismus. Damit ist der durch $\tilde{\varphi}$ induzierte Homomorphismus $\tilde{\tilde{\varphi}}_{(z', z'')}$: $\mathcal{O}_{z'} \to \mathcal{O}_{z''}$, $o \mapsto o \circ \tilde{\varphi}$ ein Isomorphismus.

Aus 4.16 folgt insbesondere, daß zwei Steinsche Mannigfaltigkeiten isomorph sind, wenn die vollen Ringe isomorph sind. Wir können jedoch sogar zeigen, daß schon ein treuflacher Ringhomomorphismus von Unterringen mit I im allgemeinen nur bei isomorphen Gebieten möglich ist:

Folgerung 4.16. *Seinen R_i Ringe mit* I *bezüglich G_i. Ist φ: $R_1 \to R_2$ ein k-Homomorphismus, so daß $\varphi(R_1)$ die Koordinatenfunktionen z_i aus R_2 enthält und so daß R_2 treuflach über R_1 wird, so ist schon $G_1 \cong G_2$.*

Beweis. Da R_2 treuflach über R_1 ist, kann φ nur injektiv sein ([13], 3.5.4). Somit ist $R_1 \cong \varphi(R_1) \subset R_2$. Nun ist $\varphi(R_1)$ ein Unterring von R_2, der den Polynomring enthält und über dem R_2 treuflach ist. Es ist somit für Ideale $J \subset \varphi(R_1)$ stets $JR_2 \cap \varphi(R_1) = J$. Daraus folgt sofort, daß auch $\varphi(R_1)$ die Eigenschaft I bezüglich G_2 erfüllt. Nach 4.16 folgt die Behauptung.

Die Forderung, daß $\varphi(R_1)$ die Koordinatenfunktionen enthält, ist übrigens notwendig: Sei $R_1 := \mathbb{C}[X]$, $R_2 := \mathbb{C}[X, Y]$. R_i erfüllt I bezüglich $\mathbb{C}^i$ und es ist R_2 treuflach über R_1 bezüglich der kanonischen Injektion $R_1 \hookrightarrow R_2$. Es ist jedoch nicht $\mathbb{C}^1 \cong \mathbb{C}^2$. Es sei hier erwähnt, daß man auch $\mathbb{C}[X]$ als Ring von holomorphen Funktionen auf $\mathbb{C}^2$ betrachten kann, so daß auch sämtliche Ideale die Eigenschaft (I) bezüglich $\mathbb{C}^2$ erfüllen. Nach unserer Definition (s. [10]) erfüllt er jedoch nicht I, da er nicht den Polynomring in zwei Variablen enthält. Das eben gebrachte Beispiel zeigt, daß diese Forderung schon sinnvoll ist.

Literatur

1. Fensch, W.: Reell-analytische Strukturen. Schriftenreihe des Math. Inst. d. Univ. Münster, 34 (1966).
2. Forster, O.: Zur Theorie der Steinschen Algebren und Moduln. Math. Z. **97**, 376–405 (1967).
3. Frisch, J.: Points des platitude d'un morphisme d'espaces analytiques complexes. Inventiones math. **4**, 118–138 (1967).
4. Grauert, H.: Analytische Faserungen über holomorph vollständigen Räumen. Math. Ann. **133**, 139–159 (1957).
5. Grauert, H., Remmert, R.: Analytische Stellenalgebren. Grundlehren der Mathematik, Bd. 176. Berlin-Heidelberg-New York: Springer 1971.
6. Gunning, R.C., Rossi, H.: Analytic functions of several complex variables. Englewood Cliffs, N.J.: Prentice Hall 1965.
7. Igusa, I.: On a property of the domain of regularity. Mem. Coll. Sci. Univ. Kyoto **27**, 95–97 (1952).
8. Langmann, K.: Ringe holomorpher Funktionen und endliche Idealverteilungen. Schriftenreihe des Math. Inst. d. Univ. Münster, Serie 2, Heft 3 (1971).
9. Langmann, K.: Konstruktionen globaler Moduln und Anwendungen. Math. Z. **127**, 235–255 (1972).
10. Langmann, K.: Globale Moduln. Math. Z. **124**, 141–168 (1972).
11. Malgrange, B.: Ideals of differentiable functions. Oxford University Press 1966.
12. Nagata, M.: Local rings. New York: Interscience 1962.
13. Nastold, H.-J.: Neuere Methoden in der lokalen Algebra. Münster 1966 (Manuskript).
14. Remmert, R.: Habilitationsschrift, Münster 1957.
15. Storch, U.: Fastfaktorielle Ringe. Schriftenreihe des Math. Inst. d. Univ. Münster, Heft 36 (1967).

Dr. Klaus Langmann
Mathematisches Institut, Abt. IV
D-4400 Münster, Roxeler Straße 64
Bundesrepublik Deutschland

(Eingegangen am 24. Januar 1972)

Math. Z. 128, 187 (1972)

Correction to

Finite Rings with a Specified Group of Units

Math. Z. **126**, 51–58 (1972)

Ian Stewart

Dr. G.L.C. Bond of the University of Reading has pointed out a mistake in my above paper. In the proof of Lemma 3.2 I made the inadvertent assumption that all rings are commutative, and in fact the lemma fails for the ring of 2×2 upper triangular matrices over $\mathbb{Z}_p$, p prime.

Fortunately the error can be confined to Lemma 3.2, as was also pointed out by Dr. Bond. The only place where 3.2 is used is in the proof of 3.3, and here its use can be avoided as follows: we have $R/J \cong T_1 \oplus \cdots \oplus T_n$ where each $T_i \cong \mathbb{Z}_2$, and we pick elements $e_j \in R$ which map by the canonical homomorphism $R \to R/J$ to the identity of T_j. We find that $e_n = x + a$ where $x \in J$ and $a \in \operatorname{Ann}(J)$. At this point we wish to assert that $e_n^\lambda \in \operatorname{Ann}(J)$ for some integer $\lambda > 0$. In the paper we invoked 3.2; instead we show that $\lambda = c-1$ will do, where c is the index of nilpotency of J. For then

$$\begin{aligned} e_n^{c-1} &= (x+a)^{c-1} \\ &= x^{c-1} + z \end{aligned}$$

where $z \in \operatorname{Ann}(J)$ (which is an ideal of R by 3.1). Thus $e_n^{c-1} J = (x^{c-1} + z) J = 0$ since $J^c = 0$, and similarly $J e_n^{c-1} = 0$. Therefore $e_n^{c-1} \in \operatorname{Ann}(J)$. The proof of 3.3 now carries through with λ replaced by $c-1$ (or by 1 when $c=1$) with no further changes.

Thus all the results can be resuscitated, with the exception of Lemma 3.2.

Dr. Ian Stewart
Mathematics Institute
University of Warwick
Coventry CV 4 7 AL
England

(Received June 26, 1972)

Math. Z. 128, 189–194 (1972)

Boundary Values and Mapping Properties of H^P Functions

Lowell J. Hansen

1. Introduction

Let Δ denote the open unit disc $\{|z|<1\}$ and let Γ denote its circumference. A classical theorem of Lindelöf [9] states that if f is analytic and bounded on Δ and has the asymptotic value w at a point $P\in\Gamma$, then f has the non-tangential limit w at P. Lehto and Virtanen extended Lindelöf's theorem to the more general case where f is meromorphic and normal [8]. A generalization of a different nature was given by Gehring and Lohwater [6]: *Let $f=u+iv$ be analytic and bounded in Δ. If u and v have the asymptotic values a and b respectively at $P\in\Gamma$, then f has the non-tangential limit $a+ib$ at P.* One might guess from Lehto and Virtanen's success with the Lindelöf Theorem that the Gehring-Lohwater Theorem would remain valid if "f bounded" were replaced by "f normal". Bagemihl [2] has shown by example that this is not the case. (McMillan [10] has also studied the unbounded case.) Our first theorem gives some less restrictive conditions than "f bounded" under which the conclusion of the Gehring-Lohwater Theorem still holds.

Theorem 1. *Let $f=u+iv$ be analytic on Δ and suppose that u approaches a along α and v approaches b along β, where α and β are arcs at $P\in\Gamma$. If either*

(I) *$\alpha\cap\beta$ does not cluster at P and f belongs to the Hardy class H^2, or*

(II) *$\alpha\cap\beta$ clusters at P,*

then f has the asymptotic value $a+ib$ at P. If, in addition, f is a normal function, then f has the non-tangential limit $a+ib$ at P.

We remark that the result of part (I) is best possible in the sense that the condition "$f\in H^2$" cannot be replaced by "$f\in H^p$ for $0<p<2$". By way of example, let $f=u+iv$ be analytic on Δ and satisfy $f^2(z)=i\left(\dfrac{1+z}{1-z}\right)$. Then f is normal, $f\in H^p$ for $0<p<2$, $u\equiv 0$ on the upper half of Γ, $v\equiv 0$ on the lower half of Γ, and yet $f(z)\to\infty$ as $z\to 1$.

Our second theorem relates the image of an H^p function to the set of its radial limits. This theorem will be used in the proof of part (I) of Theorem 1.

Theorem 2. *Let ϕ be fixed, $0<\phi<2\pi$, and let S_ϕ denote the sector $\{re^{i\theta}\colon r>0$ and $\phi<\theta<2\pi\}$. If $f\in H^{\pi/\phi}$ and satisfies $f(e^{i\theta})\in S_\phi$ a.e. on Γ, then $f(\Delta)\subset S_\phi$. The conclusion is also valid if S_ϕ is replaced by any other sector of the same size: $\{az+b\colon z\in S_\phi\}$, where $a\neq 0$. The hypothesis $f\in H^{\pi/\phi}$ cannot be weakened.*

In the special case where $\phi=\pi$, Theorem 2 implies that $f(\Delta)$ is contained in any half-plane containing almost all $f(e^{i\theta})$ if $f\in H^1$ (a simpler proof of this special case could be given using the Poisson integral representation of H^1 functions). This gives the corollary:

Corollary. *If $f\in H^1$, then $f(\Delta)$ is contained in the convex hull of the radial limits of f. The hypothesis $f\in H^1$ cannot be weakened.*

Theorem 2 is similar in character to the theorem of Neuwirth and Newman [11] which states that if $f\in H^{\frac{1}{2}}$ and $f(e^{i\theta})\geqq 0$ a.e. on Γ, then f is a constant. A slight alteration of the proof of Theorem 2 gives the following generalization of the Neuwirth-Newman Theorem:

Theorem 3. *Let $M>0$ be fixed and put $R=\{z\colon \operatorname{Re} z>-M$ and $|\operatorname{Im} z|<M\}$. If $f\in H^{\frac{1}{2}}$ and $f(e^{i\theta})\in R$ a.e. on Γ, then $f(\Delta)\subset R$ and $f\in H^p$ for $0<p<\infty$. The hypothesis $f\in H^{\frac{1}{2}}$ cannot be weakened.*

The proof of Theorem 2 (or 3) depends on knowing the Hardy classes to which a function mapping Δ conformally onto the complement of $\bar{S}_\phi$ (or $\bar{R}$) does not belong. In general, if $f(e^{i\theta})\in D$ a.e. and if one knows the Hardy number (see [7]) of each of the components of the complement of $\bar{D}$, a similar theorem can be proved.

2. Some Definitions and Notation

We shall let C denote the finite complex plane. If $K\subset C$, then ∂K and $\bar{K}$ will denote, respectively, the boundary of K and the closure of K, both taken with respect to the extended plane.

If P is a boundary point of the region $D\subset C$, and arc at P is a curve $\gamma\subset D$ such that $\gamma\cup\{P\}$ is a Jordan arc. A complex-valued function g defined on Δ is said to have the non-tangential limit w at $e^{i\theta}$ if, for each $0<\delta<\pi/2$, $g(z)\to w$ as $z\to e^{i\theta}$ in the Stolz angle $\{z\in\Delta\colon |\arg(1-ze^{-i\theta})|\leqq\delta\}$.

Let $D\subset C$ be a region and let $0<p<\infty$. If f is analytic on D, then f is said to belong to the Hardy class $H^p(D)$ if the subharmonic function $|f|^p$ possesses a harmonic majorant. If $D=\Delta$, $H^p(D)$ as just defined coincides with the usual Hardy class H^p. (See [5] for a presentation of the theory of H^p spaces.) If $f\in H^p$, then the radial limit, $f(e^{i\theta})$, exists for almost all θ. If g is analytic on Δ, g is said to be an inner function if $|g(z)|\leqq 1$ for $z\in\Delta$ and $|g(e^{i\theta})|=1$ a.e.

3. Proof of Theorem 2

Since the conclusion readily follows otherwise, we suppose that f is non-constant and $f(\Delta)\cap(C\setminus\overline{S_\Phi})$ is non-empty. Let D be a component of $f^{-1}(C\setminus\overline{S_\Phi})$. We claim that D is simply-connected. If not, then there would exist a rectifiable Jordan curve γ in D whose interior contains a point $\zeta\in\Delta\cap\partial D$. Then $f(\zeta)\in C\cap\partial S_\Phi$ and $f(\gamma)\subset C\setminus\overline{S_\Phi}$, and so $f(\zeta)$ lies in the unbounded component of $C\setminus f(\gamma)$. Hence the winding number of $f(\gamma)$ with respect to $f(\zeta)$ is zero (see [1, p. 116]). The Argument Principle [1, p. 151] then implies that f never assumes the value $f(\zeta)$ inside γ, which is a contradiction. Therefore, D is simply-connected.

We show next that the linear Lebesgue measure of $\Gamma\cap\partial D$ is zero. Let E be a subset of Γ with measure 2π such that the radial limit of f exists at each point of E and $f(E)\subset S_\Phi$. If $E_1=E\cap\partial D$, then $\Gamma\cap\partial D$ is the union of E_1 and a set of measure zero. So it suffices to show that E_1 is countable. The existence of the radial limit (and hence, the non-tangential limit) of f on E_1 implies that for each $\zeta\in E_1$ there exists a positive integer $n=n(\zeta)$ so that if

$$D_\zeta=\{(1-1/n)\leqq|z|<1 \text{ and } |\arg(1-z\bar{\zeta})|\leqq\pi/4\},$$

then $f(D_\zeta)\subset S_\Phi$. If there exist distinct points $\zeta_1,\zeta_2,\zeta_3\in E_1$ with $n(\zeta_1)=n(\zeta_2)=n(\zeta_3)$ and $D_{\zeta_i}\cap D_{\zeta_j}\neq\varnothing$ $(i\neq j)$, then D would be contained in one of the three components of $\Delta\setminus\left(\bigcup_1^3 D_{\zeta_i}\right)$ and one of the ζ_i would not belong to ∂D. Therefore, for each fixed N, the set $\{\zeta\in E_1\colon n(\zeta)=N\}$ is finite and so E_1 is countable. Hence $\Gamma\cap\partial D$ has measure zero.

Since the Lebesgue measure of $\Gamma\cap\partial D$ is zero, its harmonic measure with respect to Δ is zero. Therefore, the harmonic measure of $\Gamma\cap\partial D$ with respect to D is also zero, since $D\subset\Delta$ [12, p. 69]. Let φ denote a univalent function mapping Δ conformally onto D. Since the radial limits of a non-constant bounded analytic function on a set of positive measure in Γ cannot be contained in a set of harmonic measure zero [12, p. 209], we must have $\varphi(e^{i\theta})\in\Delta\cap\partial D$ a.e. Therefore, $f[\varphi(e^{i\theta})]\in C\cap\partial S_\Phi$ a.e.

Let F be a univalent function which maps Δ conformally onto $C\setminus\overline{S_\Phi}$ and satisfies $F(0)=f[\varphi(0)]$. Then F does not belong to the Hardy class $H^{\pi/\Phi}$ [3]. If we let $\psi=F^{-1}[f\circ\varphi]$, then $|\psi|\leqq 1$ and $|\psi(e^{i\theta})|=1$ a.e. Therefore ψ is an inner function. It follows from a lemma of Nordgren [13, p. 442] that if ψ is inner, then F and $F\circ\psi$ belong to exactly the same Hardy classes. Hence $f\circ\varphi=F\circ\psi$ does not belong to $H^{\pi/\Phi}$ either. Since φ is univalent, the restriction of f to D does not belong to $H^{\pi/\Phi}(D)$, and so $f\notin H^{\pi/\Phi}$. From this contradiction, the conclusion of Theorem 2 follows.

To prove Theorem 3, we replace S_Φ in the above proof by R. In this case, the function mapping Δ conformally onto $C\setminus\overline{R}$ does not belong to $H^{\frac{1}{2}}$ and so, if $f(\Delta)\cap(C\setminus\bar{R})$ is non-empty, we reach the contradiction that $f\notin H^{\frac{1}{2}}$. We conclude that $f(\Delta)\subset R$. Finally, any analytic function on Δ whose image is contained in an infinite strip belongs to each Hardy class H^p, $0<p<\infty$ [3, p. 473].

4. Proof of Theorem 1

As was mentioned in Section 1, any normal function which has an asymptotic value at a boundary point P automatically has that asymptotic value as a non-tangential limit at P. Therefore it suffices to show that either condition (I) or (II) implies that $a+ib$ is an asymptotic value of f at P. By considering the function $f(z\bar{P})-(a+ib)$, we may assume that $P=1$ and $a=b=0$. Since the result is trivial otherwise, we assume that f is non-constant.

Proof of Part (I). Since $\alpha\cap\beta$ does not cluster at $\zeta=1$, we may suppose that $\alpha\cap\beta\cap\Delta$ consists of a single point. We let Ω denote the interior of the Jordan curve $\alpha\cup\beta$. By a conformal mapping of Δ onto Ω, we may assume that $\Omega=\Delta$.

Since u and v are bounded on α and β respectively, there is some "cross"

$$X=\{|\operatorname{Im} z|<M\}\cup\{|\operatorname{Re} z|<M\}$$

to which $f(e^{i\theta})$ belongs for all points $e^{i\theta}\in\Gamma\setminus\{1\}$. Applications of Theorem 2 to each of the components of $C\setminus\bar{X}$ allow us to conclude that $f(\Delta)\subset X$.

McMillan [10, Theorem 2, p. 53] has shown that if u and v each have the asymptotic value zero at $\zeta=1$ but f does not, then there exists an arc σ at $\zeta=1$ on which $f'(z)\neq 0$ and which f maps in a one-to-one manner onto some ray $L_1=\{w_1+\rho e^{iB}\colon \rho\geqq 0\}$. Therefore ∞ is an asymptotic value of f at $\zeta=1$. By a corollary of the Gross-Iversen Theorem [4, Theorem GI′, p. 462], ∞ is a cluster value of f on each arc at $\zeta=1$. Hence f is not bounded on any arc at $\zeta=1$.

Since $L_1\subset f(\Delta)\subset X$, B must be an integer multiple of $\pi/2$. We consider only the case where $B=0$, since each of the other cases can be treated similarly. By omitting part of the arc σ if necessary, we may suppose that $\operatorname{Re} w_1\geqq M+2$. Let L denote the line with equation $x=M+1$, and put $H=\{z\colon \operatorname{Re} z\geqq M+1\}$. Since $f(\Delta)\subset X$ and $u\to 0$ on α, the sets $L\cap f(\Delta)$ and $f(\alpha)\cap H$ are bounded. We appeal to a lemma of McMillan [10, Lemma 2, p. 55] which implies that either

(i) there exists an arc at $\zeta=1$ which is mapped by f into $L\cap f(\Delta)$ in a one-to-one manner, or

(ii) there exists an arc at $\zeta=1$ which is mapped by f into $(f(\alpha)\cap H)\cup(L\cap f(\Delta))$.

In either case (i) or (ii), f would be bounded on an arc at $\zeta=1$. Thus we have seen that a contradiction arises if zero is not an asymptotic value of f at $\zeta=1$, and the proof of part (I) is completed.

The proof of part (II) rests on the following lemma.

Lemma. *Let $f=u+iv$ be non-constant and analytic at each point of $\{|z|\leqq 1\}$. Suppose that*

(1) $|u(e^{i\theta})|<\varepsilon$ *for* $0\leqq\theta\leqq\pi$,

(2) $|v(e^{i\theta})|<\varepsilon$ *for* $-\pi\leqq\theta\leqq 0$, *and*

(3) $f'(z)\neq 0$ *whenever* $|u(z)|=\varepsilon$ *or* $|v(z)|=\varepsilon$.

Then there exists a Jordan arc in $\Delta\cup\{-1,1\}$ with endpoints -1 and 1 on which $|u|<\varepsilon$ and $|v|<\varepsilon$.

Proof. Let $\mathcal{S}_\varepsilon=\{|\operatorname{Re} w|<\varepsilon\}\cap\{|\operatorname{Im} w|<\varepsilon\}$. Then since $f(1)\in\mathcal{S}_\varepsilon$, there exists a neighborhood of 1 which is mapped by f into $\mathcal{S}_\varepsilon$. We let D denote the unique component of $\Delta\cap f^{-1}(\mathcal{S}_\varepsilon)$ which possesses the point $\zeta=1$ as a boundary point. If K is a component of $\Delta\cap\partial D$, then $f(K)\subset\partial\mathcal{S}_\varepsilon$. Therefore, since $f'\neq 0$ on K, f maps K in a one-to-one manner onto some open segment of a side of $\mathcal{S}_\varepsilon$. Furthermore, $\bar{K}\setminus K$ consists of two points of Γ. Since these two points are mapped by f to the same side of $\mathcal{S}_\varepsilon$, they must belong to the same semi-circle $\Gamma\cap\{\operatorname{Im} z>0\}$ or $\Gamma\cap\{\operatorname{Im} z<0\}$.

We define $\theta_1=\sup\{\theta\colon 0\leqq\theta\leqq\pi$ and $e^{i\theta}\in\partial D\}$. If $\theta_1<\pi$, then $\exp(i\theta_1)$ is an endpoint of at least one component of $\Delta\cap\partial D$. Suppose first that there is only one such component, K, with $\exp(i\theta_1)$ as an endpoint. If $\exp(i\theta_2)$ is the other endpoint, then $0<\theta_2<\theta_1<\pi$. From the definition of θ_1, we must then have $(\Gamma\cap\partial D)\subset\{e^{i\theta}\colon\theta_2\leqq\theta\leqq\theta_1\}$. This is not possible, since $1\in\partial D$. Secondly, suppose that there are two components of $\Delta\cap\partial D$, K and K', which have $\exp(i\theta_1)$ as an endpoint. We denote the other endpoints by $\exp(i\theta_2)$ and $\exp(i\theta_2')$, respectively. Without loss of generality, we may assume that $\theta_2\leqq\theta_2'$. Then we must have the inequality $0<\theta_2\leqq\theta_2'<\theta_1<\pi$. Since K and K' are contained in ∂D, $\Gamma\cap\partial D$ must be contained in $\{e^{i\theta}\colon\theta_2\leqq\theta\leqq\theta_2'$ or $\theta=\theta_1\}$. Again, this is not possible since $1\in\partial D$. Since there can be no more than two components of $\Delta\cap\partial D$ having $\exp(i\theta_1)$ as an endpoint, we cannot have $\theta_1<\pi$. Therefore $\theta_1=\pi$ and $-1\in\partial D$. We conclude that -1 and 1 are accessible boundary points of D since $f(-1)$ and $f(1)\in\mathcal{S}_\varepsilon$. Hence -1 and 1 may be connected by a Jordan arc, $\mathscr{C}$, in $D\cup\{-1,1\}$. Therefore $f(\mathscr{C})\subset\mathcal{S}_\varepsilon$, which was to be proved.

Proof of Part (II). We suppose that α and β are parametrized on $[0,1]$. Let $G=\{t\in[0,1]\colon\alpha(t)\notin\beta\}$ and $A=\{a\in[0,1]\colon(a,a')$ is a component of $G\}$. If G is bounded away from 1, then $f\to 0$ on α. We therefore assume that G is not bounded away from 1. Then, since $\alpha\cap\beta$ clusters at $\zeta=1$, $A\cap(1-1/n,1)$ is infinite for each integer n. If (a,a') is a component of G, then

$\alpha(a)$, $\alpha(a') \in \beta$ and so there exist unique points d, $d' \in [0, 1)$ such that $\beta(d) = \alpha(a)$ and $\beta(d') = \alpha(a')$. Let I_a denote the closed interval with end-points d and d'. Then since $u[\alpha(t)] \to 0$ and $v[\beta(t)] \to 0$ as $t \to 1$, we can find for each $a \in A$ an $\varepsilon_a > 0$ such that

(i) $\varepsilon_a \to 0$ as $a \to 1$;
(ii) $|u[\alpha(t)]| < \varepsilon_a$ for $a \leqq t \leqq a'$;
(iii) $|v[\beta(t)]| < \varepsilon_a$ for $t \in I_a$; and,
(iv) $f'(z) \neq 0$ if $|u(z)| = \varepsilon_a$ or $|v(z)| = \varepsilon_a$.

If D_a denotes the Jordan region with $\partial D_a = \alpha([a, a']) \cup \beta(I_a)$, then $\bar{D}_a \subset \Delta$. We now map D_a conformally onto Δ in such a way that $\alpha([a, a'])$ and $\beta(I_a)$ correspond to $\{e^{i\theta}: 0 \leqq \theta \leqq \pi\}$ and $\{e^{i\theta}: -\pi \leqq \theta \leqq 0\}$, respectively. The preceding lemma then implies the existence of a curve $\mathscr{C}_a$ in $D_a \cup \{\alpha(a), \alpha(a')\}$ which connects $\alpha(a)$ to $\alpha(a')$ and on which $|u| < \varepsilon_a$ and $|v| < \varepsilon_a$. We now define the curve $\mathscr{C}$ on $[0, 1]$ so that $\mathscr{C}(t) = \alpha(t)$ if $t \in [0, 1] \setminus G$ and $\mathscr{C}([a, a']) = \mathscr{C}_a$ if (a, a') is a component of G. Then $f[\mathscr{C}(t)] \to 0$ as $t \to 1$ and so f has the asymptotic value zero at $\zeta = 1$.

References

1. Ahlfors, L.V.: Complex analysis, 2. ed. New York: McGraw-Hill 1966.
2. Bagemihl, F.: The Lindelöf theorem and the real and imaginary parts of normal functions. Michigan Math. J. **9**, 15–20 (1962).
3. Cargo, G.T.: Some geometric aspects of functions of Hardy class H^p. J. Math. Analysis Appl. **7**, 471–474 (1963).
4. Doob, J.L.: One-sided cluster-value theorems. Proc. London Math. Soc. (3) **13**, 461–470 (1963).
5. Duren, P.L.: Theory of H^p spaces. New York: Academic Press 1970.
6. Gehring, F.W., Lohwater, A.J.: On the Lindelöf theorem. Math. Nachr. **19**, 165–170 (1958).
7. Hansen, L.J.: Hardy classes and ranges of functions. Michigan Math. J. **17**, 235–248 (1970).
8. Lehto, O., Virtanen, K.I.: Boundary behavior and normal meromorphic functions. Acta Math. **97**, 47–65 (1957).
9. Lindelöf, E.: Sur un principe générale de l'analyse et ses applications à la théorie de la représentation conforme. Acta Soc. Sci. Fennicae **46**, no. 4, 1–35 (1915).
10. McMillan, J.E.: Minimum convexity of a holomorphic function, II. Michigan Math. J. **16**, 53–58 (1969).
11. Neuwirth, J., Newman, D.J.: Positive $H^{\frac{1}{2}}$ functions are constants. Proc. Amer. Math. Soc. **18**, 958 (1967).
12. Nevanlinna, R.: Eindeutige analytische Funktionen, 2. Aufl. Berlin-Göttingen-Heidelberg: Springer 1953.
13. Nordgren, E.A.: Composition operators. Canadian J. Math. **20**, 442–449 (1968).

Dr. L.J. Hansen
Department of Mathematics
Wayne State University
Detroit, Michigan 48202
USA

(Received February 21, 1972)

Math. Z. 128, 195–197 (1972)

A Product Theorem in Dimension

Manuel Castellet

All spaces will be paracompact Hausdorff spaces and all sheaves will be sheaves of abelian groups.

We ask for the weakest possible conditions on a pair of paracompact spaces X and Y under which the relation

$$\dim(X \times Y) = \dim X + \dim Y$$

holds.

Here, dim denotes the covering dimension ($\dim X \leqq n$ if and only if every locally finite open covering of X has an open refinement of order $\leqq n+1$).

If $\mathscr{F}$ is a sheaf on X, let $H^i(X, \mathscr{F})$ denote the cohomology groups of X with coefficients in $\mathscr{F}$ and supports the family of all closed sets in X. Let us denote by $\mathbb{Z}$ the constant sheaf with stalk the group $\mathbb{Z}$ of integers.

The following proposition is an easy consequence of well known facts:

Proposition. *The following conditions are equivalent:*

a) $\dim X \leqq n$,

b) *for all* $i > n$ *and every sheaf* $\mathscr{F}$ *on* X, $H^i(X, \mathscr{F}) = 0$,

c) *for all* $i \geqq n$, *every sheaf* $\mathscr{F}$ *on* X *and every closed set* C *in* X, $H^i(X, \mathscr{F}) \to H^i(C, \mathscr{F})$ *is an epimorphism,*

d) *for all* $i > n$ *and every open set* U *in* X, $H^i(X, \mathbb{Z}_U) = 0$.

Proof. a) $\Rightarrow$ b) Since X is paracompact, every covering of X has a locally finite open refinement, and, since $\dim X \leqq n$, an open refinement of order $\leqq n+1$. Hence, the open coverings of order $\leqq n+1$ provide us with a cofinal set and one can take these coverings to form the limit in the Čech cohomology.

b) $\Rightarrow$ c) From the cohomology sequence derived from the exact sequence of sheaves

$$0 \to \mathscr{F}_{X-C} \to \mathscr{F} \to \mathscr{F}_C \to 0$$

(C closed in X), since, if A is locally closed $H^i_\Phi(X, \mathscr{F}_A) = H^i_{\Phi|A}(A, \mathscr{F})$.

c) $\Rightarrow$ a) Take $\mathscr{F} = \mathbb{Z}$ and apply the Hopf's extension theorem for cohomology, [1].

b) $\Rightarrow$ d) Trivial.

d) $\Rightarrow$ a) Apply the same argument as in b) $\Rightarrow$ c) $\Rightarrow$ a).

Now we study the dimension of a product.

Let X be a paracompact space and Y a compact space of dimensions n and m respectively.

Let $\pi: X \times Y \to X$ be the projection and, for every q and every sheaf $\mathscr{F}$ on $X \times Y$, let us denote by $\mathscr{H}^q(\pi, \mathscr{F})$ the Leray's sheaf on X defined by π. $\mathscr{H}^q(\pi, \mathscr{F})$ is defined by the following presheaf: for all open $U \subset X$

$$\mathscr{H}^q(\pi, \mathscr{F})(U) = H^q(U \times Y, \mathscr{F} | U \times Y)$$

with the restriction homomorphisms

$$H^q(W \times Y, \mathscr{F} | W \times Y) \to H^q(U \times Y, \mathscr{F} | U \times Y) \quad \text{if } U \subset W.$$

Therefore, for all $x \in X$, we have

$$\mathscr{H}^q(\pi, \mathscr{F})(x) = \varinjlim_{U \ni x} H^q(U \times Y, \mathscr{F} | U \times Y) = H^q(x \times Y, \mathscr{F} | x \times Y),$$

because the compactness of Y. (When U runs over the family of neighborhoods of x, $U \times Y$ runs over a basis of neighborhoods of $x \times Y$, and one can apply 4.11.1 of [2] because $X \times Y$ is paracompact and $x \times Y$ is closed in $X \times Y$.)

Then, by a theorem of Leray, there is a spectral sequence E such that

$$E_2^{pq} = H^p(X, \mathscr{H}^q(\pi, \mathscr{F}))$$

and E converges to $H^*(X \times Y, \mathscr{F})$ with a suitable filtration. But $H^p(X, \mathscr{A}) = 0$ for all $p > n$ and every sheaf $\mathscr{A}$ on X, and $H^q(x \times Y, \mathscr{B}) = 0$ for all $q > m$ and every sheaf $\mathscr{B}$ on Y. Hence $E_2^{pq} = 0$ for $p > n$ or $q > m$, and $E_2^{nm} = H^n(X, \mathscr{H}^m(\pi, \mathscr{F}))$. Therefore

$$H^{n+m}(X \times Y, \mathscr{F}) = H^n(X, \mathscr{H}^m(\pi, \mathscr{F})),$$

and

$$\dim(X \times Y) \leqq \dim X + \dim Y.$$

Now we study the inequality

$$\dim(X \times Y) \geqq \dim X + \dim Y.$$

In order that $\dim(X \times Y) = \dim X + \dim Y$, it is necessary and sufficient that there exists a sheaf $\mathscr{F}$ on $X \times Y$ for which $H^{n+m}(X \times Y, \mathscr{F}) \neq 0$.

First, let us note that, for all open sets $U \subset X$, the sheaf $\mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y})$, which induces 0 on $X - U$, is a constant sheaf on U.

In fact, if $x \in X$, we have

$$\mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y})(x) = H^m(x \times Y, \mathbb{Z}_{(x \cap U) \times Y}) \cong \begin{cases} 0 & \text{if } x \notin U \\ H^m(Y, \mathbb{Z}) & \text{if } x \in U. \end{cases}$$

Moreover, if F is a closed set in X, contained in U, and π': $F \times Y \to F$ is the projection, then

$$\mathscr{H}^m(\pi', \mathbb{Z}_{F \times Y}) \cong \mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y})|F.$$

Thus, ([3]), for every closed set F in X such that $F \subset U$, $\mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y})|F$ is a constant sheaf. Then, the assertion follows from the normality of X because X is paracompact.

We now formulate

Theorem. *Let X be a paracompact space and Y a compact space of finite dimension. Assume that there exists an epimorphism $H^{\dim Y}(Y, \mathbb{Z}) \to \mathbb{Z}$. Then* $\dim(X \times Y) = \dim X + \dim Y$.

(For example, if $m = \dim Y$, every orientable compact topological manifold, or every m-sphere with holes, verify the hypothesis on Y.)

Proof. Set $n = \dim X$, $m = \dim Y$.

By the proposition we know that there exists an open set $V \subset X$ such that $H^n(X, \mathbb{Z}_V) \neq 0$. By the remark above there exists, for every open set U in X, an epimorphism of sheaves on X

$$\mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y}) \to \mathbb{Z}_U.$$

Then, from the exact sequence

$$0 \to \mathscr{A} \to \mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y}) \to \mathbb{Z}_U \to 0$$

we obtain an exact sequence in cohomology

$$\ldots H^n(X, \mathscr{H}^m(\pi, \mathbb{Z}_{U \times Y})) \to H^n(X, \mathbb{Z}_U) \to H^{n+1}(X, \mathscr{A}) \to \cdots.$$

But $H^{n+1}(X, \mathscr{A}) = 0$ because $\dim X = n$, and $H^n(X, \mathbb{Z}_V) \neq 0$; therefore

$$0 \neq H^n(X, \mathscr{H}^m(\pi, \mathbb{Z}_{V \times Y})) = H^{n+m}(X \times Y, \mathbb{Z}_{V \times Y}).$$

Thus, $\dim(X \times Y) \geqq n + m$. This completes the proof.

References

1. Dowker, C.H.: Mapping theorems for non-compact spaces. Amer. J. Math. **69**, 200–242 (1947).
2. Godement, R.: Topologie algébrique et théorie des faisceaux. Paris: Hermann 1958.
3. Cartan, H.: Cohomologie des groupes, suite spectral, faisceaux. Séminaire École Normale Superieure, Paris 1950–1951.

Prof. Manuel Castellet
Eidgenössische Technische Hochschule
CH-8006 Zürich
Schweiz

(Received May 3, 1972)

Math. Z. 128, 199–205 (1972)

Über Regularitätseigenschaften der Potenzräume eines Schrödinger-Operators mit singulärem Potential

Jürgen Witte

Problemstellung

Im Hilbertraum $H=\{u\,|\,u=u(x),\ \int_{R_n}|u(x)|^2\,dx<\infty\}$ mit dem üblichen Skalarprodukt wird der Differentialoperator

$$A\,u=\sum_{j=1}^{n}\left(i\frac{\partial}{\partial x_j}+b_j\right)^2 u+(q+e_0\,\|x\|^{-\alpha})\,u \qquad (n\geqq 3) \tag{1}$$

im Teilraum $\mathfrak{A}_1=\{u\,|\,u\in M_\alpha\cap H,\ A\,u\in H\}$ betrachtet. Mit $\|x\|^2=x_1^2+x_2^2+\cdots+x_p^2$ $(p\in\mathbb{N},\ 3\leqq p\leqq n)$ und $R_n^*=R_n-\{x\,|\,\|x\|=0\}$ ist hierbei

$M_\alpha=\{u\,|\,u\in C^2(R_n^*)\cap C^1(R_n)$, es gibt ein $R_0>0$ so, daß für $\|x\|<R_0$
$u_{x_j}=u_{1j}+u_{2j}\dfrac{\partial}{\partial x_j}(\|x\|^{2-\alpha})$ $(j=1,\dots,n)$ mit Funktionen $u_{1j},u_{2j}\in C^1(\|x\|<R_0)$ gilt$\}$,

falls $\alpha\in(0,1)$ ist, und für $\alpha\in[1,\frac{3}{2})$

$M_\alpha=\{u\,|\,u\in C^2(R_n^*)\cap C^0(R_n)$, es gibt ein $R_0>0$ so, daß für $0<\|x\|<R_0$

$$u_{x_j}=u_{1j}+u_{2j}\frac{\partial}{\partial x_j}(\|x\|^{2-\alpha}),$$

$$u_{x_jx_k}=u_{1jk}+u_{2jk}\|x\|^{-\alpha}+u_{3jk}\frac{\partial^2}{\partial x_j\,\partial x_\kappa}(\|x\|^{2-\alpha}) \qquad (j,k=1,2,\dots,n)$$

mit Funktionen $u_{1j},\dots,u_{3jk}\in C^0(\|x\|<R_0)$ gilt$\}$.

Über die Koeffizienten des Schrödinger-Operators sollen folgende Voraussetzungen gemacht werden:

$b_j(x)\in C^{1+\delta}(R_n)$, $0<\delta<\gamma$ [1], $j=1,2,\dots,n$, $q(x)\in C^\delta(R_n)$, $0<\delta<\gamma$,
$q(x)\geqq -a_0|x|^2-b_0(x\in R_n)$, $a_0>0$, $b_0>0$, e_0 konstant, $\alpha\in(0,\frac{3}{2})$.

[1] $f(x)\in C^{l+\delta}(R_n)$, $0<\delta<\gamma$, mit ganzzahligem $l\geqq 0$ und reellem $\gamma\in(0,1)$ bedeutet, daß die l-ten partiellen Ableitungen von $f(x)$ im R_n stetig und lokal hölderstetig mit einem Hölderexponenten $\delta\,(0<\delta<\gamma)$ sind.

Nach Ikebe-Kato [2] ist A in $\mathfrak{A}_1$ wesentlich selbstadjungiert. Ist B in $\mathfrak{B}$ die Abschließung von A in $\mathfrak{A}_1$, so interessieren Eigenschaften der Potenzräume $\mathfrak{B}^{(k)}=\{u \mid u\in\mathfrak{B},\, Bu\in\mathfrak{B}^{(k-1)}\}$, $B^k u=B(B^{k-1}u)\,(k=2,3,4,\ldots)$, $\mathfrak{B}^{(1)}=\mathfrak{B}$, $B^1 u=Bu$, $\mathfrak{B}^{(0)}=H$, $B^0 u=u$. Es gilt der folgende

Satz. *Es sei* $[n/2]=\nu$ *und* $u\in\mathfrak{B}^{(\nu+m)}$ *mit festem natürlichen* $m\geqq 3$. *Dann gilt* $u\in\mathfrak{A}_1^{(m-2)}$.

Das Ergebnis kann so gedeutet werden, daß Elemente $u\in H$, auf die der Abschließungsoperator B hinreichend oft anwendbar ist, bereits gewisse Regularitätseigenschaften besitzen. Da die Eigenfunktionen φ und Eigenpakete Φ_λ des selbstadjungierten Operators B in $\mathfrak{B}$ in $\mathfrak{B}^{(\infty)}=\bigcap_{k=1}^{\infty}\mathfrak{B}^{(k)}$ liegen, folgt nach dem Satz $\varphi, \Phi_\lambda\in\mathfrak{A}_1^{(\infty)}$. Dies bedeutet, daß die Spektralzerlegung von B in $\mathfrak{B}$ im wesentlichen in $\mathfrak{A}_1$ verläuft und der Prozeß der Abschließung für die Untersuchung des Spektrums überflüssig ist.

Am Ende der Arbeit wird noch auf entsprechende Ergebnisse bei Operatoren mit allgemeinem Hauptteil hingewiesen.

Beweis

Die Behauptung hat lokalen Charakter; es genügt daher, sie für eine geeignete Umgebung eines beliebigen Punktes $x_0\in R_n$ zu beweisen. Nach [2] ist A in $\mathfrak{A}_1$ wesentlich selbstadjungiert und damit B in $\mathfrak{B}$ selbstadjungiert. Es sei $u\in\mathfrak{B}^{(\nu+m)}$ mit einem festen natürlichen $m\geqq 3$; dann ist $u\in\mathfrak{B}^{(r)}$ für $r\leqq\nu+m$ $(r\in\mathbb{N})$. Wegen der Symmetrie von B in $\mathfrak{B}$ gilt für $\varphi\in C_0^\infty(R_n)\subset\mathfrak{A}_1$

$$(\varphi, B^r u)=(B\varphi, B^{r-1}u)=(A\varphi, B^{r-1}u)$$

oder mit der Abkürzung $\varphi_r=\overline{B^{r-1}u}$ $(r=1,2,3,\ldots)$

$$\int_{R_n}\varphi_r(x)\{-A\varphi\}\,dx=-\int_{R_n}\varphi_{r+1}(x)\,\varphi(x)\,dx. \tag{2}$$

Aus dem Bestehen der Relation (2) für alle $\varphi\in C_0^\infty(R_n)$ kann man auf eine Integraldarstellung von $\varphi_r(x)$ schließen (vgl. [3]). Für zwei beliebige konzentrische Kugeln $K_j\subset K_{j+1}$ $(j=1,2,3,\ldots)$ mit Mittelpunkt x_0 gilt

$$\varphi_r(y)=-\int_{K_{j+1}}\varphi_r(x)\,A\big(\rho_{j+1}(x)\,s(y,x)\big)\,dx+\int_{K_{j+1}}\varphi_{r+1}(x)\,\rho_{j+1}(x)\,s(y,x)\,dx \tag{3}$$

fast überall in K_j. Hierbei ist $\rho_{j+1}(x)\in C_0^\infty(K_{j+1})$, $\rho_{j+1}(x)=1$ für $x\in K_j$ und $s(y,x)=\dfrac{1}{(n-2)\,\omega_n}|x-y|^{2-n}$ die Singularitätenfunktion des Differentialoperators (1); ω_n bedeutet die Oberfläche der n-dimensionalen Einheitskugel. Aus (3) folgt auf Grund bekannter Abschätzungen (vgl. [1], IV. 3.8) fast überall in K_j

$$|\varphi_r(y)|^2 \leqq c_j \int_{K_{j+1}} \frac{|\varphi_r(x)|^2}{|x-y|^{n-\eta}}\,dx + d_j \int_{K_{j+1}} \frac{|\varphi_{r+1}(x)|^2}{|x-y|^{n-2}}\,dx \tag{4}$$

mit festem irrationalen $\eta \in (0,1)$ und Konstanten c_j, d_j. Wählt man $\mu \in \mathbb{N}$ so, daß $\mu\,\eta < n$ und $(\mu+1)\,\eta > n$ gilt, dann liefert die μ-fache Iteration von (4) fast überall in K_j

$$|\varphi_r(y)|^2 \leqq c_{rj} + d_{rj} \int_{K_{2\mu+2}} \frac{|\varphi_{r+1}(x)|^2}{|x-y|^{n-2}}\,dx \tag{5}$$

mit Konstanten c_{rj}, d_{rj}. ν-fache Iteration von (5) ergibt fast überall in K_1

$$|\varphi_r(y)|^2 \leqq \gamma_r + \delta_r \int_{R_n} |\varphi_{r+\nu+1}(x)|^2\,dx \leqq \text{konst}$$

für $r=1,2,\ldots,m$ wegen $u \in \mathfrak{B}^{(\nu+m)}$, d.h. $\varphi_r(y)$ $(r=1,2,\ldots,m)$ stimmt fast überall in jeder Kugel K mit Mittelpunkt x_0 mit einer beschränkten Funktion $\tilde{\varphi}_r(y)$ überein. Nach [3] folgt daher aus der Integraldarstellung (3), daß im Fall $\alpha \in (0,1)$ $\varphi_1, \varphi_2, \ldots, \varphi_{m-1}$ fast überall mit hölderstetig differenzierbaren Funktionen $\hat{\varphi}_r$ $(r=1,2,\ldots,m-1)$ übereinstimmen; sind $K_j \supset K_{j+1}$ $(j=1,2,3,\ldots)$ konzentrische Kugeln mit Mittelpunkt x_0, so besteht nach [1], IV.4.2, für $w_{rl}(x) = \partial\hat{\varphi}_r/\partial x_l$ $(r=1,2,\ldots,m-2;\ l=1,2,\ldots,n)$ eine zu (3) analoge Integraldarstellung

$$\begin{aligned} w_{rl}(y) = {} & \int_{K_j} w_{rl}(x)\,\Delta_n\big(\rho_j(x)\,s(y,x)\big)\,dx - \int_{K_j} \hat{\varphi}_{r+1}(x)\big(\rho_j(x)\,s(y,x)\big)_{x_l}\,dx \\ & -2i \int_{K_j} \left\{\sum_{j=1}^{n} \big(w_{rj}\,b_j + \hat{\varphi}_r (b_i)_{x_j}\big)\right\} \big(\rho_j(x)\,s(y,x)\big)_{x_l}\,dx \\ & + \int_{K_j} \hat{\varphi}_r(x) \left\{\sum_{j=1}^{n} \big(i(b_j)_{x_j} + b_j^2\big) + q(x)\right\} \big(\rho_j(x)\,s(y,x)\big)_{x_l}\,dx \\ & + e_0 \int_{K_j} \hat{\varphi}_r(x)\,\|x\|^{-\alpha} \big(\rho_j(x)\,s(y,x)\big)_{x_l}\,dx \end{aligned} \tag{6}$$

für alle $y \in K_{j+1}$ mit $\rho_j(x) \in C_0^\infty(K_j)$, $\rho_j(x)=1$ für $x \in K_{j+1}$. Für $\|y\|>0$ folgt aus (6) nach den in [1], IV.4.1, angegebenen Hilfssätzen, daß $w_{rl}(y) = \partial\hat{\varphi}_r/\partial y_l$ einmal hölderstetig differenzierbar ist. Um die Regularitätsaussagen bei Annäherung an die Punktmenge $x_1^2 + x_2^2 + \cdots + x_p^2 = 0$ zu beweisen, genügt es, ein $R_1 > 0$ zu finden, derart daß zu jedem $\mathring{x} = (\mathring{x}_1, \mathring{x}_2, \ldots, \mathring{x}_n)$ mit $\sum_{\nu=1}^{p} \mathring{x}_\nu^2 = 0$ die behauptete Eigenschaft in der Kugel $|x - \mathring{x}| < R_1$ gilt. Da die Voraussetzungen keinen Punkt von $\|x\|=0$ auszeichnen, darf man o.B.d.A. den Fall $\mathring{x}=0$ betrachten. Es wird deswegen die Relation (6) für die Kugeln $K_j = \{x \mid x \in R_n,\ |x| < 1/j\}$ $(j=1,2,3,\ldots)$ betrachtet. Die ersten vier Integrale in (6) sind nach

[1], IV.4.1, einmal hölderstetig differenzierbar. Für $l=p+1,\dots,n$ ist nach [3] $\int\limits_{K_1} \hat{\varphi}_r(x)\,\|x\|^{-\alpha}(\rho_1(x)\,s(y,x))_{x_l}\,dx$ einmal hölderstetig differenzierbar. Für $l=1,2,\dots,p$ ist

$$\int\limits_{K_1-K_2} \hat{\varphi}_r(x)\,\|x\|^{-\alpha}(\rho_1(x)\,s(y,x))_{x_l}\,dx\in C^\infty(K_2)$$

und

$$\int\limits_{K_2} \hat{\varphi}_r(x)\,\|x\|^{-\alpha}\frac{\partial s}{\partial x_l}\,dx$$

$$=\int\limits_{K_2} (\hat{\varphi}_r(x)-\hat{\varphi}_r(y))\,\|x\|^{-\alpha}\frac{\partial s}{\partial x_l}\,dx+\hat{\varphi}_r(y)\int\limits_{K_2}\|x\|^{-\alpha}\frac{\partial s}{\partial x_l}\,dx,$$

wobei das erste Integral auf der rechten Seite einmal hölderstetig differenzierbar ist. Nach [4] ist

$$\int\limits_{K_2}\|x\|^{-\alpha}\frac{\partial s}{\partial x_l}\,dx=-\frac{\partial}{\partial y_l}\{h_0(y)+\zeta_0\cdot\|y\|^{2-\alpha}\}=-\int\limits_{K_2}\|x\|^{-\alpha}\frac{\partial s}{\partial y_l}\,dx \tag{7}$$

mit $h_0(y)\in C^\infty(K_2)$, $\zeta_0=\text{konst}$. Nach (6) ist daher

$$\frac{\partial\hat{\varphi}_r}{\partial y_l}=w_{rl}(y)=h_{rl}(y)-e_0\,\zeta_0\,\hat{\varphi}_r(y)\frac{\partial}{\partial y_l}\{\|y\|^{2-\alpha}\}$$

($l=1,2,\dots,n$; $r=1,2,\dots,m-2$) mit $h_{rl}(y)\in C^1(K_2)$. Daher ist $\hat{\varphi}_r\in\mathfrak{A}_1$ für $r=1,2,\dots,m-2$ bzw. $u\in\mathfrak{A}_1^{(m-2)}$.

Es sei nun $\alpha\in[1,\frac{3}{2})$. Nach [3] stimmen $\varphi_1,\varphi_2,\dots,\varphi_{m-1}$ fast überall mit Funktionen $\hat{\varphi}_1,\hat{\varphi}_2,\dots,\hat{\varphi}_{m-1}\in C^\delta(R_n)$, $0<\delta<2-\alpha$, überein, für die nach (3) die Integralgleichung

$$\hat{\varphi}_r(y)=-\int\limits_{K_1}\hat{\varphi}_r(x)\,A(\rho_1(x)\,s(y,x))\,dx+\int\limits_{K_1}\hat{\varphi}_{r+1}(x)\,\rho_1(x)\,s(y,x)\,dx \tag{8}$$

($r=1,2,\dots,m-2$) bzw.

$$\hat{\varphi}_{m-1}(y)=-\int\limits_{K_1}\hat{\varphi}_{m-1}(x)\,A(\rho_1(x)\,s(y,x))\,dx+\int\limits_{K_1}\tilde{\varphi}_m(x)\,\rho_1(x)\,s(y,x)\,dx \tag{9}$$

für alle $y\in K_2\subset K_1$ besteht. Wie in [1], IV.4.2, erschließt man aus (8) bzw. (9), daß $\hat{\varphi}_r(y)$ ($r=1,2,\dots,m-2$) für $\|y\|>0$ zweimal hölderstetig differenzierbar ist. Um das Verhalten von $\hat{\varphi}_r(y)$ bei Annäherung an die Punktmenge $x_1^2+\dots+x_p^2=0$ zu untersuchen, wird (8) bzw. (9) für die Kugeln $K_j=\{x\,|\,|x|<1/j\}$ ($j=1,2,3,\dots$) betrachtet. Wegen $\rho_1(x)=1$ für $x\in K_2$ und $\Delta_n\,s(y,x)=0$ ($x\neq y$) folgt aus (8) bzw. (9)

$$\hat{\varphi}_r(y)=h_r(y)-e_0\int\limits_{K_2}\hat{\varphi}_r(x)\,\|x\|^{-\alpha}\,s(y,x)\,dx \tag{10}$$

mit $h_r(y) \in C^{1+\delta}(K_2)$, $0 < \delta < \min(2-\alpha, \gamma)$. Für $\|y\| > 0$ liefert die Differentiation von (10)

$$\begin{aligned}\frac{\partial \hat{\varphi}_r}{\partial y_l} = \frac{\partial h_r}{\partial y_l} - e_0 \int_{K_2} (\hat{\varphi}_r(x) - \hat{\varphi}_r(y)) \|x\|^{-\alpha} \frac{\partial s}{\partial y_l} dx \\ - e_0 \hat{\varphi}_r(y) \int_{K_2} \|x\|^{-\alpha} \frac{\partial s}{\partial y_l} dx\end{aligned} \tag{11}$$

für $y \in K_2$. Wegen $\hat{\varphi}_r \in C^{\delta}(K_2)$, $0 < \delta < 2-\alpha$, ergibt eine leichte Zwischenrechnung

$$\int_{K_2} (\hat{\varphi}_r(x) - \hat{\varphi}_r(y)) \|x\|^{-\alpha} \frac{\partial s}{\partial y_l} dx \in C^{\delta^*}(K_2), \qquad 0 < \delta^* < 3 - 2\alpha.$$

Aus (11) folgt daher unter Verwendung von (7)

$$w_{rl}(y) = \frac{\partial \hat{\varphi}_r}{\partial y_l} = h_{rl}(y) - e_0 \zeta_0 \hat{\varphi}_r(y) \frac{\partial}{\partial y_l} \{\|y\|^{2-\alpha}\} \tag{12}$$

$(l = 1, 2, \ldots, n;\ r = 1, 2, \ldots, m-1)$ mit einer hölderstetigen Funktion $h_{rl}(y)$. Aus (12) folgt insbesondere $w_{rl}(y) \in L_{2,\mathrm{loc}}(R_n)$ und damit nach (6) für $y \in K_2$

$$\begin{aligned} w_{rl}(y) &= \int_{K_1} w_{rl}(x)\, \Delta_n(\rho_1(x)\, s(y,x))\, dx \\ &\quad - \int_{K_1} \left\{ \hat{\varphi}_{r+1} + 2i \sum_{j=1}^{n} w_{rj} b_j + \hat{\varphi}_r \left(i \sum_{j=1}^{n} (b_j)_{x_j} - \sum_{j=1}^{n} b_j^2 - q - e_0 \|x\|^{-\alpha} \right) \right\} \\ &\qquad \cdot (\rho_1(x)\, s(y,x))_{x_l}\, dx \\ &= \tilde{h}_{rl}(y) - 2i \int_{K_2} \sum_{j=1}^{n} w_{rj} b_j \frac{\partial s}{\partial x_l} dx + e_0 \int_{K_2} \hat{\varphi}_r \|x\|^{-\alpha} \frac{\partial s}{\partial x_l} dx \end{aligned}$$

$(r = 1, 2, \ldots, m-2;\ l = 1, 2, \ldots, n)$ mit $\tilde{h}_{rl}(y) \in C^1(K_2)$. Mit (12) ergibt sich daher

$$\begin{aligned} w_{rl}(y) = \hat{h}_{rl}(y) + 2i\, e_0 \zeta_0 \int_{K_2} \sum_{j=1}^{p} \hat{\varphi}_r(x)\, b_j(x) \frac{\partial}{\partial x_j} \{\|x\|^{2-\alpha}\} \frac{\partial s}{\partial x_l} dx \\ + e_0 \int_{K_2} \hat{\varphi}_r(x) \|x\|^{-\alpha} \frac{\partial s}{\partial x_l} dx \end{aligned} \tag{13}$$

mit $\hat{h}_{rl}(y) \in C^1(K_2)$. Es werden nun die Integrale

$$I_1 = \int_{K_2} \hat{\varphi}_r(x) \|x\|^{-\alpha} \frac{\partial s}{\partial x_l} dx,$$

$$I_2 = \int_{K_2} \hat{\varphi}_r(x)\, b_j(x) \frac{\partial}{\partial x_j} \{\|x\|^{2-\alpha}\} \frac{\partial s}{\partial x_l} dx \qquad (j = 1, 2, \ldots, p)$$

betrachtet. Für $\|y\|>0$ ist

$$\begin{aligned}\frac{\partial I_1}{\partial y_k} &= \int_{K_2} (\hat{\varphi}_r(x)-\hat{\varphi}_r(y))\, \|x\|^{-\alpha} \frac{\partial^2 s}{\partial x_l\, \partial y_k}\, dx \\ &\quad + \hat{\varphi}_r(y) \frac{\partial}{\partial y_k} \left\{ \int_{K_2} \|x\|^{-\alpha} \frac{\partial s}{\partial x_l}\, dx \right\}.\end{aligned} \tag{14}$$

Wegen $\hat{\varphi}_r(y)\in C^{\delta}(K_2)$, $0<\delta<2-\alpha$, ist für $\|y\|>0$ mit festem $\varepsilon\in(0,\delta)$ und

$$F(y)= \int_{K_2} \|y\|^{\alpha-\varepsilon}(\hat{\varphi}_r(x)-\hat{\varphi}_r(y))\, \|x\|^{-\alpha} \frac{\partial^2 s}{\partial x_l\, \partial y_k}\, dx,$$

$$\frac{|F(y)|}{\|y\|^{\alpha-\varepsilon}} \leqq \frac{1}{\|y\|^{\alpha-\varepsilon}} \int_{K_2} \frac{\operatorname{konst}(\|x\|^{\alpha-\varepsilon}+\|x-y\|^{\alpha-\varepsilon})}{\|x\|^{\alpha}\, |x-y|^{n-\delta}}\, dx \leqq \frac{\operatorname{konst}}{\|y\|^{\alpha-\varepsilon}}.$$

Da für $\|y\|>0$ $F(y)$ stetig ist, gilt nach (14)

$$\frac{\partial I_1}{\partial y_k} = \frac{\|y\|^{\varepsilon/2} f(y)}{\|y\|^{\alpha}} + g(y) \frac{\partial^2}{\partial y_l\, \partial x_k} \{\|y\|^{2-\alpha}\}$$

mit $f(y)$, $g(y)\in C^0(K_2)$. Analog findet man

$$\frac{\partial I_2}{\partial y_k} = \frac{\|y\|^{\varepsilon/2} \tilde{f}(y)}{\|y\|^{\alpha}} + \tilde{g}(y) \frac{\partial^2}{\partial y_l\, \partial y_k} \{\|y\|^{2-\alpha}\}$$

mit $\tilde{f}(y)$, $\tilde{g}(y)\in C^0(K_2)$. Nach (13) ist daher

$$\begin{aligned}\frac{\partial w_{rl}}{\partial y_k} &= \frac{\partial^2 \hat{\varphi}_r}{\partial y_l\, \partial y_k} \\ &= h_{rlk}(y)+\varphi_{rlk}(y)\, \|y\|^{-\alpha}+\psi_{rlk}(y) \frac{\partial^2}{\partial y_l\, \partial y_k} \{\|y\|^{2-\alpha}\}\end{aligned}$$

($k, l=1,2,\ldots,n$; $r=1,2,\ldots,m-2$) mit stetigen Funktionen h_{rlk}, φ_{rlk}, ψ_{rlk}, d.h. es gilt $\hat{\varphi}_r\in\mathfrak{A}_1$ bzw. $u\in\mathfrak{A}_1^{(m-2)}$.

Operatoren mit allgemeinem Hauptteil

Das hier benutzte Verfahren läßt sich auch für allgemeine elliptische Differentialoperatoren in einem Gebiet G

$$A\,u = \sum_{j,k}^{1,n} \left(i \frac{\partial}{\partial x_j} + b_j(x)\right) a_{jk}(x) \left(i \frac{\partial}{\partial x_k} + b_k(x)\right) u + q(x)\, u$$

im Teilraum $\mathfrak{A}=\{u\,|\,u\in C^2(G)\cap H,\ A\,u\in H\}$ übertragen; dabei sei $a_{jk}\in C^{2+\delta}(G)$, $b_j\in C^{1+\delta}(G)$ und A in $\mathfrak{A}$ wesentlich selbstadjungiert. Ist $q\in C^{\delta}(G)$, so gilt $u\in\mathfrak{A}^{(m-2)}$, falls $u\in\overline{\mathfrak{A}}^{(m+\nu)}$ ist[2] (vgl. auch [1], IV.4.2). Gehört q

[2] $\bar{A}$ in $\overline{\mathfrak{A}}$ ist die (selbstadjungierte) Abschließung des Operators A in $\mathfrak{A}$.

nur zur Stummelschen Klasse (d.h. es ist $\int_{|x-y|<R} |q(x)|^2 |x-y|^{-n+4-\kappa} dx \leq \text{konst}$), so folgt aus $u \in \overline{\mathfrak{A}}^{(m+\nu)}$ nur $\{u, \bar{A}u, \ldots, \bar{A}^{m-2}u\} \subset C^{\delta}(G)$ bzw. $C^{1+\delta}(G)$, je nachdem, ob $\kappa \in (0, 2]$ oder $\kappa \in (2, 4)$ gilt (vgl. [3]). Genaue Angaben über das Verhalten der Ableitungen, wie sie oben gemacht wurden, kann man natürlich in diesem Fall nicht erwarten.

Literatur

1. Hellwig, G.: Partielle Differentialgleichungen. Stuttgart: Teubner 1960.
2. Ikebe, T., Kato, T.: Uniqueness of the self-adjoint extension of singular elliptic differential operators. Arch. Rat. Mech. Analysis **9**, 77–92 (1962).
3. Witte, J.: Über die Regulariät der Spektralschar eines singulären elliptischen Differentialoperators. Math. Z. **107**, 116–126 (1968).
4. Witte, J.: Über das Verhalten der Spektralschar eines elliptischen Differentialoperators in der Umgebung der Singularität des Potentials $q(x) = |x|^{-\alpha}$. Math. Z. **115**, 140–152 (1970).

Dr. Jürgen Witte
Mathematisches Institut
der Rheinisch-Westfälischen Technischen Hochschule
D-5100 Aachen, Templergraben 55
Bundesrepublik Deutschland

(Eingegangen am 8. Mai 1972)

Math. Z. 128, 207–216 (1972)

Die Einbettung von Mittelwertstrukturen in $\mathbb{Q}$-Vektorräume

Werner Bos

Sei A eine Menge und $f=(f_n)_{n\in\mathbb{N}}$ eine Folge von Abbildungen, und zwar $f_n: A^n \to A$ mit $A^n = A\times\cdots\times A$ (n Faktoren). f heißt eine *Mittelwertstruktur*, wenn die folgenden vier Axiome erfüllt sind (vgl. [6]):

(I) f_n ist für alle $n\in\mathbb{N}$ symmetrisch.

(II) Für jedes Element $a\in A$ und $n\in\mathbb{N}$ ist die Abbildung

$$x\mapsto f_n(a,\dots,a,x)$$

eine injektive Selbstabbildung von A.

(III) $f_n(a,\dots,a)=a$ für alle $a\in A$ und alle $n\in\mathbb{N}$.

(IV) $f_n(a_1,\dots,a_n)=f_n(f_{n-1}(a_1,\dots,a_{n-1}),\dots,f_{n-1}(a_1,\dots,a_{n-1}),a_n)$ für $n>1$.

Die Mittelwertstrukturen bilden eine Kategorie; die Definition von Morphismen ist naheliegend ([6], Definition 3). Beispiele für Mittelwertstrukturen erhält man aus $\mathbb{Q}$-Vektorräumen, wobei f_n je n Vektoren deren arithmetisches Mittel zuordnet. In [6] werden diese speziellen „vollen" Mittelwertstrukturen bis auf Isomorphie durch eine nur f_2 betreffende Eigenschaft (V) charakterisiert. Ferner wird dort unter einer etwas schwächeren Voraussetzung (V′) die Einbettbarkeit in eine volle Mittelwertstruktur oder anders ausgedrückt „in einen $\mathbb{Q}$-Vektorraum" (abus de langage) bewiesen.

In der vorliegenden Arbeit wird demgegenüber ganz allgemein bewiesen, daß sich jede Mittelwertstruktur in einen $\mathbb{Q}$-Vektorraum einbetten läßt. Zum Beweis wird hierbei die gesamte Mittelwertstruktur, d.h. die ganze Folge (f_n) benutzt, während es in [6] unter der Voraussetzung (V′) gelingt, allein mit Hilfe von f_2 eine solche Einbettung zu konstruieren. Hieraus folgt z.B. insbesondere, daß unter der Voraussetzung (V′) die Abbildung f_2 die nachfolgenden Abbildungen f_n, $n>2$, bestimmt. Dieses Korollar läßt also der vorliegende allgemeine Beweis nicht zu. Ferner läßt sich das Verfahren von [6] unter Umständen leichter auf „stetige" Mittelwertstrukturen (d.h. A = topologischer Raum, f_n stetig für alle n) übertragen (vgl. [6] Beispiel 2 im Anschluß an Satz 1).

Wir verabreden die folgenden Abkürzungen der Notation:

$$f_n(a,\ldots)=f_n(a,\ldots,a),\quad f_n(a,\ldots;b)=f_n(a,\ldots,a,b),$$

$$f_n(a,\ldots;b,\underset{k}{\ldots})=f_n(\underbrace{a,\ldots,a}_{(n-k)\text{-mal}},\underbrace{b,\ldots,b}_{k\text{-mal}}),$$

$$f_n(b_1,\ldots,b_k;\ldots;;a_1,\ldots,a_s)$$
$$=f_n(b_1,\ldots,b_k,b_1,\ldots,b_k,\ldots,b_1,\ldots,b_k,a_1,\ldots,a_s).$$

Als Vorbereitung beweisen wir einige Lemmata. Lemma 1 stimmt mit [6] Hilfssatz 1 überein. Der Vollständigkeit halber wird auch der kurze Beweis dieses Lemmas hier nochmals angegeben.

Lemma 1. *Es ist für* $p, q=0,1,2,\ldots, p+q>0$:

$$f_{p+q}(a_1,\ldots,a_p,b_1,\ldots,b_q)=f_{p+q}(a_1,\ldots,a_p,f_q(b_1,\ldots,b_q),\ldots).$$

Beweis. (Induktion nach p.) Wegen (III) gilt das Lemma für $p=0$, $q=1,2,3,\ldots$ und wegen (I), (IV) für $p=1$, $q=0,1,2,3,\ldots$. Angenommen, das Lemma ist richtig für $p\geqq 1$, $q=0,1,2,3,\ldots$, dann erhalten wir unter Benutzung von (I) und (IV)

$$f_{p+1+q}(a_1,\ldots,a_{p+1},b_1,\ldots,b_q)$$
$$=f_{p+1+q}(f_{p+q}(a_1,\ldots,a_p,b_1,\ldots,b_q),\ldots;a_{p+1})$$
$$=f_{p+1+q}(f_{p+q}[a_1,\ldots,a_p,f_q(b_1,\ldots,b_q),\ldots],\ldots;a_{p+1})$$
$$=f_{p+1+q}(a_1,\ldots,a_{p+1},f_q(b_1,\ldots,b_q),\ldots).$$

Lemma 2. *Aus* $f_n(a,\underset{l}{\ldots};b,\ldots)=f_n(a,\underset{l}{\ldots};c,\ldots)$ *folgt* $b=c$ *oder* $l=n$.

Beweis. Für $l=n-1$ ergibt sich die Aussage unmittelbar aus (II). Sei $l\leqq n-2$. Unter Benutzung von Lemma 1 und (III) erhält man

$$f_n(a,\underset{l}{\ldots};b,\ldots)$$
$$=f_{n(n-l)}(f_n(a,\underset{l}{\ldots};b,\ldots),\ldots;f_n(a,\underset{l}{\ldots};c,\ldots),\underset{n}{\ldots})$$
$$=f_{n(n-l)}(a,\underset{l(n-l)}{\ldots};b,\ldots;c,\underset{n-l}{\ldots})=f_{n(n-l)}(f_n(a,\underset{l}{\ldots};b,\ldots;c),\ldots)$$
$$=f_n(a,\underset{l}{\ldots};b,\ldots;c).$$

Daraus folgt mit (IV) wegen (II) $b=c$.

Korollar. *Die Selbstabbildungen* $a\mapsto f_n(a,\ldots;b_1,\ldots,b_k)$ *von* A_f *sind für alle* $n>k$ *und* $(b_1,\ldots,b_k)\in A_f^k$ *injektiv*[1].

Dies ergibt sich wegen Lemma 1 aus dem obigen Lemma.

[1] A_f ist die der Mittelwertstruktur f zugrundeliegende Menge A.

Lemma 3. *Aus* $f_n(a, \ldots; b, \underset{i}{\ldots}) = f_n(a, \ldots; b, \underset{j}{\ldots})$, $a \neq b$, *folgt* $i = j$.

Beweis. Angenommen es ist $i > j$, dann erhalten wir aus der Voraussetzung wegen Lemma 1 mit $a' = f_{n-j}(a, \ldots; b, \underset{i-j}{\ldots})$ die Gleichung

$$f_n(a', \ldots; b, \underset{j}{\ldots}) = f_n(a, \ldots, b, \underset{j}{\ldots});$$

wegen Lemma 2 ist also $a' = a$. Andererseits folgt damit aus der Definition von a' wegen $a \neq b$ aus dem gleichen Lemma $i - j = 0$ im Widerspruch zur obigen Annahme.

Korollar. *Aus* $f_n(a, \ldots; b, \underset{l}{\ldots}) = f_m(a, \ldots; b, \underset{k}{\ldots})$, $a \neq b$, *folgt* $m\, l = n\, k$.

Beweis. Mit (III) und Lemma 1 erhält man nämlich aus der Voraussetzung $f_{nm}(a, \ldots; b, \underset{ml}{\ldots}) = f_{nm}(a, \ldots; b, \underset{nk}{\ldots})$.

Lemma 4. *Sei* $b' = f_n(b_1, \ldots, b_k; c, \underset{j}{\ldots})$ *und* $a' = f_m(a_1, \ldots, a_l; c, \underset{j}{\ldots})$, *dann ist*

$$f_{n+m-j}(a_1, \ldots, a_l, b', \ldots) = f_{n+m-j}(b_1, \ldots, b_k, a', \ldots).$$

Beweis. Durch Anwendung von Lemma 1 ergibt sich:

$$\begin{aligned} f_{n+m-j}(a_1, \ldots, a_l, b', \ldots) &= f_{n+m-j}(a_1, \ldots, a_l, b_1, \ldots, b_k, c, \underset{j}{\ldots}) \\ &= f_{n+m-j}(b_1, \ldots, b_k, a', \ldots). \end{aligned}$$

Sei im folgenden Satz f eine Mittelwertstruktur, $A = A_f$ und $m^* = m_{\mathbb{Q}}$ die Mittelwertstruktur des arithmetischen Mittels auf $\mathbb{Q}$; also $A_{m^*} = \mathbb{Q}$.

Satz 1. *Zu* $a, b \in A, a \neq b$, *gibt es einen Homomorphismus* φ: $f \to m^*$ *mit* $\varphi(a) = 0$, $\varphi(b) = 1$.

Der Beweis dieses Satzes benutzt Definitionen und Hilfssätze, die wir voranstellen:

$\mathfrak{B} \subset \mathfrak{P}A$ sei die Menge aller Teilmengen B von A mit den Eigenschaften:

(i) $a, b \in B$,

(ii) aus $(b_1, \ldots, b_n) \in B^n$ folgt $f_n(b_1, \ldots, b_n) \in B$.

Für $B \in \mathfrak{B}$ bezeichnen wir mit $f|B$ die Mittelwertstruktur $(f_n|B^n)_{n \in \mathbb{N}}$. Es sei ferner

$$\Gamma = \{\gamma \mid \gamma\colon\ f|B \to m^*,\ B \in \mathfrak{B},\ \gamma(a) = 0,\ \gamma(b) = 1\}.$$

Hilfssatz 1. *Γ ist nicht leer.*

Beweis. Es sei $B_0 = \{x \mid x \in A,\ x = f_m(a, \ldots; b, \underset{k}{\ldots}), m \geqq k \geqq 0, m \in \mathbb{N}\}$.

1) Es ist $B_0 \in \mathfrak{B}$: (i) ist klar. Zu (ii): Sei $(b_1, \ldots, b_n) \in B_0^n$, also

$$b_j = f_{m_j}(a, \ldots; \underset{k_j}{b, \ldots}), \quad j = 1, \ldots, n.$$

Wir setzen $m = m_1 \cdot \cdots \cdot m_n$ und erhalten mit Hilfe von (III) und Lemma 1

$$\begin{aligned} f_n(b_1, \ldots, b_n) &= f_{nm}(f_n(b_1, \ldots, b_n), \ldots) \\ &= f_{nm}(\underset{m}{b_1, \ldots}; \underset{m}{b_2, \ldots}; \ldots; \underset{m}{b_n, \ldots}) \\ &= f_{nm}(a, \ldots; \underset{r}{b, \ldots}) \in B_0. \end{aligned}$$

Dabei ist $r = m'_1 k_1 + \cdots + m'_n k_n$ mit $m'_j = m/m_j$, $j = 1, \ldots, n$.

2) Wir definieren im Hinblick auf Lemma 3, Korollar, γ: $B_0 \to \mathbb{Q}$ durch $f_m(a, \ldots; \underset{k}{b, \ldots}) \mapsto k/m$.

γ ist ein Homomorphismus $f|B_0 \to m^*$, denn nach 1) folgt aus $\gamma(b_j) = k_j/m_j$, $j = 1, \ldots, n$, die Gleichung

$$\gamma(f_n(b_1, \ldots, b_n)) = \frac{r}{nm} = \frac{1}{n}\left(\frac{k_1}{m_1} + \cdots + \frac{k_n}{m_n}\right) = m^*(\gamma(b_1), \ldots, \gamma(b_n)).$$

3) Es ist $\gamma(a) = \gamma(f_1(a)) = 0$ und $\gamma(b) = \gamma(f(b)) = 1$. Also ist wegen 1), 2), 3) $\gamma \in \Gamma$.

Für $\gamma \in \Gamma$, γ: $B \to \mathbb{Q}$ setzen wir $B(\gamma) := B$.

Auf Γ wird die folgende Ordnung eingeführt: $\gamma_1 < \gamma_2$ gelte genau dann, wenn $B_1 \subset B_2$ und $\gamma_1 = \gamma_2 | B_1$ ist.

Hilfssatz 2. *Γ besitzt ein maximales Element.*

Beweis. Wir benutzen den Satz von Zorn. Sei $\emptyset \neq \Delta \subset \Gamma$, Δ vollständig geordnet. Wir setzen $B_\Delta := \bigcup_{\Delta} B(\gamma)$.

Da Δ vollständig geordnet ist, gilt:

$$\text{aus } b \in B(\gamma_1) \cap B(\gamma_2),\ \gamma_i \in \Delta,\ i = 1, 2, \quad \text{folgt} \quad \gamma_1(b) = \gamma_2(b).$$

Dies rechtfertigt die folgende Definition der Abbildung γ_Δ: $B_\Delta \to \mathbb{Q}$: für $b \in B(\gamma)$, $\gamma \in \Delta$ sei $\gamma_\Delta(b) = \gamma(b)$.

Es gilt:

1) $B_\Delta \in \mathfrak{B}$. (i) ist klar wegen $\Delta \neq \emptyset$. Zu (ii): Es sei $(b_1, \ldots, b_n) \in B_\Delta^n$. Da Δ vollständig geordnet ist, gibt es ein $\gamma \in \Delta$ mit der Eigenschaft

$$(b_1, \ldots, b_n) \in (B(\gamma))^n.$$

Daher ist $f_n(b_1, \ldots, b_n) \in B(\gamma) \subset B_\Delta$.

2) γ_Δ ist ein Homomorphismus $f|B_\Delta \to m^*$. Sei $(b_1, \dots, b_n) \in B_\Delta^n$ und wie in 1) $\gamma \in \Delta$ mit $(b_1, \dots, b_n) \in (B(\gamma))^n$, dann ist

$$\begin{aligned}\gamma_\Delta f_n(b_1, \dots, b_n) = \gamma f_n(b_1, \dots, b_n) &= m^*(\gamma(b_1), \dots, \gamma(b_n)) \\ &= m^*(\gamma_\Delta(b_1), \dots, \gamma_\Delta(b_n)).\end{aligned}$$

3) Offenbar ist $\gamma_\Delta(a) = 0$, $\gamma_\Delta(b) = 1$, also $\gamma_\Delta \in \Gamma$.
Ferner ist klar, daß γ_Δ eine obere Schranke für Δ ist.

Hilfssatz 3. *Für jedes maximale Element $\gamma \in \Gamma$ gilt $B(\gamma) = A$.*

Beweis. Indirekt sei γ ein maximales Element und sei $c \in A \setminus B$, wobei wir in diesem Beweis zur Abkürzung $B = B(\gamma)$ setzen.

Wegen $B(\gamma) \in \mathfrak{B}$ ist $c \neq f_n(b_1, \dots, b_n)$ für jede Wahl von $n \in \mathbb{N}$ und $(b_1, \dots, b_n) \in B^n$.

1) Sei $\bar{B} = \{y \mid y \in A,\ y = f_l(x_1, \dots, x_l), l \in \mathbb{N}, (x_1, \dots, x_l) \in (B \cup \{c\})^l\}$. $\bar{B} \in \mathfrak{B}$, denn: (i) ist klar; zu (ii): sei $(y_1, \dots, y_n) \in \bar{B}^n$, $y_i = f_{l_i}(x_{i1}, \dots, x_{il_i})$, $x_{ij} \in B \cup \{c\}$, $j = 1, \dots, l_i$, $i = 1, \dots, n$, dann ergibt sich analog wie beim Abschnitt 1) des Beweises von Hilfssatz 1 mit $l = l_1 \cdot \dots \cdot l_n$ und $l_i' = l/l_i$, $i = 1, \dots, n$,

$$f_n(y_1, \dots, y_n) = f_{nl}(x_{11}, \dots, x_{1l_1}; \underset{l_1'}{\dots}; \dots; x_{n1}, \dots, x_{nl_n}; \underset{l_n'}{\dots}) \in \bar{B}.$$

2) Wir setzen zunächst $\bar{\gamma}(b) = \gamma(b)$ für $b \in B$, $\bar{\gamma}(c) = u$ und zwar sei vereinbart:

(α) $u = 0$, wenn für jede Wahl von $n, k \in \mathbb{N}, n > k$ und $(b_1, \dots, b_k) \in B^k$ gilt $f_n(b_1, \dots, b_k, c, \dots) \notin B$.

(β) u ist die Lösung der Gleichung

$$\gamma(f_n(b_1, \dots, b_k, c, \dots)) = \frac{1}{n}\left[\sum_{i=1}^{k} \gamma(b_i) + (n-k)\,u\right]$$

für eine feste Wahl von $n, k \in \mathbb{N}$, $n > k$, $(b_1, \dots, b_k) \in B^k$ mit der Eigenschaft $b' = f_n(b_1, \dots, b_k, c, \dots) \in B$.

Ist $y \in \bar{B}$, $y = f_l(x_1, \dots, x_l)$, $(x_1, \dots, x_l) \in (B \cup \{c\})^l$, dann setzen wir außerdem $\bar{\gamma}(y) = \frac{1}{l}\sum_{i=1}^{l} \bar{\gamma}(x_i)$.

Wir werden zeigen, daß in beiden Fällen $\bar{\gamma}(y)$ von der speziellen Wahl der Gleichung $y = f_l(x_1, \dots, x_l)$ und im Fall (β) überdies auch von der dort getroffenen Wahl unabhängig ist.

Wir setzen also voraus

$$\begin{aligned}&f_l(w_1, \dots, w_l) = f_m(x_1, \dots, x_m);\\ &w_i, x_j \in B \cup \{c\},\ i = 1, \dots, l,\ j = 1, \dots, m.\end{aligned} \tag{1}$$

und behaupten:
$$\frac{1}{l}\sum_{i=1}^{l}\bar{\gamma}(w_i)=\frac{1}{m}\sum_{j=1}^{m}\bar{\gamma}(x_j). \tag{2}$$

Da γ ein Homomorphismus ist, ist dies trivialerweise richtig, falls $w_i, x_j \in B$, $i=1, \dots, l$, $j=1, \dots, m$ gilt.

Der Fall (α): Außer in dem gerade erwähnten trivialen Fall folgt aus der Voraussetzung von (α) und (1), daß auf beiden Seiten von (1) c als Argument vorkommt. Wir haben also den folgenden Sachverhalt

$$\begin{gathered} f_l(d_1, \dots, d_{l-k}, c, \dots)=f_m(h_1, \dots, h_{m-j}, c, \dots), \\ 1 \leqq k < l,\ 1 \leqq j < m,\ d_1, \dots, d_{l-k}, h_1, \dots, h_{m-j} \in B. \end{gathered} \tag{1'}$$

Aus Lemma 1 und (III) ergibt sich daraus mit

$$d=f_{l-k}(d_1, \dots, d_{l-k}), \qquad h=f_{m-j}(h_1, \dots, h_{m-j}), \tag{3}$$

und somit $d, h \in B$:

$$f_{lm}(d, \dots; c, \underset{mk}{\dots})=f_{lm}(h, \dots; c, \underset{lj}{\dots}). \tag{1''}$$

Angenommen es wäre hierin $mk \neq lj$, etwa $mk > lj$, dann ergibt sich in gleicher Weise mit

$$p=f_{l(m-j)}(d, \dots; c, \underset{mk-lj}{\dots}) \tag{4}$$

$f_{lm}(p, \dots; c, \underset{lj}{\dots})=f_{lm}(h, \dots; c, \underset{lj}{\dots})$. Also ist nach Lemma 2 $h=p$; d.h. es wäre $B \ni h=p \notin B$, letzteres wegen (α).

Es ist also in (1'') $mk=lj$ und daher nach Lemma 2 $d=h$. Aus (3) folgt damit durch Anwendung von $\gamma=\bar{\gamma}|B$

$$\frac{1}{l-k}\sum_{i=1}^{l-k}\gamma(d_i)=\frac{1}{m-j}\sum_{i=1}^{m-j}\gamma(h_i).$$

Daraus folgt wegen $mk=lj$ und $\bar{\gamma}(c)=0$

$$\frac{1}{l}\left(\sum_{i=1}^{l-k}\bar{\gamma}(d_i)+k\bar{\gamma}(c)\right)=\frac{1}{m}\left(\sum_{i=1}^{m-j}\bar{\gamma}(h_i)+j\bar{\gamma}(c)\right). \tag{2'}$$

Der Fall (β): Wir zeigen zunächst, daß u nicht von der in (β) getroffenen speziellen Wahl von n, k und $(b_1, \dots, b_k) \in B^k$ abhängt.

Sei $a'=f_m(a_1, \dots, a_l, c, \dots) \in B$, $\quad m > l$, $(a_1, \dots, a_l) \in B^l$.

Aus (β) und hieraus erhalten wir durch Anwendung von (III) und Lemma 1

$$b'=f_{n(m-l)}(b_1, \dots, b_k; \dots; ; c, \underset{(n-k)(m-l)}{\dots})$$

$$a'=f_{m(n-k)}(a_1, \dots, a_l; \dots; ; c, \underset{(n-k)(m-l)}{\dots}).$$

Daraus ergibt sich mit Lemma 4

$$f_{nm-kl}(b_1, \dots, b_k; \dots; a', \underset{m(n-k)}{\dots}) = f_{nm-kl}(a_1, \dots, a_l; \dots; b', \underset{n(m-l)}{\dots})$$

und daraus mit Hilfe des Homomorphismus γ:

$$(m-l)\sum_{i=1}^{k} \gamma(b_i) + m(n-k)\,\gamma(a') = (n-k)\sum_{i=1}^{l} \gamma(a_i) + n(m-l)\,\gamma(b'),$$

oder durch Einsetzen von u aus (β) ergibt sich

$$\gamma(a') = \frac{1}{m}\left(\sum_{i=1}^{l} \gamma(a_i) + (m-l)\,u\right),$$

d.h. die Unabhängigkeit von u von der in (β) getroffenen Wahl.

Wir setzen nun wieder (1) voraus und beweisen (2): Wir können jetzt ohne Beschränkung der Allgemeinheit von folgendem Sachverhalt ausgehen:

$$f_l(d_1, \dots, d_{l-k}, c, \dots) = f_m(h_1, \dots, h_{m-j}, c, \dots), \tag{1*}$$
$$0 \leqq k < l,\ 0 \leqq j < m,\ d_1, \dots, d_{l-k}, h_1, \dots, h_{m-j} \in B,\ mk \geqq lj.$$

Wir benutzen wieder (3) und (4) als Definition und erhalten wie dort (auch für den Fall $mk = lj$) $h = p$, d.h.

$$f_{m-j}(h_1, \dots, h_{m-j}) = f_{l(m-j)}(d, \dots; c, \underset{mk-lj}{\dots}).$$

Die linke Seite ist ein Element von B. Sofern $mk - lj > 0$ ist, kann also, wie wir gerade gezeigt haben, diese Gleichung wie in (β) zur Definition von $u = \bar{\gamma}(c)$ benutzt werden. Im Falle $mk - lj = 0$ ergeben sich die folgenden Gleichungen aus der Homomorphismuseigenschaft von γ:

$$\begin{aligned}\frac{1}{m-j}\sum_{i=1}^{m-j} \gamma(h_i) &= \frac{m(l-k)}{l(m-j)}\,\gamma(d) + \frac{mk-lj}{l(m-j)}\,\bar{\gamma}(c)\\ &= \frac{m}{l(m-j)}\sum_{i=1}^{l-k} \gamma(d_i) + \frac{mk-lj}{l(m-j)}\,\bar{\gamma}(c).\end{aligned}$$

Dies ergibt, wobei wir $\bar{\gamma}|B = \gamma$ berücksichtigen:

$$\frac{1}{m}\left(\sum_{i=1}^{m-j} \bar{\gamma}(h_i) + j\,\bar{\gamma}(c)\right) = \frac{1}{l}\left(\sum_{i=1}^{l-k} \bar{\gamma}(d_i) + k\,\bar{\gamma}(c)\right). \tag{2*}$$

3) $\bar{\gamma}: \bar{B} \to \mathbb{Q}$ ist ein Homomorphismus $\bar{\gamma}: f|\bar{B} \to m^*$: Sei

$$(y_1, \dots, y_m) \in \bar{B}^m, \quad y_i = f_{l_i}(z_{i1}, \dots, z_{ik_i}, c, \underset{j_i}{\dots}),$$
$$k_i + j_i = l_i, \quad i = 1, \dots, m.$$

Nach Definition von $\bar{\gamma}$ ist

$$\bar{\gamma}(y_i)=\frac{1}{l_i}\left(\sum_{\nu=1}^{k_i}\bar{\gamma}(z_{i_\nu})+j_i\bar{\gamma}(c)\right).$$

Mit den Abkürzungen

$$l=l_1\cdot\cdots\cdot l_m,\qquad l_i'=\frac{l}{l_i},\qquad i=1,\ldots,m,$$

und $r=l_1' j_1+\cdots+l_m' j_m$ erhalten wir mit Hilfe von (III) und Lemma 1

$$f_m(y_1,\ldots,y_m)=f_{ml}(z_1,\ldots,z_{1k_1};\underset{l_1'}{\ldots};\ldots;\ldots;z_{m1},\ldots,z_{mk_m};\underset{l_m'}{\ldots};c,\underset{r}{\ldots})$$

und somit nach Definition von $\bar{\gamma}$:

$$\begin{aligned}\bar{\gamma} f_m(y_1,\ldots,y_m)&=\frac{1}{m\,l}\left(\sum_{i=1}^{m} l_i'\left(\sum_{\nu=1}^{k_i}\bar{\gamma}(z_{i_\nu})+j_i\bar{\gamma}(c)\right)\right)\\&=\frac{1}{m}\sum_{i=1}^{m}\bar{\gamma}(y_i).\end{aligned}$$

4) Wegen $\bar{\gamma}\neq\gamma$, $\bar{\gamma}>\gamma$ ist somit γ entgegen unserer Voraussetzung nicht maximal.

Also ist $B(\gamma)=A$.

Beweis von Satz 1. $\varphi\in\Gamma$, φ maximal existiert nach Hilfssatz 2 und genügt wegen Hilfssatz 3 den Erfordernissen von Satz 1.

Satz 2. *Jede Mittelwertstruktur f läßt sich in einen $\mathbb{Q}$-Vektorraum einbetten.*

Beweis. 1) Sei $A=A_f$, $\mathbb{V}$ der $\mathbb{Q}$-Vektorraum $\bigoplus_A \mathbb{Q}$ und U der von der Menge

$$\left\{x\,|\,x\in\mathbb{V},\ x=f_n(a_1,\ldots,a_n)-\frac{1}{n}\sum_{i=1}^{n}a_i,\ (a_1,\ldots,a_n)\in A^n,\ n\in\mathbb{N}\right\}$$

erzeugte Untervektorraum.

Dabei ist die Menge A als in der kanonischen Weise in $\mathbb{V}$ eingebettet vorausgesetzt. Sei $V=\mathbb{V}/U$ und φ: $A\subset\mathbb{V}\to V$ die zusammengesetzte kanonische Abbildung. m_V sei die Mittelwertstruktur des arithmetischen Mittels im $\mathbb{Q}$-Vektorraum V. Wir werden zeigen, daß φ eine Einbettung $f\to m_V$ ist. Jedenfalls ist φ ein Homomorphismus $f\to m_V$, denn

$$\varphi f_n(a_1,\ldots,a_n)-\frac{1}{n}\sum_{i=1}^{n}\varphi(a_i)=0.$$

2) Sei W ein beliebiger $\mathbb{Q}$-Vektorraum und m_W die Mittelwertstruktur des arithmetischen Mittels auf W, $\psi: f\to m_W$ ein Homomorphismus, dann gibt es einen Homomorphismus χ: $m_V\to m_W$, so daß das folgende Dreieck kommutativ ist:

$$\begin{tikzcd} m_V \arrow[rr, "\chi"] & & m_W \\ & f \arrow[ul, "\varphi"] \arrow[ur, "\psi"'] & \end{tikzcd} \tag{$*$}$$

Da A eine Basis von $\mathbb{V}$ ist, liefert die Festsetzung $a \mapsto \psi(a)$ eine lineare Abbildung $\tilde{\chi}$: $\mathbb{V} \to W$; da ψ ein Homomorphismus $f \to m_W$ ist, ist

$$\tilde{\chi}\, f_n(a_1, \ldots, a_n) = \psi\, f_n(a_1, \ldots, a_n) = \frac{1}{n}\sum_{i=1}^{n} \psi(a_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \tilde{\chi}(a_i) = \tilde{\chi}\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right).$$

Mit anderen Worten:

$$\tilde{\chi}\left(f_n(a_1, \ldots, a_n) - \frac{1}{n}\sum_{i=1}^{n} a_i\right) = 0,$$

also $\tilde{\chi}(U)=0$. $\tilde{\chi}$ induziert somit eine lineare Abbildung χ: $V \to W$, also einen Homomorphismus χ: $m_V \to m_W$. Nach Definition von χ ist

$$\chi \circ \varphi(a) = \psi(a).$$

3) φ: $f \to m_V$ ist eine Einbettung, d.h. φ: $A \to V$ ist injektiv: Sei $a, b \in A$, $a \neq b$, dann gibt es nach Satz 1 einen Homomorphismus ψ: $f \to m_{\mathbb{Q}}$ mit $\psi(a) \neq \psi(b)$. Setzen wir in 2) $W = \mathbb{Q}$, so ergibt sich

$$\chi\, \varphi(a) = \psi(a) \neq \psi(b) = \chi\, \varphi(b)$$

und daher $\varphi(a) \neq \varphi(b)$.

Bemerkung[2]. Zu einem $\mathbb{Q}$-Vektorraum W und einem Homomorphismus ψ: $f \to m_W$ kann es sehr wohl verschiedene Homomorphismen χ: $m_V \to m_W$ geben, die $(*)$ kommutativ machen, z.B. stets dann, wenn f selbst Mittelwertstruktur des arithmetischen Mittels auf einem $\mathbb{Q}$-Vektorraum und ψ eine konstante Abbildung $\neq 0$ ist. Unter diesen Homomorphismen ist der in 2) definierte jedoch der einzige, der zugleich eine lineare Abbildung $V \to W$ von $\mathbb{Q}$-Vektorräumen ist.

Bezeichnet S den Funktor von der Kategorie der $\mathbb{Q}$-Vektorräume und linearen Abbildungen in die Kategorie der Mittelwertstrukturen und Homomorphismen, der einem $\mathbb{Q}$-Vektorraum W die Mittelwertstruktur m_W des arithmetischen Mittels auf W zuordnet und eine lineare Abbildung $W \to W'$ als Homomorphismus $m_W \to m_{W'}$ von Mittelwertstrukturen auffaßt, so haben wir in der Tat folgendes bewiesen:

Zu jeder Mittelwertstruktur f existiert ein $\mathbb{Q}$-Vektorraum V und eine Einbettung φ: $f \to m_V$, so daß $\langle V, \varphi \rangle$ im Sinn von [8], III.1 *ein universeller Pfeil von f nach S ist.*

[2] Diese Bemerkung verdanke ich einem freundlichen Hinweis von Herrn K.H. Kamps.

Literatur

1. Aczél, J.: Lectures on functional equations and their applications. New York and London: Academic Press 1966, 234–244[3].
2. Aumann, G.: Aufbau von Mittelwerten mehrerer Argumente I. Math. Ann. **109**, 235–253 (1934).
3. Aumann, G.: Aufbau von Mittelwerten mehrerer Argumente II. Math. Ann. **111**, 713–730 (1935).
4. Aumann, G.: Über Räume mit Mittelbildungen. Math. Ann. **119**, 210–215 (1943).
5. Bos, W.: Axiomatische Charakterisierung des arithmetischen Mittels. Math.-phys. Semesterber. **18**, 45–53 (1971).
6. Bos, W.: Mittelwertstrukturen. Erscheint in Math. Ann.
7. Eckmann, B.: Räume mit Mittelbildungen. Commentarii Math. Helvet. **28**, 329–340 (1954).
8. MacLane, S.: Kategorien. Berlin-Heidelberg-New York: Springer-Verlag 1972.
9. Whittaker, E., Robinson, G.: The calculus of observations 4. ed. New York: Dover Publications Inc. 1967.

[3] Hier sind zahlreiche weitere Literaturangaben zu finden.

Prof. Werner Bos
Universität Fachbereich Mathematik
D-7750 Konstanz
Bundesrepublik Deutschland

(Eingegangen am 18. Mai 1972)

Math. Z. 128, 217–230 (1972)

Ergodische Theorie von Ziffernentwicklungen in Wahrscheinlichkeitsräumen

Roland Fischer

1. Einleitung

$(I, \mathfrak{F}, m)$ sei ein Wahrscheinlichkeitsraum, d.h. I eine Menge, $\mathfrak{F}$ eine σ-Algebra von Teilmengen von I und m ein normiertes vollständiges Maß auf $\mathfrak{F}$. T sei eine Transformation von I in sich, die folgende Eigenschaften (a) bis (e) habe:

(a) Mit $E \in \mathfrak{F}$ seien auch TE, $T^{-1}E \in \mathfrak{F}$ (d.h. T ist bimeßbar). Die Aussagen $m(E)=0$, $m(TE)=0$, $m(T^{-1}E)=0$ seien für alle $E \in \mathfrak{F}$ gleichbedeutend.

(b) Es existiere eine meßbare Partition $\mathfrak{J}_1 = \{I(k) \mid k \in \mathbb{N}\}$ von I so, daß die Einschränkung von T auf $I(k)$ eine injektive Abbildung ist für alle $k \in \mathbb{N}$.

Beginnend mit den Mengen $I(k)$ definieren wir rekursiv Bereiche $I(k_1, \ldots, k_n)$ $(k_i \in \mathbb{N})$ durch

$$T^{-1} I(k_2, \ldots, k_n) \cap I(k_1) = I(k_1, \ldots, k_n) \quad \text{für } n = 2, 3, \ldots. \tag{1}$$

Es gilt dann

$$I(k_1, \ldots, k_n) = \bigcup_{k_{n+1}} I(k_1, \ldots, k_n, k_{n+1}). \tag{2}$$

Eine Menge $I(k_1, \ldots, k_n) \neq \emptyset$ heiße Zylinder n-ter Ordnung und $\mathfrak{J}_n$ sei die Menge der Zylinder n-ter Ordnung und $\mathfrak{J}$ die Menge aller Zylinder. Die Mengen $\mathfrak{J}_n$ $(n=1, 2, \ldots)$ sind Partitionen von I, wobei $\mathfrak{J}_{n+1}$ eine Verfeinerung von $\mathfrak{J}_n$ ist. Die Restriktion von T^n auf einen Zylinder n-ter Ordnung ist injektiv. Für $x \in I$ sei $I_n(x)$ jener Zylinder n-ter Ordnung, der x enthält. Jedem $x \in I$ wird nun eine Folge $(a_n)_{n \in \mathbb{N}}$ natürlicher Zahlen zugeordnet, indem man definiert: $a_n(x) = a_n$, wenn

$$I_n(x) = I(a_1, \ldots, a_n). \tag{3}$$

Gilt nun

$$\lim_{n \to \infty} I_n(x) = \{x\} \tag{4}$$

für alle $x \in I$, dann ist die Zuordnung $x \mapsto (a_n)_{n \in \mathbb{N}}$ injektiv und man kann sie als „Ziffernentwicklung von x bezüglich T" auffassen. T ist dann ein Verschiebungsoperator. In dieser Arbeit wird jedoch (4) nicht vorausgesetzt werden.

(c) Die Vervollständigung bezüglich m der von $\mathfrak{J}$ erzeugten τ-Algebra stimme mit $\mathfrak{F}$ überein.

In metrischen Wahrscheinlichkeitsräumen folgt (c) aus

$$\lim_{n \to \infty} \operatorname{diam} I_n(x) = 0,$$

wenn $\mathfrak{F}$ die Vervollständigung der σ-Algebra der Borelmengen ist. Wegen (a) und der Injektivität von T^n auf jedem Zylinder $I(k_1, \ldots, k_n)$ existiert eine Radon-Nikodym-Ableitung $\dfrac{dm\,T^n}{dm}(x)$ auf dieser Menge.

(d) Es gebe eine Konstante $C > 1$, so daß für jedes $I(k_1, \ldots, k_n) \in \mathfrak{J}n$

$$\operatorname{ess.\,sup} \frac{dm\,T^n}{dm}(x) \leqq C.\ \operatorname{ess.\,inf} \frac{dm\,T^n}{dm}(x)$$

gilt, wobei ess. sup und ess. inf über alle $x \in I(k_1, \ldots, k_n)$ zu erstrecken sind.

Ein Zylinder $I(k_1, \ldots, k_n)$ heiße *eigentlich*, wenn

$$T^n I(k_1, \ldots, k_n) = I$$

ist, sonst *uneigentlich*. Damit definieren wir:

$$\mathfrak{A}_n = \{I_n \in \mathfrak{J}_n \mid I_n \text{ eigentlich}\}, \tag{5}$$

$$\mathfrak{C}_n = \mathfrak{J}_n \setminus \mathfrak{A}_n, \tag{6}$$

$$\mathfrak{D}_n = \{I(k_1, \ldots, k_n) \in \mathfrak{J}_n \mid I(k_1), I(k_1, k_2), \ldots, I(k_1, k_2, \ldots, k_n) \text{ uneigentlich}\}, \tag{7}$$

$$\mathfrak{B}_n = \{I(k_1, \ldots, k_n) \mid I(k_1, \ldots, k_{n-1}) \in \mathfrak{D}_{n-1},\ I(k_1, \ldots, k_n) \in \mathfrak{A}_n\}; \tag{8}$$

$$B_n = \bigcup_{B \in \mathfrak{B}_n} B, \qquad D_n = \bigcup_{D \in \mathfrak{D}_n} D. \tag{9}$$

(e) Es sei

$$\sum_{n=1}^{\infty} m(D_n) < \infty.$$

Es folgen einige Beispiele von Transformationen T, die die Eigenschaften (a) bis (e) besitzen:

1) Sei $I = (0, 1)$, m das Lebesquesche Maß,

$$Tx = \beta\, x \bmod 1$$

mit $\beta > 1$. (Resttransformation bei β-adischen Ziffernentwicklungen, vgl. Cigler [1].)

2) I sei das Einheitsquadrat in der komplexen Zahlenebene, m das zweidimensionale Lebesguesche Maß.

$$Tx = s\,x \bmod 1$$

mit $s \in \mathbb{C}$, $|s| > 2(\sqrt{2} + \sqrt{3})$. (Komplexe Ziffernentwicklungen, Fischer [2].)

3) Sei $I = (0, 1)^d$, m das d-dimensionale Lebesguesche Maß und

$$Tx = A\,x \bmod 1,$$

wobei A eine nichtsinguläre symmetrische $(d \times d)$-Matrix sei, deren Eigenwerte alle größer als $d \cdot 2^{d+1}$ sind. (Philipp [5].)

4) $I = (0, 1)^d$ vermindert um eine gewisse Nullmenge, m das d-dimensionale Lebesguesche Maß und

$$T(x_1, x_2, \ldots, x_d) = \left(\frac{x_2}{x_1}, \frac{x_3}{x_1}, \ldots, \frac{x_{n-1}}{x_1}, \frac{1}{x_1}\right) \bmod 1.$$

(Resttransformation des Jacobischen Algorithmus bzw. für $d = 1$ des Kettenbruchalgorithmus, vgl. Schweiger [8].)

Transformationen T von der oben beschriebenen Art wurden von Schweiger unter dem Namen „Zahlentheoretische Transformationen" in [7] eingeführt. Allerdings betrachtet Schweiger nur solche Transformationen, bei denen sämtliche Zylinder eigentlich sind, wodurch (e) immer erfüllt ist. Waterman verallgemeinert in [9] einige Ergebnisse von Schweiger auf Transformationen, bei denen auch uneigentliche Zylinder vorkommen dürfen, benützt aber weitere Voraussetzungen, die bei den angeführten Beispielen 1) − 3) im allgemeinen nicht zutreffen. In der vorliegenden Arbeit werden Resultate von Schweiger und Waterman verallgemeinert, wobei zum Teil andere Beweismethoden verwendet werden.

2. Ergodizität von T

Sei $\mathfrak{A} = \bigcup_{n=1}^{\infty} \mathfrak{A}_n$ die Menge aller eigentlichen Zylinder.

Lemma 1. *Jeder Zylinder kann modulo einer Nullmenge als disjunkte Vereinigung von eigentlichen Zylindern dargestellt werden.*

Beweis. Wir vermerken zunächst, daß für $n = 1, 2, \ldots$

$$I = B_1 \cup B_2 \cup \cdots \cup B_n \cup D_n \tag{10}$$

gilt, wobei $\bigcup_{i=1}^{n} B_i$ eine Vereinigung eigentlicher Zylinder ist. Außerdem gilt

$$D_n = B_{n+1} \cup B_{n+2} \cup \cdots \cup B_{n+k} \cup D_{n+k}. \tag{11}$$

Aus (e) folgt somit die Behauptung für Zylinder $I_n \in \mathfrak{D}_n$. Ist nun $I(k_1, k_2, \ldots, k_n) \notin \mathfrak{D}_n$, dann gibt es ein r, $1 \leqq r \leqq n$, so daß $I(k_1, \ldots, k_r) \in \mathfrak{A}_r$. Wenn wir r maximal wählen, dann gilt

$$I(k_{r+1}, \ldots, k_n) \in \mathfrak{D}_{n-r}.$$

Denn wäre $I(k_{r+1}, \ldots, k_{r+i}) \in \mathfrak{A}_i$ $(i \leqq n-r)$, dann wäre auch $I(k_1, \ldots, k_{r+i}) = I(k_1, \ldots, k_r) \cap T^{-r} I(k_{r+1}, \ldots, k_{r+i})$ ein eigentlicher Zylinder im Widerspruch zur Maximalität von r. Damit wissen wir aber, daß

$$I(k_{r+1}, \ldots, k_n) = \bigcup I_s \bmod 0$$

ist, wobei $I_s \in \mathfrak{A}$. Wegen der Nichtsingularität von T können wir jetzt schließen:

$$\begin{aligned} I(k_1, \ldots, k_n) &= I(k_1, \ldots, k_r) \cap T^{-r} I(k_{r+1}, \ldots, k_n) = I(k_1, \ldots, k_r) \cap \bigcup T^{-r} I_s \\ &= \bigcup I(k_1, \ldots, k_r) \cap T^{-r} I_s \bmod 0. \end{aligned}$$

Da die Mengen $I(k_1, \ldots, k_r) \cap T^{-r} I_s$ eigentliche Zylinder sind, ist alles gesagt. (W.z.b.w.)

Lemma 2. *Für fast alle $x \in I$ gilt: In der Folge $(I_n(x))_{n \in \mathbb{N}}$ kommen unendlich viele eigentliche Zylinder vor.*

Beweis. Wir zeigen: Für alle $n \in \mathbb{N}$ ist die Menge der x, die in keinem eigentlichen Zylinder einer Ordnung $\geqq n$ liegen, vom Maß Null. Dies folgt aber aus Lemma 1, da sich jedes Atom der Partition $\mathfrak{Z}_n$ modulo 0 als disjunkte Vereinigung von eigentlichen Zylindern mit Ordnung $\geqq n$ darstellen läßt, w.z.b.w.

Lemma 3. *Die disjunkten Vereinigungen von Mengen aus $\mathfrak{A}$ liegen in $\mathfrak{F}$ dicht.*

Beweis. Dies folgt aus (c), Lemma 1 und der Tatsache, daß Vereinigungen von Zylindern immer als disjunkte Vereinigungen geschrieben werden können, w.z.b.w.

Da die Restriktion von T^n auf $I(k_1, \ldots, k_n)$ injektiv ist, können wir die Umkehrabbildung

$$V(k_1, \ldots, k_n)\colon\ T^n I(k_1, \ldots, k_n) \to I(k_1, \ldots, k_n) \tag{12}$$

definieren. Aus (a) folgt: Es existiert

$$\frac{dm\,V(k_1,\ldots,k_n)}{dm}(x)=\Delta(k_1,\ldots,k_n)(x) \tag{13}$$

für $x\in T^n I(k_1,\ldots,k_n)$. Weiter gilt

$$\Delta(k_1,\ldots,k_n)(T^n x)=\left[\frac{dm\,T^n}{dm}(x)\right]^{-1} \quad \text{f.ü.} \tag{14}$$

und damit gilt wegen (d):

$$\text{ess. sup}\,\Delta(k_1,\ldots,k_n)(x)\leqq C\,\text{ess. inf}\,\Delta(k_1,\ldots,k_n)(x). \tag{15}$$

Lemma 4. *Für alle $Y\in\mathfrak{F}$ mit $m(Y)>0$ gilt:* $\limsup_{n\to\infty} m(T^n Y)=1$.

Beweis. Es sei $I(k_1,\ldots,k_n)$ eigentlich und $X\subseteq I(k_1,\ldots,k_n)$, dann ist

$$\begin{aligned} m(X)&=\int_{T^n X}\Delta(k_1,\ldots,k_n)\,dm\geqq \text{ess. inf}\,\Delta(k_1,\ldots,k_n)\cdot m(T^n X)\\ &\geqq C^{-1}\,\text{ess. sup}\,\Delta(k_1,\ldots,k_n)\cdot m(T^n X)\geqq C^{-1}\,m\big(I(k_1,\ldots,k_n)\big)\,m(T^n X), \end{aligned}$$

das bedeutet

$$m(T^n X)\leqq C\cdot\frac{m(X)}{m\big(I(k_1,\ldots,k_n)\big)}. \tag{16}$$

Sei nun $Y\in\mathfrak{F}$ und $m(Y)>0$. Wir wählen ein $\delta>0$. Es kann die Ungleichung

$$m(Y\cap A)\leqq\left(1-\frac{\delta}{C}\right)m(A)$$

nicht für alle $A\in\mathfrak{A}$ richtig sein, sonst wäre sie auch für alle abzählbaren disjunkten Vereinigungen von Mengen aus $\mathfrak{A}$ und wegen Lemma 3 auch für alle $A\in\mathfrak{F}$, was mit $A=Y$

$$m(Y)\leqq\left(1-\frac{\delta}{C}\right)m(Y)$$

ergäbe. Es gibt daher ein $A_n\in\mathfrak{A}_n$ mit

$$\frac{m(Y\cap A_n)}{m(A_n)}>\left(1-\frac{\delta}{C}\right). \tag{17}$$

Da $A_n\setminus(Y\cap A_n)\subseteq A_n$ gilt nach (16) und (17):

$$\begin{aligned} m\big(T^n(A_n\setminus(Y\cap A_n))\big)&\leqq C\cdot\frac{m\big(A_n\setminus(Y\cap A_n)\big)}{m(A_n)}\\ &=C\cdot\frac{m(A_n)-m(Y\cap A_n)}{m(A_n)}=C\left(1-\frac{m(Y\cap A_n)}{m(A_n)}\right)\leqq\delta. \end{aligned} \tag{18}$$

Daraus sehen wir:

$$m(T^n Y) \geqq m\big(T^n(Y \cap A_n)\big) = m\big(T^n(A_n \setminus (A_n \setminus (Y \cap A_n)))\big)$$
$$= m(T^n A_n) - m\big(T^n(A_n \setminus (Y \cap A_n))\big) \geqq 1 - \delta. \quad \text{(W.z.b.w.)}$$

Satz 2. *T ist ergodisch bezüglich m.*

Beweis. E sei eine invariante Menge, d.h. $T^{-1}E = E$. Dann gilt $T(I \setminus E) \subseteq I \setminus E$. Ist $m(I \setminus E) > 0$, so können wir Lemma 4 mit $Y = I \setminus E$ anwenden und es folgt $m(I \setminus E) = 1$. (W.z.b.w.)

In diesem Abschnitt wurde statt (e) nur $\lim_{n\to\infty} m(D_n) = 0$ verwendet.

3. Das invariante Maß

In diesem Abschnitt soll die Existenz eines zu m äquivalenten, bezüglich T invarianten Wahrscheinlichkeitsmaßes nachgewiesen werden.

Satz 2. *Es sei $\rho(x) \in L^1(m)$ die Dichtefunktion eines bezüglich m absolut stetigen Maßes μ. Dann gilt: μ ist genau dann invariant bezüglich T, wenn*

$$\rho(x) = \sum_{x \in TI(k)} \rho\big(V(k)(x)\big)\,\Delta(k)(x) \qquad \text{m.-f. ü.} \tag{19}$$

gilt. (Kuzminsche Gleichung, siehe [9].)

Beweis. Sei $E \in \mathfrak{F}$. Folgende Rechnung zeigt die Richtigkeit der Behauptung:

$$\begin{aligned}
\mu(T^{-1}E) &= \int_{T^{-1}E} \rho(x)\,dm = \sum_k \int_{T^{-1}} \int_{E \cap I(k)} \rho(x)\,dm \\
&= \sum_k \int_{TI(k) \cap E} \rho\big(V(k)(x)\,\Delta(k)(x)\,dm \\
&= \sum_k \int_E \chi_{TI(k)}(x)\,\rho\big(V(k)(x)\big)\,\Delta(k)(x)\,dm \\
&= \int_E \sum_k \chi_{TI(k)}(x)\,\rho\big(V(k)(x)\big)\,\Delta(k)(x)\,dm \\
&= \int_E \sum_{x \in TI(k)} \rho\big(V(k)(x)\big)\,\Delta(k)(x)\,dm.
\end{aligned}$$

Summation und Integration dürfen dabei wegen des Satzes von der monotonen Konvergenz vertauscht werden. (W.z.b.w.)

Wir definieren nun eine neue Transformation $\hat{T}$ von I in sich:

$$\hat{T}x = T^n x, \qquad x \in B_n. \tag{20}$$

Aus (10) und (e) folgt, daß

$$I = \bigcup_{n=1}^{\infty} B_n \bmod 0 \tag{21}$$

und damit ist $\hat{T}$ für fast alle $x \in I$ definiert. Man sieht sofort, daß $T = \hat{T}$ genau dann gilt, wenn es nur eigentliche Zylinder gibt.

Lemma 5. *$\hat{T}$ hat die Eigenschaften* (a)–(e). *Es gibt eine Partition $\hat{\mathfrak{Z}}_1$ für $\hat{T}$, so daß nur eigentliche Zylinder auftreten.*

Beweis. Wir setzen

$$\hat{\mathfrak{Z}}_1 = \bigcup_{n=1}^{\infty} \mathfrak{B}_n. \tag{22}$$

$\hat{\mathfrak{Z}}_1$ ist dann eine Partition von I(mod 0) mit der in (b) geforderten Eigenschaft bezüglich $\hat{T}$. Außerdem sind alle diese Zylinder erster Ordnung eigentlich. Durch vollständige Induktion sieht man, daß auch die Zylinder höherer Ordnung eigentlich sind. Damit sind die Eigenschaften (a), (b) und (e) für $\hat{T}$ nachgewiesen. Sei $\hat{\mathfrak{Z}}_n$ die Menge der Zylinder n-ter Ordnung von $\hat{T}$ bezüglich $\hat{\mathfrak{Z}}_1$, dann zeigen wir weiter durch vollständige Induktion: Die Mengen von $\hat{\mathfrak{Z}}_n$ sind Mengen von $\mathfrak{Z}$ von der Form:

$$\hat{I}_n = I(k_1, \dots, k_s, l_1, \dots, l_t, m_1, \dots, m_u, \dots), \tag{23}$$

wobei $I(k_1, \dots, k_s) \in \mathfrak{B}_s$, $I(l_1, \dots, l_t) \in \mathfrak{B}_t$ $I(m_1, \dots, m_u) \in \mathfrak{B}_u$ und wobei n „Zifferngruppen" vorkommen, und die Transformation $\hat{T}^n$ stimmt auf $\hat{I}_n$ mit $T^{s+t+u\dots}$ überein.

Für $n=1$ ist dies per definitionem richtig. Für $\hat{I}_n \in \hat{\mathfrak{Z}}_n$ gilt nun:

$$\begin{aligned}\hat{I}_n &= \hat{T}^{-1}\hat{I}_{n-1} \cap \hat{I}_1 = T^{-s} I(l_1, \dots, l_t, m_1, \dots, m_u, \dots) \cap I(k_1, \dots, k_s) \\ &= I(k_1, \dots, k_s, l_1, \dots, l_t, m_1, \dots, m_u, \dots),\end{aligned}$$

womit alles gezeigt ist. Damit hat $\hat{T}$ auch die Eigenschaft (d). Wir zeigen nun:

$$\bigcup_{n=1}^{\infty} \hat{\mathfrak{Z}}_n = \mathfrak{A}. \tag{24}$$

Unter Beachtung von Lemma 1 folgt dann, daß $\hat{T}$ auch (c) erfüllt. Es ist klar, daß

$$\bigcup_{n=1}^{\infty} \hat{\mathfrak{Z}}_n \subseteq \mathfrak{A}$$

gilt. Ist andererseits $I(k_1, k_2, \dots, k_n) \in \mathfrak{A}$, dann gibt es ein s $(1 \leqq s \leqq n)$, so daß $I(k_1, \dots, k_s) \in \mathfrak{B}_s$ und $I(k_{s+1}, \dots, k_n) \in \mathfrak{A}$ gelten, da $I(k_1, k_2, \dots, k_n) = T^{-s} I(k_{s+1}, \dots, k_n) \cap I(k_1, \dots, k_s)$ ist. Wendet man denselben Schluß nun auf $I(k_{s+1}, \dots, k_n)$ an, usw., so sieht man, daß man die Ziffernsequenz $(k_1, \dots, k_n)$ in Stücke $(k_i, k_{i+1}, \dots, k_{i+k})$ zerlegen kann, wobei

$$I(k_i, k_{i+1}, \dots, k_{i+k}) \in \mathfrak{B}_{k-1}$$

gilt. Es ist also $I(k_1, \dots, k_n) \in \bigcup\limits_{n=1}^{\infty} \hat{\mathfrak{J}}_n$,

$$\bigcup_{n=1}^{\infty} \hat{\mathfrak{J}}_n \supseteq \mathfrak{A}$$

und Lemma 5 ist bewiesen.

Lemma 6. *Es gibt ein bezüglich* $\hat{T}$ *invariantes, zu* m *äquivalentes Wahrscheinlichkeitsmaß* $\hat{\mu}$, $\hat{\mu}(E) = \int\limits_E \hat{\rho}(x)\, dm$, $(E \in \mathfrak{F})$, *wobei*

$$C^{-1} \leqq \hat{\rho}(x) \leqq C \quad f.ü. \tag{25}$$

gilt.

Beweis. Der Beweis kann genauso geführt werden wie bei Schweiger [7]. Daß I dort der n-dimensionale reelle Einheitswürfel ist, geht in den Beweis nicht ein. (W.z.b.w.)

Bemerkung. Ist $\dfrac{dm\, T}{dm} = \text{const}$ (dies ist in den Beispielen 1, 2 und 3 der Fall), dann ist $\hat{\mu} = m$.

Mit Hilfe der Funktion $\hat{\rho}$ soll nun eine Dichtefunktion für ein bezüglich T invariantes, zu m äquivalentes Maß konstruiert werden. Wir setzen

$$\phi_0(x) = \hat{\rho}(x), \qquad \phi_n(x) = \sum_{I_n \in \mathfrak{D}_n} \sum_{x \in T^n I_n} \phi_0\big(V(I_n)(x)\big)\, \Delta(I_n)(x) \qquad (n = 1, 2, \dots) \tag{26}$$

wobei

$$V(I_n) = V(k_1, \dots, k_n), \qquad \Delta(I_n) = \Delta(k_1, \dots, k_n) \tag{27}$$

für $I_n = I(k_1, \dots, k_n)$ bedeutet.

Lemma 7. *Die Funktionen* $\phi_n(x)$ $(n = 0, 1, 2, \dots)$ *und*

$$\sigma(x) = \sum_{n=0}^{\infty} \phi_n(x) \tag{28}$$

sind integrierbar.

Beweis. Die Glieder der in (26) auftretenden Summe sind nichtnegativ, daher gilt:

$$\begin{aligned}\int_I \phi_n(x)\,dm &= \int_I \sum_{I_n\in\mathfrak{D}_n} \phi_0\big(V(I_n)(x)\big)\,\Delta(I_n)(x)\,dm\\ &= \int_I \sum_{I_n\in\mathfrak{D}_n} \chi_{T^n I_n}(x)\,\phi_0\big(V(I_n)(x)\big)\,\Delta(I_n)(x)\,dm\\ &= \sum_{I_n\in\mathfrak{D}_n} \int_I \chi_{T^n I_n}(x)\,\phi_0\big(V(I_n)(x)\big)\,\Delta(I_n)(x)\,dm\\ &= \sum_{I_n\in\mathfrak{D}_n} \int_{T^n I_n} \phi_0\big(V(I_n)(x)\big)\,\Delta(I_n)(x)\,dm = \sum_{I_n\in\mathfrak{D}_n} \int_{I_n} \phi_0\,dm\\ &= \int_{D_n} \hat\rho\,dm = \hat\mu(D_n) < \infty\end{aligned}$$

nach dem Satz von Levi, also ist ϕ_n integrierbar. Analog sieht man:

$$\int_I \sigma(x)\,dm = \int_I \sum_{n=0}^{\infty} \phi_n(x)\,dm = \sum_{n=0}^{\infty} \int_I \phi_n(x)\,dm = 1 + \sum_{n=1}^{\infty} \hat\mu(D_n),$$

da $\phi_n(x) \geqq 0$. Da $\hat\mu(D_n) \leqq C\,m(D_n)$ folgt nun aus (e) die zweite Behauptung des Lemmas. (W.z.b.w.)

Satz 3. *Das durch*

$$\mu_1(E) = \int_E \sigma(x)\,dm \qquad (E\in\mathfrak{F}) \tag{29}$$

auf $\mathfrak{F}$ definierte Maß ist invariant bezüglich T und äquivalent zu m.

Beweis. Aus $\sigma(x) \geqq \phi_0(x) \geqq C^{-1}$ und der Integrierbarkeit von $\sigma(x)$ folgt die Äquivalenz von m und μ_1. Wir zeigen, daß $\sigma(x)$ die Gl. (19) erfüllt, womit dann Satz 3 bewiesen ist. Für $n = 0, 1, 2, \ldots$ gilt:

$$\begin{aligned}&\sum_{x\in TI(k)} \phi_n\big(V(k)(x)\big)\,\Delta(k)(x)\\ &= \sum_{x\in TI(k)} \sum_{\substack{V(k)(x)\in T^n I(k_1,\ldots,k_n)\\ I(k_1,\ldots,k_n)\in\mathfrak{D}_n}} \phi_0\big(V(k_1,\ldots,k_n)(V(k)(x))\big)\\ &\quad\cdot \Delta(k_1,\ldots,k_n)\big(V(k)(x)\big)\,\Delta(k)(x)\\ &= \sum_k \sum_{I(k_1,\ldots,k_n)\in\mathfrak{D}_n} \chi_{TI(k)}(x)\,\chi_{T^n I(k_1,\ldots,k_n)}\big(V(k)(x)\big)\\ &\quad\cdot \phi_0\big(V(k_1,\ldots,k_n,k)(x)\big)\,\Delta(k_1,\ldots,k_n,k)(x)\\ &= \sum_{I(k_1,\ldots,k_n)\in\mathfrak{D}_n} \sum_k \chi_{T^{n+1} I(k_1,\ldots,k_n,k)}(x)\\ &\quad\cdot \phi_0\big(V(k_1,\ldots,k_n,k)(x)\big)\,\Delta(k_1,\ldots,k_n,k)(x)\\ &= \sum_{I(k_1,\ldots,k_{n+1})\in\mathfrak{D}_{n+1}} \chi_{T^{n+1} I(k_1,\ldots,k_n,k)}(x)\\ &\quad\cdot \phi_0\big(V(k_1,\ldots,k_{n+1})(x)\big)\,\Delta(k_1,\ldots,k_{n+1})(x)\\ &\quad+ \sum_{I(k_1,\ldots,k_{n+1})\in\mathfrak{B}_{n+1}} \phi_0\big(V(k_1,\ldots,k_{n+1})(x)\big)\,\Delta(k_1,\ldots,k_{n+1})(x)\\ &= \phi_{n+1}(x) + \sum_{I(k_1,\ldots,k_{n+1})\in\mathfrak{B}_{n+1}} \phi_0\big(V(k_1,\ldots,k_{n+1})(x)\big)\,\Delta(k_1,\ldots,k_{n+1})(x).\end{aligned}$$

Damit können wir weiterrechnen:

$$\sum_{x\in TI(k)} \sigma(V(k)(x))\,\Delta(k)(x) = \sum_{x\in TI(k)} \sum_{n=0}^{\infty} \phi_n(V(k)(x))\,\Delta(k)(x)$$

$$= \sum_{n=0}^{\infty} \sum_{x\in TI(k)} \phi_n(V(k)(x))\,\Delta(k)(x)$$

$$= \sum_{n=0}^{\infty} \phi_{n+1}(x) + \sum_{I(k_1,\ldots,k_{n+1})\in\mathfrak{B}_{n+1}} \phi_0(V(k_1,\ldots,k_{n+1})(x))\,\Delta(k_1,\ldots,k_{n+1})(x) \tag{30}$$

$$= \sum_{n=1}^{\infty} \phi_n(x) + \sum_{n=1}^{\infty} \sum_{I(k_1,\ldots,k_n)\in\mathfrak{B}_n} \phi_0(V(k_1,\ldots,k_n)(x))\,\Delta(k_1,\ldots,k_n)(x) \quad \text{f.ü.}$$

Dabei durfte die Summationsreihenfolge vertauscht werden, da die betrachtete Doppelreihe f.ü. absolut konvergent ist (dies folgt aus der Integrierbarkeit von $\sigma(x)$).

Wir müssen noch zeigen, daß

$$\sum_{n=1}^{\infty} \sum_{I(k_1,\ldots,k_n)\in\mathfrak{B}_n} \phi_0(V(k_1,\ldots,k_n)(x))\,\Delta(k_1,\ldots,k_n)(x) = \phi_0(x) \quad \text{f.ü.} \tag{31}$$

ist. Dies ist aber gerade die Gl. (19) angewandt auf die Funktion $\hat{\rho} = \phi_0$ und die Transformation $\hat{T}$ mit der Partition $\hat{\mathfrak{Z}}_1$. Da $\hat{\rho}$ Dichte eines bezüglich $\hat{T}$ invarianten Maßes ist, gilt also (31) und damit

$$\sum_{x\in TI(k)} \sigma(V(k)(x))\,\Delta(k)(x) = \sum_{n=1}^{\infty} \phi_n(x) + \phi_0(x) = \sigma(x) \quad \text{f.ü.} \quad \text{(w.z.b.w.)}$$

Von nun an sei

$$\mu(E) = [\mu_1(I)]^{-1}\,\mu_1(E) \tag{32}$$

das normierte invariante zu m äquivalente Wahrscheinlichkeitsmaß auf I. Aus der Ergodizität von T folgt, daß μ das einzige derartige Maß ist.

4. Die Exaktheit von T

T ist ein Endomorphismus, d.h. maßtreue Transformation, des Wahrscheinlichkeitsraumes $(I, \mathfrak{F}, \mu)$. Nach Rohlin [6] heißt ein Endomorphismus exakt, wenn die σ-Algebra

$$\bigcap_{n=0}^{\infty} T^{-n}\,\mathfrak{F}$$

nur aus Mengen von Maß 0 oder 1 besteht.

Satz 4. *T ist ein exakter Endomorphismus.*

Beweis. Aus Lemma 4 und der Äquivalenz von m und μ folgt

$$\limsup_{n\to\infty} \mu(T^n Y) = 1 \tag{33}$$

für alle $Y \in \mathfrak{F}$ mit $\mu(Y) > 0$. Da die Folge $\mu(Y), \mu(TY), \ldots$ monoton wächst, gilt sogar

$$\lim_{n\to\infty} \mu(T^n Y) = 1. \tag{34}$$

Dies ist nach [6] gleichbedeutend mit der Exaktheit von T. (W.z.b.w.)

Folgerung. *T ist mischend von jedem Grad* (s. Rohlin [6]).

5. Die Entropie von T

Zur Definition der Entropie $h(T)$ der Transformation T verweise ich auf [3]. Dort ist auch der folgende Satz zu finden:

Satz von MacMillan. *Ist*

$$-\sum_{I_1 \in \mathfrak{J}_1} \mu(I_1) \log \mu(I_1) < \infty, \tag{36}$$

dann gilt

$$h(T) = -\lim_{n\to\infty} \frac{1}{n} \log \mu(I_n(x)) \quad \textit{f.ü.} \tag{37}$$

Wir setzen $\rho(x) = \dfrac{d\mu}{dm}(x)$, dann gilt:

$$\rho(x) = \tau^{-1} \sigma(x) \geqq \tau^{-1} C^{-1} > 0 \quad \text{f.ü.} \quad \text{mit } \tau = \mu_1(I). \tag{38}$$

Lemma 8. *$\log \dfrac{dm\, T}{dm}$ ist genau dann integrierbar, wenn $\log \dfrac{d\mu\, T}{d\mu}$ integrierbar ist. Ist eine der beiden Funktionen integrierbar, dann gilt:*

$$\int_I \log \frac{dm\, T}{dm}\, d\mu = \int_I \log \frac{d\mu\, T}{d\mu}\, d\mu. \tag{39}$$

Weiter gilt: Ist

$$\inf_{I_1 \in \mathfrak{J}_1} m(T I_1) > 0 \tag{40}$$

und ist $\rho(x)$ beschränkt, dann ist (36) *gleichbedeutend mit der Integrierbarkeit von $\log \dfrac{dm\, T}{dm}$.*

Beweis. Aus der Kettenregel für Ableitungen folgt

$$\frac{d\mu\, T}{d\mu}(x) = \frac{\rho(Tx)}{\rho(x)} \frac{dm\, T}{dm}(x). \tag{41}$$

Unter Beachtung von (38) gilt ferner

$$\log \tau^{-1} C^{-1} \leqq \log \rho(x) \leqq \rho(x),$$

weshalb $\log \rho(x)$ integrierbar ist. Mit Rücksicht auf die Invarianz von μ folgt (39), womit der erste Teil von Lemma 8 bewiesen ist. Für die zweite Behauptung sei auf [9], Lemma 5.1, p. 89 verwiesen. (W.z.b.w.)

Satz 5. *Gilt* (36) *und ist* $\log \frac{dm\,T}{dm}$ *integrierbar, dann ist*

$$h(T) = -\lim_{n\to\infty} \frac{1}{n} \log \mu(I_n(x)) \quad f.ü., \tag{42}$$

$$= -\lim_{n\to\infty} \frac{1}{n} \log m(I_n(x)) \quad f.ü., \tag{43}$$

$$= \int_I \log \frac{dm\,T}{dm}\, d\mu, \tag{44}$$

$$= \int_I \log \frac{d\mu\,T}{d\mu}\, d\mu. \tag{45}$$

Bemerkung. Bei allen im Abschnitt 1 angeführten Beispielen sind die Voraussetzungen von Satz 6 erfüllt, da dort immer $\log \frac{dm\,T}{dm}$ integrierbar und $\rho(x)$ beschränkt ist und (40) gilt. Die Entropieformeln (44) und (45) folgen auch aus einem Ergebnis von Parry [4], wenn man dort die Voraussetzung der Endlichkeit der Partition $\mathfrak{J}_1$ durch die Bedingung (36) ersetzt.

Beweis. (42) folgt aus dem Satz von MacMillan, (45) aus (44) und (39). Es genügt also, (43) und (44) zu beweisen. Sei $E(f|\mathfrak{P})(x)$ die bedingte Erwartung der integrierbaren Funktion f bezüglich der Partition $\mathfrak{P}$ (s. [3]), dann gilt

$$E(\rho|\mathfrak{J}_n)(x) = \frac{\mu(I_n(x))}{m(I_n(x))} \tag{46}$$

und aus dem Martingalsatz ([3]) folgt nun

$$\lim_{n\to\infty} \frac{\mu(I_n(x))}{m(I_n(x))} = \rho(x) \quad \text{f.ü.} \tag{47}$$

Wegen $0 < \rho(x) < \infty$ f.ü. ergibt sich daraus

$$-\lim_{n\to\infty} \frac{1}{n} \log \mu(I_n(x)) = -\lim_{n\to\infty} \frac{1}{n} \log m(I_n(x)) \quad \text{f.ü.}, \tag{48}$$

womit (43) gezeigt ist. Lemma 2 lehrt, daß für fast alle $x \in I$ eine Folge (l_i), die von x abhängt, existiert mit $I_{l_i}(x) \in \mathfrak{A}$. Aus (d) folgt

$$C^{-1} \left[\frac{dm\,T^{l_i}}{dm}(x)\right]^{-1} \leqq m(I_{l_i}(x)) \leqq C \left[\frac{dm\,T^{l_i}}{dm}(x)\right]^{-1} \tag{49}$$

und aus der Kettenregel für Derivierte

$$\frac{dm\,T^{l_i}}{dm}(x)=\prod_{j=0}^{l_i-1}\frac{dm\,T}{dm}(T^j x). \tag{50}$$

(49) und (50) liefern

$$\begin{aligned}-\log C+\sum_{j=0}^{l_i-1}\log\frac{dm\,T}{dm}(T^j x)&\leqq -\log m\big(I_{l_i}(x)\big)\leqq\\ &\leqq \log C+\sum_{j=0}^{l_i-1}\log\frac{dm\,T}{dm}(T^j x).\end{aligned} \tag{51}$$

Der individuelle Ergodensatz gibt nun

$$-\lim_{i\to\infty}\frac{1}{l_i}\log m\big(I_{l_i}(x)\big)=\int_I\log\frac{dm\,T}{dm}\,d\mu \quad \text{f.ü.}, \tag{52}$$

womit wegen der Konvergenz von $\frac{1}{n}\log m\big(I_n(x)\big)$ Satz 6 bewiesen ist.

Folgerung. *Sei $A(x,n)$ die Anzahl der eigentlichen Zylinder unter den Zylindern $I_1(x), I_2(x), \ldots, I_n(x)$. Dann gilt:*

$$\lim_{n\to\infty}\frac{1}{n}A(x,n)=\frac{h(T)}{h(\hat{T})} \quad \textit{f.ü.}$$

Beweis. Satz 6 lehrt

$$h(\hat{T})=-\lim_{n\to\infty}\frac{1}{n}\log m\big(\hat{I}_n(x)\big) \quad \text{f.ü.}$$

Auf Grund des Beweises von Lemma 5 ist klar, daß $\hat{I}_n(x)=I_{l_n}(x)$ ist, wenn $I_{l_n}(x)$ den n-ten eigentlichen Zylinder in der Folge $I_1(x), I_2(x), \ldots$ bezeichnet. Daher ist $A(x,l_n)=n$ und da außerdem

$$h(T)=-\lim_{n\to\infty}\frac{1}{l_n}\log m\big(I_{l_n}(x)\big) \quad \text{f.ü.}$$

gilt, ist alles gesagt. (W.z.b.w.)

Bemerkung. Ist $\frac{dm\,T}{dm}=\text{const.}$ (Beispiele 1, 2 und 3), dann gilt

$$h(\hat{T})=\Big(1+\sum_{n=1}^{\infty}m(D_n)\Big)\,h(T).$$

Literatur

1. Cigler, J.: Ziffernverteilung in ϑ-adischen Brüchen. Math. Z. **75**, 8–13 (1961).
2. Fischer, F.: Ergodische Eigenschaften komplexer Ziffernentwicklungen. Sitzungsber. der Österr. Akad. d. Wiss., math.-naturw. Kl., Abt. II, **180**, 49–68 (1972).
3. Parry, W.: Entropy and generators in ergodic theory. New York and Amsterdam: Benjamin Inc. 1969.
4. Parry, W.: On Rohlin's formula for entropy. Acta math. Acad. Sci. Hungar. **15**, 107–113, (1964).
5. Philipp, W.: Mischungseigenschaften gewisser, auf dem Torus definierter Endomorphismen. Math. Z. **101**, 369–374 (1967).
6. Rohlin, V.A.: Exact endomorphisms of a Lebesgue space. Izvestija Akad. Nauk. SSSR, Ser. mat. **24**, 499–530 (1960); Amer. math. Soc. Translat. II. Ser. **39**, 1–36 (1964).
7. Schweiger, F.: Metrische Theorie einer Klasse zahlentheoretischer Transformationen. Acta Arithmetica **15**, 1–18 (1968); Corrigendum, Acta Arithmetica **16**, 217–219 (1969).
8. Schweiger, F.: Mischungseigenschaften und Entropie beim Jacobischen Algorithmus. J. reine angew. Math. **229**, 50–56 (1968).
9. Waterman, M.S.: Some ergodic properties of multi-dimensional f-expansions. Z. Wahrscheinlichkeitstheorie verw. Geb. **16**, 77–103 (1970).

Dr. Roland Fischer
Mathematisches Institut der Universität
5020 Salzburg, Ferdinand-Porsche-Straße 1/I
Österreich

(Eingegangen am 18. Februar 1972)

Math. Z. 128, 231–233 (1972)

Remarques sur l'application $\mathrm{Hom}(G, H) \to [B_G, B_H]$

Daniel Lehmann et Renzo Piccinini

Soient G et H deux groupes topologiques et $\mathrm{Hom}(G, H)$ l'ensemble des classes de conjugaison d'homomorphismes continus de G dans H (deux homomorphismes ρ, ρ': $G \to H$ étant conjugués s'il existe $h \in H$ tel que, $\forall g \in G$, $\rho'(g) = h \cdot \rho(g) \cdot h^{-1}$). On note $[B_G, B_H]$ l'ensemble des classes d'homotopie libres d'applications continues entre espaces classifiants, et $\alpha_{G,H}$: $\mathrm{Hom}(G, H) \to [B_G, B_H]$ l'application induite par le foncteur B. (Si $E_G \to B_G$ désigne un G-fibré principal universel et $\dot{\rho}$ la classe de conjugaison d'un homomorphisme continu ρ: $G \to H$, $\alpha_{G,H}(\dot{\rho})$ désigne encore la classe d'homotopie d'une application classifiante du H-fibré principal $H \underset{\rho}{\times} E_G \to B_G$ associé à E_G par l'action ρ de G sur H.)

Nous nous posons le problème de savoir si $\alpha_{G,H}$ est ou non surjectif (resp. injectif ou bijectif).

1. Rappel de quelques exemples

(i) cas où G *est fini* (muni de la topologie discrète), et $H = S^1$: α_{G,S^1} est alors *bijectif*.

En effet

$$\mathrm{Hom}(G, S^1) \cong H^2(G, \mathbb{Z}),$$
$$[B_G, B_{S^1}] \cong H^2(B_G, \mathbb{Z}),$$

et α_{G,S^1} n'est autre que l'identification entre la cohomologie d'un groupe fini et celle de son espace classifiant.

(ii) cas où *G et H sont tous deux discrets*: $\alpha_{G,H}$ est alors *bijectif*.

[C'est une conséquence immédiate du théorème de classification des revêtements galoisiens.]

(iii) cas $G = \mathbb{Z}$, *H groupe de Lie*: $\alpha_{\mathbb{Z},H}$ est alors *surjectif*.

En effet, on peut prendre $B_{\mathbb{Z}} = S^1$; tout fibré de base S^1 et de groupe structural H est isomorphe à un fibré C^∞, lequel admet une connexion; une telle connexion est certainement sans courbure puisque $\dim S^1 = 1$: c'est dire qu'un tel fibré appartient à l'image de $\alpha_{\mathbb{Z},H}$.

Il est facile de se convaincre que $\alpha_{\mathbb{Z},H}$ n'est pas *nécessairement injectif* (par exemple, dans le cas $H = \mathbb{R}$, le fibré trivial $\mathbb{R} \times S^1 \to S^1$ peut être obtenu comme image par $\alpha_{\mathbb{Z},\mathbb{R}}$ d'un homomorphisme non trivial).

(iv) cas $G = \Pi_g$ (groupe fondamental d'une surface compacte orientable de genre $g \geqq 1$, muni de la topologie discrète), $H = GL^+(2, \mathbb{R})$ *ou*

$SO(2)$. Il résulte d'un théorème de Milnor [7] que $\alpha_{\Pi_g, GL^+(2,\mathbb{R})}$ n'est *jamais surjectif*. Il en est à fortiori ainsi de $\alpha_{\Pi_g, SO(2)}$ [pour $g=1$,

$$\alpha_{\mathbb{Z}\times\mathbb{Z}, GL^+(2,\mathbb{R})} \quad \text{et} \quad \alpha_{\mathbb{Z}\times\mathbb{Z}, SO(2)}$$

sont même des applications constantes, et donc, en particulier, non injectives].

2. Cas où G est un groupe de Lie compact (non nécessairement connexe) et où $H = U = \varinjlim_n U(n)$ (groupe unitaire infini)

Proposition 1. *$\alpha_{G,U}$ n'est jamais surjectif.*

Notons $R(G)$ l'anneau des caractères unitaires de G,

$$I(G) = \operatorname{Ker}(R(G) \to \mathbb{Z})$$

l'idéal des caractères de dimension 0,

$$\hat{I}(G) = \varprojlim_n I(G)/(I(G))^n$$

le complété de $I(G)$ pour la topologie $I(G)$-adique, et $\iota\colon I(G) \to \hat{I}(G)$ l'application canonique.

On a une application injective évidente

$$\varinjlim_n \operatorname{Hom}(G, U(n)) \to \operatorname{Hom}(G, U).$$

Puisque G est compact,

$$\varinjlim_n \operatorname{Hom}(G, U(n)) = \operatorname{Hom}(G, U),$$

d'où une application injective

$$\sigma\colon \operatorname{Hom}(G, U) \rightarrowtail I(G).$$

Notant $(B_G)_r$ le r-squelette d'une réalisation de B_G comme CW-complexe à squelettes finis, on a un diagramme commutatif

$$\begin{tikzcd}
{[B_G, B_{U(n)}]} \arrow[r] & {[B_G, B_U]} \arrow[r, tail, two heads, "\rho"] & \varprojlim_r [(B_G)_r, B_U] \\
 & & \hat{I}(G) \arrow[u, two heads, tail, "\hat{\alpha}_{G,U}"'] \\
\operatorname{Hom}(G, U(n)) \arrow[r, "\theta_n"] \arrow[uu, "\alpha_{G,U(n)}"'] & \operatorname{Hom}(G, U) \arrow[r, "\sigma"] \arrow[uu, "\alpha_{G,U}"'] & I(G) \arrow[u, "\iota"']
\end{tikzcd}$$

où $\hat{\alpha}_{G,U}$ désigne l'isomorphisme défini par Atiyah-Hirzebruch ([1–3]), et où ρ désigne l'application naturelle induite, pour tout CW-complexe X à squelettes finis, par les restrictions $[X, BU] \to [X_r, BU]$. On sait que, pour $X = B_G$ (G groupe de Lie compact connexe), ρ est un isomorphisme (Buhstaber et Miscenko [4]).

Or ι n'est jamais surjective, donc $\alpha_{G,U}$ non plus, C.Q.F.D.

Sur le même diagramme, on vérifie immédiatement la

Proposition 2. *Si, en outre, l'application ι est injective (par exemple si G est un p-groupe ou si G est connexe), $\alpha_{G,U}$ est alors injective, de même que $\alpha_{G,U(n)}$ $(n \geqq 1)$.*

3. Remarques finales

a) Notons $\alpha^r_{G,H}$ l'application composée

$$\mathrm{Hom}(G, H) \xrightarrow[\alpha_{G,H}]{} [B_G, B_H] \to [(B_G)_r, B_H]$$

où B_G a été muni d'une structure de CW-complexe, où $(B_G)_r$ désigne son r-squelette, et où la flèche de droite désigne la restriction à $(B_G)_r$ des applications $B_G \to B_H$.

Les résultats du § 2 sont en contraste, lorsque G est un groupe fini à cohomologie périodique, avec le fait que les applications $\alpha^r_{G,U}$ sont toujours surjectives, de même que les applications $\alpha^r_{G,U(n)}$ pour $n \gg r$ (cf. Kamber-Tondeur [5], Lehmann [6]); tandis que pour G fini, $\alpha^r_{G,U(n)}$ n'est jamais injective dès que $n \gg r$ (cf. [6]).

b) Le fait que $\alpha_{G,U}$ ne soit pas surjectif suggère la *conjecture:*

Si G est un groupe de Lie compact, $\alpha_{G,U(n)}$ n'est pas surjectif pour $n \gg 0$.

Bibliographie

1. Atiyah, M.: Characters and cohomology of finite groups. Inst. Hautes Études Sci., Publ. Math. **9**, 23–64 (1961).
2. Atiyah, M., Hirzebruch, F.: Vector bundles and homogeneous spaces. Amer. Math. Soc., Proc. Sympos. Pure Math. **3**, 7–38 (1962).
3. Atiyah, M., Segal, G.: Equivariant K-theory and completion. J. Diff. Geom. **3**, 1–18 (1969).
4. Buhstaber, V., Miscenko, A.: K-theory on the category of infinite cell complexes. Math. USSR – Izvestija **2**, n^0 3, 515–556 (1968).
5. Kamber, F., Tondeur, P.: Flat manifolds. Lecture Notes in Math., vol. 67. Berlin-Heidelberg-New York: Springer 1969.
6. Lehmann, D.: Connexions à coubure nulle et K-théorie. Anais Acad. Bras. Ciências **40**, 1–6 (1968).
7. Milnor, J.: On connexions with zero curvature. Commentarii Math. Helvet. **32**, 215–223 (1957).

Daniel Lehmann
Département de Mathématiques
Université des Sciences et Techniques de Lille
Lille
France

Renzo Piccinini
Forschungsinstitut für Math.
Eidg. Technische Hochschule
Zürich
Schweiz

(Reçu le 13 avril 1972)

Math. Z. 128, 235–243 (1972)

Über die cohomologische Dimension der auflösbaren Gruppen

Robert Bieri

1. Einleitung

Gruenberg [5, § 8.8, Theorem 5] hat gezeigt, daß die cohomologische Dimension cdG einer *lokal nilpotenten* Gruppe G wie folgt mit der Hirschzahl hG von G verknüpft ist: cdG ist genau dann endlich, wenn G torsionsfrei, nilpotent und von endlicher Hirschzahl hG ist. In diesem Fall gilt

$$(1.1)\qquad \begin{aligned} &cdG = hG, \text{ wenn } G \text{ endlich erzeugt ist},\\ &cdG = hG+1, \text{ wenn } G \text{ nicht endlich erzeugbar ist.}\end{aligned}$$

Darüber hinaus gilt $cdG = hG$ nicht nur für jede endlich erzeugte, torsionsfreie nilpotente Gruppe, sondern allgemeiner für jede torsionsfreie, polyzyklische Gruppe G ([5, § 8.8, Lemma 8], vergleiche auch [2, Satz 3.1.2]).

In der vorliegenden Arbeit stellen wir uns die Frage, wie das Resultat von Gruenberg auf *lokal auflösbare* Gruppen zu verallgemeinern ist. Das ist bis zu einem gewissen Grad nicht schwierig. Es gilt: (3.1) *Die cohomologische Dimension cdG einer lokal auflösbaren Gruppe G ist genau dann endlich, wenn G torsionsfrei, auflösbar und von endlicher Hirschzahl hG ist. In diesem Falle ist cdG entweder gleich hG oder gleich $hG+1$.* Die Frage nach einem gruppentheoretischen Kriterium für $cdG = hG < \infty$ scheint aber nicht leicht zu sein. An Hand von zwei einfachen Gegenbeispielen zeigen wir vorerst: (4.1) *Es gibt auflösbare, nicht-polyzyklische Gruppen G mit $cdG = hG < \infty$.* (4.2) *Es gibt endlich erzeugte auflösbare Gruppen G mit $cdG = hG+1 < \infty$.*

Anschließend beweisen wir das folgende Hauptresultat: (5.2) *Sei G eine torsionsfreie Gruppe, sei $N \trianglelefteq G$ ein nilpotenter Normalteiler von endlicher Hirschzahl $hN = n$ und mit frei-abelscher Faktorgruppe G/N von endlichem Rang $h(G/N) = m$. Dann gilt: Der Funktor $H^{n+m+1}(G, -)$ verschwindet genau dann auf allen G/N-Moduln, wenn $H_n(N, \mathbb{Z})$ ein zyklischer G/N-Modul ist* (G/N operiert durch Konjugation auf $H_n(N, \mathbb{Z})$). Da dieses Resultat auch dann noch richtig ist, wenn die Faktorgruppe G/N eine frei-abelsche Untergruppe vom Rang m mit endlichem Index enthält, liefert es eine notwendige Bedingung für $cdG = hG$. Ein Korollar davon ist: (6.3) *Eine auflösbare Gruppe G mit* $cdG = hG < \infty$ *ist eine Minimax-Gruppe* (im Sinne von Robinson [7]).

Zum Beweis von (3.1) und (5.2) verwenden wir zwei Hilfsresultate aus einer früheren Arbeit [2]. Die Gruppe (4.2) ist nicht endlich präsentierbar, und überdies ist jede torsionsfreie, auflösbare, endlich erzeugte Gruppe mit endlicher Hirschzahl eine Minimax-Gruppe. Daher bleibt die Frage offen, ob die auflösbaren Gruppen G mit $cdG = hG < \infty$ genau die endlich präsentierbaren sind.

2. Bezeichnungen

Der Rang $r(G)$ (im Sinne von Prüfer) einer Gruppe G ist die kleinste positive ganze Zahl mit der Eigenschaft, daß jede endlich erzeugte Untergruppe von G durch $r(G)$ Elemente erzeugt werden kann. Wenn es keine solche Zahl gibt, dann ist $r(G) = \infty$.

Der torsionsfreie Rang hA einer abelschen Gruppe A ist der Rang von A/tA, wo tA die Torsionsuntergruppe von A bezeichnet.

Eine Gruppe G heißt auflösbar, wenn sie eine (endliche) Normalreihe mit abelschen Faktoren besitzt. Die Summe hG der torsionsfreien Ränge dieser Faktoren (ev. $hG = \infty$) heißt die Hirschzahl von G. hG ist eine Invariante von G. Nach einem Resultat von Zaičev [9] ist $hG = r(G)$, wenn G eine (endliche) Normalreihe mit torsionsfreien, abelschen Faktoren besitzt.

Die Gruppe G heißt polyzyklisch, wenn sie eine (endliche) Normalreihe mit zyklischen Faktoren besitzt.

Das Hirsch-Plotkin-Radikal R einer Gruppe G ist der eindeutig bestimmte, maximale lokal nilpotente Normalteiler von G.

$I(G)$ ist das Augmentationsideal im Gruppenring $\mathbb{Z}G$. Ist A ein G-Modul, dann bezeichnet A^G die Fixpunkte von A unter G und A_G die Faktorgruppe $A/I(G)A$.

Rationale Gruppen. Q bezeichnet die additive Gruppe der rationalen Zahlen. Eine Gruppe heißt rational, wenn sie isomorph zu einer Untergruppe von Q ist. Die rationalen Gruppen können wie folgt klassifiziert werden: Man betrachtet die Menge $\mathfrak{M}$ der formalen Produkte $\prod p^{\alpha_p}$ über alle Primzahlen p wobei α_p entweder eine nichtnegative ganze Zahl oder ∞ ist. Zwei solche Produkte seien äquivalent, wenn sie sich nur in endlichvielen endlichen Exponenten α_p unterscheiden, und sei $\mathfrak{M}^*$ die Menge der dadurch definierten Äquivalenzklassen $[\prod p^{\alpha_p}]$. Sei $L \leqq Q$ und sei e die kleinste positive ganze Zahl in L. Für jede Primzahl p sei $\alpha_p = \sup\{k \in \mathbb{Z} \mid e p^{-k} \in L\}$. Die Klasse $L^* = [\prod p^{\alpha_p}] \in \mathfrak{M}^*$ heißt der Typ von L. Die Zuordnung $L \leftrightarrow L^*$ liefert eine eineindeutige Korrespondenz zwischen $\mathfrak{M}^*$ und den Isomorphietypen der rationalen Gruppen.

Es sei L eine rationale Gruppe, $L^* = [\prod p^{\alpha_p}]$ und sei o.B.d.A. $1 \in L$. Jeder Automorphismus $\varphi\colon L \to L$ ist gegeben durch $\varphi(1) = a/b$, a und b

ganze, teilerfremde Zahlen. Wir bezeichnen mit $\pi(\varphi)$ die Menge aller positiven Primteiler von ab. Ist A eine Gruppe von Automorphismen von L, dann sei $\pi(A)=\bigcup_{\varphi\in A}\pi(\varphi)$. Offensichtlich gilt

$$\pi(\mathrm{Aut}(L))=\{p\,|\,\alpha_p=\infty\}.$$

Ist n eine positive ganze Zahl, dann bezeichne $Q^{(n)}$ die additive Gruppe aller rationalen Zahlen deren Nenner eine Potenz von n ist. Es gilt $Q^{(n)*}=[p_1^\infty\, p_2^\infty \dots p_s^\infty]$, mit $\pi(n)=\{p_1, p_2, \dots, p_s\}$.

3. Homologische und cohomologische Dimension

Sei G eine Gruppe. Die cohomologische Dimension cdG (bzw. die homologische Dimension hdG) von G ist die größte ganze Zahl n mit der Eigenschaft, daß es einen G-Modul A mit $H^n(G,A)\neq 0$ (bzw. $H_n(G,A)\neq 0$) gibt. Wenn keine solche Zahl existiert, dann schreiben wir $cdG=\infty$ (bzw. $hdG=\infty$). Für jede Gruppe G gilt $cdG\geqq hdG$; ob es aber Gruppen mit $cdG>hdG+1$ gibt, ist unseres Wissens eine offene Frage. Für jede torsionsfreie, auflösbare Gruppe G gilt $hdG=hG$ [2, Lemma 3.3.3].

(3.1) **Satz.** *Die cohomologische Dimension cdG einer lokal auflösbaren Gruppe G ist genau dann endlich, wenn G torsionsfrei, auflösbar und von endlicher Hirschzahl hG ist. In diesem Falle ist*

$$hG\leqq cdG\leqq hG+1.$$

Beweis. Sei G lokal auflösbar mit $cdG<\infty$. Der Rang der abelschen Untergruppen ist durch cdG beschränkt, also ist G nach einem Resultat von Merzljakov [6] von endlichem Rang. G ist trivialerweise torsionsfrei. Nach Čarin [3] ist aber jede torsionsfreie, lokal auflösbare Gruppe von endlichem Rang auflösbar. Es folgt $hG=hdG\leqq cdG<\infty$. Sei umgekehrt G torsionsfrei, auflösbar und $hG<\infty$, und sei R das Hirsch-Plotkin-Radikal von G. Nach Baer-Heineken [1, Prop. 5.5] ist R nilpotent und G/R enthält eine endlich erzeugte, frei-abelsche Untergruppe G_1/R von endlichem Index. Nach dem „Maximumprinzip" folgt aus der Spektralreihe von Lyndon-Hochschild-Serre $cdG_1\leqq cd(G_1/R)+cdR$, also gilt nach (1.1) $cdG_1\leqq h(G_1/R)+hR+1=hG+1$. Dank dem Satz von Serre ist damit auch $cdG\leqq hG+1<\infty$.

4. Gegenbeispiele

(4.1) *Sei n eine ganze Zahl, $n>1$. Dann ist die Gruppe*

$$G=\langle x, y;\, y^{-1}xy=x^n\rangle$$

auflösbar aber nicht polyzyklisch, und es gilt $cdG=hG=2$.

Das ist fast trivial. Sei N der von x erzeugte Normalteiler. Offensichtlich ist $N \cong Q^{(n)}$ und $G/N \cong \mathbb{Z}$, also ist G auflösbar, torsionsfrei, nicht polyzyklisch und $hG=2$. Weil G eine 1-Relationen-Gruppe ist, ist auch $cdG=2$.

(4.2) *Sei S unendlich-zyklisch, von $s \in S$ erzeugt; sei $n>1$ eine ganze Zahl. Operiert s auf $N=Q^{(n)} \times Q^{(n)}$ durch $s \circ (a,b)=(na, n^{-1}b)$, $(a,b) \in N$, dann ist das semidirekte Produkt $G=S \sqsubset N$ eine torsionsfreie, endlich erzeugte, auflösbare Gruppe mit $cdG=hG+1=4$. G ist aber nicht endlich präsentierbar.*

Beweis. Offensichtlich ist G torsionsfrei, auflösbar mit $hG=3$ und wird von den Elementen $s \sqsubset (0,0)$, $0 \sqsubset (1,0)$, $0 \sqsubset (0,1)$ erzeugt. Wir berechnen nun die ganzzahlige Homologiegruppe $H_3(G,\mathbb{Z})$. Wegen $hS=1$ und $hN=2$ folgt mit dem „Maximumprinzip" aus der Spektralreihe von Lyndon-Hochschild-Serre:

$$H_3(G,\mathbb{Z}) \cong H_1(S, H_2(N,\mathbb{Z})).$$

Die ganzzahlige Homologie eines direkten Produkts kann mit Hilfe der Künnethformel berechnet werden:

$$H_2(N,\mathbb{Z}) \cong H_1(Q^{(n)},\mathbb{Z}) \otimes H_1(Q^{(n)},\mathbb{Z}) = Q^{(n)} \otimes Q^{(n)} = Q^{(n)}.$$

Dabei operiert s auf $Q^{(n)} \otimes Q^{(n)}$ wie $s \circ (a \otimes b) = na \otimes n^{-1} b = a \otimes b$. $H_2(N,\mathbb{Z})$ ist also isomorph zum trivialen S-Modul $Q^{(n)}$, somit ist $H_3(G,\mathbb{Z}) \cong H_1(S, Q^{(n)}) \cong Q^{(n)}$. Nun folgt aus dem Universelle-Koeffizienten-Theorem

$$H^4(G,\mathbb{Z}) \cong \operatorname{Ext}(H_3(G,\mathbb{Z}),\mathbb{Z}) \cong \operatorname{Ext}(Q^{(n)},\mathbb{Z}) \neq 0,$$

also $cdG \geqq 4$. Nach (3.1) gilt daher $cdG=4$.

Da $H_1(S, H_1(N,\mathbb{Z})) \cong N^S = 0$ ist, folgt aus der Lyndonspektralreihe ferner: $H_2(G,\mathbb{Z}) \cong H_0(S, H_2(N,\mathbb{Z})) \cong H_0(S, Q^{(n)}) \cong Q^{(n)}$. G kann daher nicht endlich präsentierbar sein.

5. Das Hauptresultat

Sei G eine torsionsfreie Gruppe, sei $N \trianglelefteq G$ ein nilpotenter Normalteiler von endlicher Hirschzahl $hN=n$ und mit frei-abelscher Faktorgruppe G/N vom Rang $h(G/N)=m<\infty$, $hG=m+n=h$. In [2, Lemma 3.3.2] haben wir gezeigt: Die Homologiegruppe $L=H_n(N,\mathbb{Z})$ ist rational, und sie ist genau dann zyklisch (als abelsche Gruppe), wenn N endlich erzeugbar ist.

(5.1) **Lemma.** *Wenn es eine rationale Gruppe K mit $\operatorname{Ext}(L,K) \neq 0$ und $\pi(\operatorname{Im}(G \to \operatorname{Aut}(L))) \subset \pi(\operatorname{Aut}(K))$ gibt, dann existiert ein G/N-Modul A mit $H^{h+1}(G,A) \neq 0$.*

Beweis. Die zweite Voraussetzung erlaubt, K mit einer G/N-Struktur zu versehen, indem man ein Element $g \in G$ auf K und auf L durch Multiplikation mit derselben rationalen Zahl r_g operieren läßt. Wegen $\mathrm{Ext}(L, K) \neq 0$ ist L nicht zyklisch, also N nicht endlich erzeugbar, also $cdN = n+1$. Nach dem „Maximumprinzip" folgt aus der Spektralreihe von Lyndon-Hochschild-Serre

$$H^{h+1}(G, K) \cong H^m(G/N, H^{n+1}(N, K)).$$

N operiert trivial auf K, also ist $H^{n+1}(N, K) \cong \mathrm{Ext}(L, K)$. Dabei operiert $g \in G$ auf $e \in \mathrm{Ext}(L, K)$ wie $g \circ e = \mathrm{Ext}(r_g^{-1}, r_g)(e) = r_g^{-1} r_g \mathrm{Ext}(\mathrm{Id}, \mathrm{Id})(e) = e$, d.h. $H^{n+1}(N, K)$ ist ein trivialer G/N-Modul. Da G/N frei-abelsch vom Rang m ist folgt daraus $H^{h+1}(G, K) \cong H^{n+1}(N, K) \cong \mathrm{Ext}(L, K) \neq 0$, was zu beweisen war.

(5.2) **Satz.** *Sei G eine torsionsfreie Gruppe, sei $N \trianglelefteq G$ ein nilpotenter Normalteiler mit $hN = n < \infty$ und frei-abelscher Faktorgruppe G/N vom Rang $h(G/N) = m < \infty$, $hG = m + n = h$. Dann sind die folgenden drei Aussagen äquivalent:*

(i) $H^{h+1}(G, A) = 0$, *für jeden G/N-Modul A,*

(ii) $H_n(N, \mathbb{Z}) = L$ *ist zyklisch als G/N-Modul,*

(iii) L *ist rational vom Typ* $[p_1^\infty p_2^\infty \dots p_s^\infty]$, $0 \leqq s < \infty$ *und*

$$\pi(\mathrm{Im}(G \to \mathrm{Aut}(L))) = \{p_1, p_2, \dots, p_s\}.$$

Beweis. Die Implikation (ii) $\Rightarrow$ (iii) ist evident: Wenn L von (o.B.d.A.) $1 \in L$ als G/N-Modul erzeugt wird, dann treten im Nenner der Zahlen $r \in L$ genau die Primzahlen $p_i \in \pi(\mathrm{Im}(G/N \to \mathrm{Aut}(L)))$ auf, und zwar mit beliebigen Exponenten. Da G/N endlich erzeugbar ist, folgt

$$L^* = [p_1^\infty p_2^\infty \dots p_s^\infty], \qquad 0 \leqq s < \infty.$$

Erfüllt L die Voraussetzung (iii), dann suchen wir ein Element $x \in G$, das auf L durch Multiplikation mit der Rationalen a/b operiert, wobei a und b ganz und teilerfremd sind, mit $\pi(a, b) = \{p_1, p_2, \dots, p_s\}$. Offensichtlich ist $L \cong Q^{(ab)}$. Sei V die von N und x erzeugte Untergruppe von G. V ist normal in G, $G \trianglerighteq V \triangleright N$. Ist $x_1 \in G$ mit $x_1^k \in x^l N$ dann ist auch x_1 ein derartiges Element; wir können also o.B.d.A. annehmen, daß die Faktorgruppe G/V torsionsfrei, also frei-abelsch vom Rang $m-1$ ist. V/N ist unendlich-zyklisch.

L wird von den Elementen $x^k \circ 1 = (a/b)^k$, $k \in Z$ erzeugt. In der Tat: sind λ, μ ganze Zahlen mit $\lambda a^2 + \mu b^2 = 1$, dann gilt

$$\frac{1}{(ab)^k} = \left(\lambda \frac{a}{b} + \mu \frac{b}{a}\right)^k, \qquad k \in \mathbb{Z}.$$

L ist somit zyklisch als V/N-Modul, also erst recht als G/N-Modul, womit die Implikation (iii) $\Rightarrow$ (ii) bewiesen ist.

Nun beweisen wir (iii) $\Rightarrow$ (i): Sei A ein beliebiger G/N-Modul. Nach dem „Maximumprinzip für Spektralreihen" gilt

$$H^{h+1}(G,A) \cong H^{m-1}(G/V, H^{n+2}(V,A)),$$

es genügt also zu zeigen, daß $H^{n+2}(V,A)=0$ ist. Nun ist aber

$$H^{n+2}(V,A) \cong H^1(V/N, H^{n+1}(N,A)) \cong \operatorname{Ext}(L,A)_V.$$

Wir berechnen $\operatorname{Ext}(L,A)$ mit Hilfe der folgenden kurzen exakten Folge:

$$0 \to R \xrightarrow{\alpha} X \xrightarrow{\beta} L \to 0,$$

X ist die freie abelsche Gruppe über der Basis $\{x_i\}$, $i=1,2,\ldots$, R ist die freie abelsche Gruppe über der Basis $\{r_j\}$, $j=1,2,\ldots$, und die Homomorphismen α, β sind gegeben durch

$$\beta(x_i)=\frac{1}{(a\,b)^i}, \qquad \alpha(r_j)=a\,b\,x_{j+1}-x_j.$$

Nun ist $\operatorname{Ext}(L,A)_V \cong \operatorname{Hom}(R,A)/\alpha^* \operatorname{Hom}(X,A)+I(V)\operatorname{Hom}(R,A)$, und die Operation von x auf $\operatorname{Hom}(R,A)$ kann dem folgenden Diagramm entnommen werden:

$$\begin{array}{ccccccccc}
0 & \to & R & \xrightarrow{\alpha} & X & \xrightarrow{\beta} & L & \to & 0 \\
 & & \downarrow{\scriptstyle\eta} & & \downarrow{\scriptstyle\xi} & & \downarrow{\scriptstyle x} & & \\
0 & \to & R & \xrightarrow{\alpha} & X & \xrightarrow{\beta} & L & \to & 0
\end{array}$$

$$x(r)=\frac{a}{b}\,r, \quad r\in L, \qquad \xi(x_i)=a^2\,x_{i+1}, \qquad \eta(r_j)=a^2\,r_{j+1}.$$

Es gilt also $(x\circ f)(r_j)=a^2\,x\,f(r_{j+1})$, $f\in\operatorname{Hom}(R,A)$. Nun betrachten wir die Homomorphismen $f\in\operatorname{Hom}(R,A)$, $g\in\operatorname{Hom}(X,A)$:

$$f(r_i)=\lambda^2\,x^{-1}\,a_i+\lambda\,\mu\,a_i-\mu\,b\,a_{i+1},$$
$$g(x_i)=\mu^2\,x\,a_i+\lambda\,\mu\,a_i+\mu\,a\,x\,a_{i+1}.$$

Dabei sind λ und μ zwei ganze Zahlen mit $\lambda\,a+\mu\,b=1$, und die a_i, $i=1,2,\ldots$, sind beliebige Elemente aus A. Eine kleine Rechnung zeigt nun, daß gilt

$$((x-1)\,f+\alpha^*\,g)(r_i)=a_{i+1}-(\lambda+\mu\,x)^2\,x^{-1}\,a_i.$$

Wählt man zu vorgegebenem $h\in\operatorname{Hom}(R,A)$

$$a_1=0, \qquad a_{i+1}=h(r_i)+(\lambda+\mu\,x)^2\,x^{-1}\,a_i \quad i\geqq 1,$$

dann ist $(x-1)f+\alpha^* g=h$. Somit gilt

$$\mathrm{Hom}(R,A)=\alpha^*\,\mathrm{Hom}(X,A)+I(V)\,\mathrm{Hom}(R,A),$$

was zu beweisen war.

Schließlich nehmen wir an, (iii) sei nicht erfüllt; dann sind zwei Fälle zu unterscheiden:

a) L ist vom Typ $[\prod p^{\alpha_p}]$, wobei für eine unendliche Menge P von Primzahlen $0<\alpha_p<\infty$ ist. Wir wollen zeigen, daß in diesem Falle $\mathrm{Ext}(L,L)\neq 0$ ist: Die kurze exakte Folge $0\to L\to Q\xrightarrow{\sigma} Q/L\to 0$ induziert eine exakte Sequenz

$$\mathrm{Hom}(L,Q)\to\mathrm{Hom}(L,Q/L)\to\mathrm{Ext}(L,L)\to 0.$$

Sei $\varphi\in\mathrm{Hom}(L,Q/L)$ gegeben durch $\varphi(x)=\sum_{p\in P} x p^{-1}+L$, $x\in L$. Die unendliche Summe ist sinnvoll, da $x p^{-1}$ nur für endlichviele $p\in P$ nicht in L liegt. Für alle $p\in P$ ist $\varphi(p^{-\alpha_p})=p^{-1-\alpha_p}+L\neq L$. Wenn es nun eine Abbildung $\psi\in\mathrm{Hom}(L,Q)$ mit $\sigma\psi=\varphi$ geben würde, dann wäre für alle $p\in P$ $p^{-\alpha_p}\psi(1)-p^{-1-\alpha_p}\in L$, also wäre der Nenner von $\psi(1)-p^{-1}$ teilerfremd zu p. Dann müßte aber der Nenner von $\psi(1)$ durch p teilbar sein, und das ist nicht für alle $p\in P$ möglich. φ faktorisiert daher nicht über Q, also ist $\mathrm{Ext}(L,L)\neq 0$.

b) L ist vom Typ $[p_1^\infty p_2^\infty\dots p_s^\infty]$, und eine der Primzahlen $p_i=q$ liegt nicht in $\pi(\mathrm{Im}(G\to\mathrm{Aut}(L)))$. In diesem Falle betrachten wir die Gruppe Q_q der rationalen Zahlen, deren Nenner zu q teilerfremd ist. Offensichtlich ist $\pi(\mathrm{Im}(G\to\mathrm{Aut}(L)))\subset(\mathrm{Aut}(Q_q))$, und wir wollen zeigen, daß $\mathrm{Ext}(L,Q_q)\neq 0$ ist: Die kurze exakte Folge $0\to L\to Q\to Q/L\to 0$ induziert eine exakte Folge

$$\mathrm{Ext}(Q/L,Q_q)\to\mathrm{Ext}(Q,Q_q)\to\mathrm{Ext}(L,Q_q)\to 0,$$

ferner induziert $0\to Q_q\to Q\to Q/Q_q\to 0$ die exakte Folge

$$\mathrm{Hom}(Q/L,Q/Q_q)\to\mathrm{Ext}(Q/L,Q_q)\to 0.$$

Jedes Element $x\in Q/L$ hat zu q teilerfremde Ordnung, die Ordnungen der Elemente $y\in Q/Q_q$ sind dagegen gerade die Potenzen von q. Daher ist $\mathrm{Hom}(Q/L,Q/Q_q)=0$, also auch $\mathrm{Ext}(Q/L,Q_q)=0$. Damit folgt

$$\mathrm{Ext}(L,Q_q)\cong\mathrm{Ext}(Q,Q_q),$$

und nach Fuchs [4, IX.52 Exercise 15] ist $\mathrm{Ext}(Q,Q_q)\neq 0$.

In beiden Fällen a) und b) sind die Voraussetzungen von Lemma (5.1) erfüllt, es gibt also einen G/N-Modul A mit $H^{h+1}(G,A)\neq 0$. Damit ist Satz (5.2) vollständig bewiesen.

6. Eine notwendige Bedingung für $cdG = hG$

Satz (5.2) ist auch dann noch richtig, wenn zwar G/N nicht frei-abelsch ist, jedoch eine frei-abelsche Untergruppe U/N vom Rang $h(U/N) = m < \infty$ mit endlichem Index enthält. In der Tat: Sei o.B.d.A. U normal in G. Dann operiert G/U trivial auf $L = H_n(N, \mathbb{Z})$, und somit sind die Aussagen „L ist zyklisch als G/N-Modul" und „L ist zyklisch als U/N-Modul" äquivalent. Ist $H^{n+m+1}(G, A) = 0$ für jeden G/N-Modul A, dann gilt für jeden U/N-Modul B $H^{n+m+1}(U, B) \cong H^{n+m+1}(G, \mathrm{Hom}_U(\mathbb{Z}G, B)) = 0$. Sei umgekehrt $H^{n+m+1}(U, B) = 0$ für jeden U/N-Modul B, und sei A ein G/N-Modul. Da der Index $|G:U|$ endlich ist, sind die G-Moduln $\mathrm{Hom}_U(\mathbb{Z}G, A)$ und $\mathbb{Z}G \otimes_U A$ isomorph. Es gibt also einen Epimorphismus

$$\rho\colon \mathrm{Hom}_U(\mathbb{Z}G, A) \twoheadrightarrow A.$$

Nach (3.1) ist $cdG \leqq n + m + 1$, also induziert ρ einen Epimorphismus

$$H^{n+m+1}(U, A) \xrightarrow{\cong} H^{n+m+1}(G, \mathrm{Hom}_U(\mathbb{Z}G, A)) \twoheadrightarrow H^{n+m+1}(G, A),$$

also muß $H^{n+m+1}(G, A) = 0$ sein.

Nun sei G eine torsionsfreie, auflösbare Gruppe mit endlicher Hirschzahl $hG = h$, und sei $R \trianglelefteq G$ das Hirsch-Plotkin-Radikal von G, $hR = n$. Nach Baer-Heineken [1, Prop. 5.5] ist R nilpotent, und G/R enthält eine endlich erzeugte, frei-abelsche Untergruppe von endlichem Index, es folgt also: *Der Funktor $H^{h+1}(G, -)$ verschwindet genau dann auf allen G/R-Moduln, wenn $H_n(R, \mathbb{Z})$ als G/R-Modul zyklisch ist.*

(6.1) **Korollar.** *Sei G eine torsionsfreie, auflösbare Gruppe mit $cdG = hG < \infty$. Sei $R \trianglelefteq G$ das Hirsch-Plotkin-Radikal von G, $hR = n$. Dann ist $H_n(R, \mathbb{Z})$ zyklisch als G/R-Modul.*

Die Homologiegruppe $H_n(R, \mathbb{Z})$ läßt sich wie folgt berechnen:

(6.2) **Lemma.** *Jede torsionsfreie, nilpotente Gruppe R von endlicher Hirschzahl $hR = n$ besitzt eine Zentralreihe $R = R_0 \trianglerighteq R_1 \trianglerighteq \cdots \trianglerighteq R_n = 1$ mit rationalen Faktoren, und es gilt:*

$$H_n(R, \mathbb{Z}) \cong R/R_1 \otimes R_1/R_2 \otimes \cdots \otimes R_{n-1}.$$

Beweis. Die aufsteigende Zentralreihe von R hat torsionsfreie Faktoren, also läßt sie sich zu einer Zentralreihe mit rationalen Faktoren verfeinern. Der Rest der Behauptung folgt mit Induktion nach n. Ist $n = 1$, dann ist R rational und die Aussage trivial. Für $n \geqq 2$ folgt nach dem „Maximumprinzip" aus der Lyndonspektralreihe

$$H_n(R, \mathbb{Z}) \cong H_{n-1}(R/R_{n-1}, H_1(R_{n-1}, \mathbb{Z})).$$

R_{n-1} ist im Zentrum von R, also operiert R/R_{n-1} trivial auf $H_1(R_{n-1}, \mathbb{Z}) \cong R_{n-1}$. Nach dem Theorem der universellen Koeffizienten gilt daher $H_n(R, \mathbb{Z}) \cong H_{n-1}(R/R_{n-1}, \mathbb{Z}) \otimes R_{n-1}$, und aus der Induktionsvoraussetzung folgt daraus die Behauptung.

Die Gruppe G heißt eine Minimax-Gruppe, wenn sie eine (endliche) Normalreihe besitzt, deren Faktoren entweder der Minimal- oder der Maximalbedingung genügen. Offensichtlich sind Untergruppen, Faktorgruppen und Extensionen mit Minimax-Gruppen von Minimax-Gruppen wieder Minimax-Gruppen. Die rationalen Minimax-Gruppen sind genau die Gruppen $Q^{(n)}$, $n \geqq 1$, $n \in \mathbb{Z}$.

(6.3) **Korollar.** *Jede torsionsfreie, auflösbare Gruppe G mit* $cdG = hG < \infty$ *ist eine Minimax-Gruppe.*

Beweis. Sei R das Hirsch-Plotkin-Radikal von G. Nach (6.2) hat R eine Zentralreihe, deren Faktoren Untergruppen von $H_n(R, \mathbb{Z})$, $n = hR$, sind. Nach (6.1) ist aber $H_n(R, \mathbb{Z})$ zyklisch als G/R-Modul, also als abelsche Gruppe minimax (vgl. (5.2)(iii)). Da ferner G/R polyzyklisch ist, folgt daraus die Behauptung.

Bemerkungen. Sei G eine torsionsfreie, auflösbare Gruppe mit endlicher Hirschzahl. Wenn G polyzyklisch ist, dann gilt $cdG = hG$; ist dagegen G keine Minimax-Gruppe, dann gilt $cdG = hG + 1$. Die polyzyklischen Gruppen sind endlich präsentierbar. dagegen können Nicht-Minimax-Gruppen nach Robinson [8] nicht einmal endlich erzeugbar sein. Auch im Hinblick auf die Beispiele (4.1) und (4.2) drängt sich damit die attraktive Frage auf: *Gilt* $cdG = hG$ *genau dann, wenn G endlich präsentierbar ist?* Es ist mir nicht gelungen, darauf eine Antwort zu finden.

Literatur

1. Baer, R., Heineken, H.: Radical groups of finite abelian subgroup rank. Illinois J. Math., erscheint demnächst.
2. Bieri, R.: Gruppen mit Poincaré-Dualität. Promotionsarbeit an der Eidgenössischen Technischen Hochschule Zürich (1972). Comment. Math. Helv., erscheint demnächst.
3. Čarin, V.S.: On locally soluble groups of finite rank. Mat. Sbornik **41**, 37–48 (1957) [Russisch].
4. Fuchs, L.: Infinite abelian groups, vol. 1. New York: Academic Press 1970.
5. Gruenberg, K.W.: Cohomological topics in group theory. Lecture Notes in Math. 143, Berlin-Heidelberg-New York: Springer 1970.
6. Merzljakov, Ju.I.: Locally soluble groups of finite rank. Algebra i Logika **3**, 5–16 (1964) [Russisch].
7. Robinson, D.J.S.: On soluble minimax groups. Math. Z. **89**, 30–51 (1965).
8. Robinson, D.J.S.: A note on groups of finite rank. Compositio Math. **31**, 240–246 (1969).
9. Zaičev, D.I.: On solvable groups of finite rank. Doklady Akad. Nauk. SSSR **181**, 13–14 (1968) [Russisch] = Soviet Math. Doklady **9**, 783–785 (1968).

Dr. R. Bieri
Forschungsinstitut für Mathematik
Eidgenössische Technische Hochschule
8006 Zürich
Schweiz

(Eingegangen am 1. April 1972)

Math. Z. 128, 245–251 (1972)

Two Generator Fuchsian Groups of Genus One

Norman Purzitsky and Gerhard Rosenberger

Introduction

Let $A, B \in Lf(2, \mathbb{R})$, the group of linear fractional transformations with real entries and determinant 1. The results of Knapp [2], Purzitsky [8], and Rosenberger [10] give conditions so that we may determine in all cases but one whether the group $\{A, B\}$ generated by A and B is a Fuchsian group. The remaining case is when A and B are hyperbolic transformations whose commutator $[A, B]$ is of finite order. We give necessary and sufficient conditions for this group to be discrete. We also obtain all faithful representations of a group whose presentation is $\{A, B \mid [A, B]^n = 1\}$ by a discrete subgroup of $PSL(2, \mathbb{R}) = SL(2, \mathbb{R})/\{\pm 1\} \cong Lf(2, \mathbb{R})$ and partition the representations into disjoint conjugacy classes. Here 1 is the identity. Results in this direction were first given by Lehner and Newman [4, 5] who considered the case of two elliptic generators. Their results were extended to all two generator free products by Purzitsky [9] and Rosenberger [10].

The form of these distinct conjugacy classes provides an explicit solution to a problem which appears in [1]. Let

$$S = \{(x, y, z) \mid x^2 + y^2 + z^2 - xyz = 2 - 2\cos \pi/n\} \subseteq \mathbb{R}^3.$$

Let x, y, z be the traces of A, B, AB respectively i.e. $x = \sigma(A)$, $y = \sigma(B)$ and $z = \sigma(AB)$. Then $-2\cos \pi/n = \sigma([A, B])$. Thus the birational transformations on (x, y, z) induced by the Nielsen transformations on A, B preserve S. Fricke and Klein assert that this group acts discontinuously on S. We prove this by providing a fundamental domain.

We also mention that a corollary to the results of this paper and those of [9, 10] is an extension of Nielsen's Theorem on free groups to all two generator Fuchsian groups, $G = \{A, B\}$. That is, if $X, Y \in G$ and $G = \{X, Y\}$, then X and Y are Nielsen transformations of A, B.

Section 1. Definitions

1) Let $X(z) = \dfrac{az+b}{cz+d}$, $c \neq 0$. The isometric circle of X is the circle $|cz+d| = 1$.

2) Let $G = \{A, B \mid R_\alpha(A, B) = 1\}$. Let $W_1(A, B)$, $W_2(A, B)$ be words in A, B. We shall write

a) $W_1(A, B) \equiv W_2(A, B)$ if W_1 and W_2 are the same word.

b) $W_1(A, B) = W_2(A, B)$ if W_1 and W_2 are equal in the free group of rank 2, $F_2 = \{A, B\}$.

c) $W_1(A, B) \sim W_2(A, B)$ if W_1 and W_2 are equal in G.

3) Let $W(A, B) \equiv A^{\alpha_1} B^{\beta_1} \ldots A^{\alpha_n} B^{\beta_n}$, where $\alpha_i \neq 0$ if $i = 2, \ldots, n$ and $\beta_j \neq 0$ if $j = 1, \ldots, n-1$. We define the length of W as $L(W) = \sum_{i=1}^{n} |\alpha_i| + |\beta_i|$. If W is not of the proper form we cancel the terms $X X^{-1}$ to obtain the proper form and then define $L(W)$.

4) Let $\{A, B\} = G$. An elementary Nielsen transformation of A, B is the replacement of A, B by the new generators $A^n B, A$ ($n \in Z$, the set of integers). A Nielsen transformation is the replacement of A, B by two other generators obtained from finitely many applications of elementary Nielsen transformations.

5) Let $G = \{A, B\} = \{C, D\}$ be a two generator group. The generators C, D are Nielsen equivalent to A, B if there is a pair of generators W_1, W_2 which are Nielsen transformations of A, B such that $C \sim T W_1 T^{-1}$, $D \sim T W_2 T^{-1}$, where $T \in G$.

6) Let $\sigma(A) > 2$, $\sigma(B) > 2$. The fixed points of A, B separate each other if and only if

$$\Lambda(A, B) = \frac{w_A - w_B}{w_A - z_B} \cdot \frac{z_A - z_B}{z_A - w_B} < 0,$$

where w_x, z_x are the fixed points of x.

7) $H^+ = \{x + iy \mid y > 0, -\infty < x < \infty\}$.

8) A hyperbolic line in H^+ is a semicircle perpendicular to the real axis or a line of the form $\{x_0 + it \mid t > 0\}$.

Remark. We note that if A and B are hyperbolic transformations, then $\sigma([A, B]) < 2$ if and only if $\Lambda(A, B) < 0$.

Section 2. Main Theorems

Theorem 1. *Let A, B be hyperbolic transformations such that $[A, B]^n = 1$. Then $\{A, B\}$ is a Fuchsian group if and only if $\sigma[A, B] = -2\cos \pi/n$ ($n \geqq 2$).*

Theorem 2. *Let $A, B \in PSL(2, \mathbb{R})$ and $G = \{A, B \mid [A, B]^n = 1\}$ be a Fuchsian group. Then G is conjugate over $GL(2, \mathbb{R})$ to one and only one*

$$G(\lambda, \mu, \rho) = \left\{ \begin{pmatrix} 0 & 1 \\ -1 & \lambda \end{pmatrix}, \begin{pmatrix} 0 & \rho \\ -\rho^{-1} & \mu \end{pmatrix} \right\}$$

where $\lambda^2 + \mu^2 + \left(\rho + \frac{1}{\rho}\right)^2 - \lambda\mu\left(\rho + \frac{1}{\rho}\right) = 2 - 2\cos \pi/n$ ($n \geqq 2$), $\rho > 1$, *and* $\lambda \leqq \mu \leqq \rho + \frac{1}{\rho} \leqq \frac{1}{2}\lambda\mu$. *Moreover* $2 < \lambda < 3$.

Section 3

In this section we give sufficient conditions for two hyperbolic transformations whose fixed points separate each other and $[A,B]^n=1$ $(n\geqq 2)$ to generate a Fuchsian group.

Lemma 1. *Let $\sigma(A)>2$, $\sigma(B)>2$ and $A^{-1}B$ be elliptic. If $\sigma(A^{-1}B)\leqq 0$, then $0\leqq|\theta|\leqq\pi/2$, where 2θ is the angle of rotation for $A^{-1}B$.*

Proof. We conjugate the group $\{A,B\}$ so that the fixed points of A are ± 1 and those of B are $\pm k$. We note that this is possible since $\sigma(A^{-1}B)\leqq 0$ implies the fixed points of A and B do not separate each other. So we may assume $A=\begin{pmatrix} a & c\\ c & a\end{pmatrix}$ and $B=\begin{pmatrix} a_1 & b_1\\ c_1 & a_1\end{pmatrix}$, where $c, c_1, b_1>0$, $a^2-c^2=1$ and $a_1^2-b_1c_1=1$.

Since $A^{-1}B$ is elliptic, the isometric circle of A intersects that of B at an angle θ, where $2\theta<2\pi$ is the angle of rotation. Since $\sigma(A^{-1}B)\leqq 0$, we see that $c(b_1+c_1)\geqq 2aa_1$. Thus $1/c^2+1/c_1^2\leqq(a/c-a_1/c_1)^2$. This last inequality implies $0\leqq|\theta|\leqq\pi/2$.

Lemma 2. *Let $G=\{A,B\}\subseteq Lf(2,\mathbb{R})$, where $\sigma(A)>2$, $\sigma(B)>2$, and $\sigma([A,B])=-2\cos\pi/n$. Then G is a Fuchsian group.*

Proof. We conjugate the group G so that $A=\begin{pmatrix} p & 0\\ 0 & p^{-1}\end{pmatrix}$, $B=\begin{pmatrix} a & ad-1\\ 1 & d\end{pmatrix}$, where $p>1$, $ad>1$. We now proceed to construct a fundamental domain for G.

Let
$$C_1=\{z\mid |z-a|=\sqrt{a/d}\}\quad\text{and}\quad C_2=\{z\mid |z+d|=\sqrt{d/a}\}.$$

Since $\sigma([A,B])>-2$ we see that $H^+\cap A(C_1)\cap C_1\neq\varnothing$ and
$$H^+\cap A(C_2)\cap C_2\neq\varnothing\quad[8].$$

Thus $H^+\cap A^{-1}(C_i)\cap C_i\neq\varnothing$ $(i=1,2)$. Let $w_{i,k}\in A^i(C_k)\cap C_k\cap H^+$ for $i=\pm 1$, $k=1,2$. Let C_3 be the hyperbolic line joining $w_{-1,1}$ to $w_{-1,2}$ and C_4 the hyperbolic line joining $w_{1,1}$ to $w_{1,2}$. The domain D, which is exterior to C_1, C_2, C_3 but interior to C_4 will form a fundamental domain for $\{A,B\}$.

From [8] we see that $w_{-1,1}$ is a fixed point of $[A,B^{-1}]$. If we set θ_{ij} to be the angle of intersection between the circles C_i and C_j, where $(i,j)=(1,3)$, $(1,4)$, $(2,3)$, $(2,4)$. Then $\theta_{13}+\theta_{14}+\theta_{23}+\theta_{24}=\gamma$ is the angle of rotation of $[A,B^{-1}]$. By Lemma 1 we see that $|\gamma|=2\pi/n$. Thus by the Poincare Theorem D is a fundamental domain for G, and hence G is discrete.

Section 4. The Necessary Conditions and Classification

Lemma 3. *Let* $G=\{A, B \mid [A, B]^n=1\}$. *Let* X, Y *be generators for* G *such that* $[X, Y]^r=1$. *Then* X, Y *are Nielsen equivalent to* A, B.

Proof. We consider X, Y as words in A, B. Let $M=\{\alpha, \beta) \mid \alpha, \beta$ are Nielsen equivalent in G to $X, Y\}$ and $N=\{L([\alpha, \beta]) \mid (\alpha, \beta) \in M\}$. We may assume $L([X, Y])$ is a minimum of N. We shall proceed to show $X=A$, $Y=B$.

We remark that $L([X, Y])$ is minimal with respect to all choices of α, β. We show that in the class of words equivalent to $[X, Y]$ in G, we have $[X, Y]$ is also of minimal length.

Let $W_1 \sim [X, Y]$ be of minimal length. We assume $L(W_1) < L([X, Y])$. Since $W_1 \sim [X, Y]$, we have $[X, Y]=W_1 \prod_{i=1}^{m} T_i [A, B]^n T_i^{-1}$. Now there may be some cancelation between T_i and $[A, B]$. We take this into account by noting that if there is cancelation, $T_i[A, B]^n T_i^{-1}=S_i[C_i, D_i]^n S_i^{-1}$, where $(C_i, D_i)=(A^{\pm 1}, B^{\pm 1})$ or $(B^{\pm 1}, A^{\pm 1})$. We now assume that there is no cancelation between S_i and $[C_i, D_i]$.

Also there may be some cancelation between W_1 and

$$\prod_{i=1}^{m} S_i [C_i, D_i]^n S_i^{-1},$$

but $W_1 \neq W[C_1, D_1]^n S_1^{-1}$. Thus if $m \geqq 2$ or if there is no cancelation between W_1 and $S_1[C_1, D_1] S_1^{-1}$ then $[X, Y]=P[C, D]^n Q$, where $(C, D)=(C_i, D_i)$ for some i and $P[C, D]\,Q$ does not have any terms of the form ZZ^{-1}, $Z=A, B$. Write $X=UVS$, $Y=URS$, where X and Y are internally void of free cancelation and

$$L(VR^{-1})=L(R^{-1}V)=L(V^{-1}R)=L(RV^{-1})=L(V)+L(R).$$

Then $[X, Y]=UVSURV^{-1}U^{-1}S^{-1}R^{-1}U^{-1}$. Now if we consider the generators $\alpha=VSU, \beta=RSU$ we see that $[\alpha, \beta]=VSUR\,V^{-1}S^{-1}U^{-1}R^{-1}$. Since $[X, Y]$ is of a minimal length $U=1$. So $X=VS$, $Y=RS$ and $[X, Y]=VSRV^{-1}S^{-1}R^{-1} \equiv P[C, D]^n Q$. We now show this to be impossible.

If $L(V) \leqq L(P)$ by a conjugation we may consider $SRV^{-1}S^{-1}R^{-1}V \equiv P_1[C, D]^n Q V$. By the symmetry of R, V, S we may assume $L(V) > L(P)$.

If $L(VS) > L(P[C, D]^n)$, i.e. $VS \equiv P[C, D]^n Q_1$, then we may replace $X \equiv VS$ by PQ_1 and obtain a contradiction to the choice of X, Y. So we may assume $[C, D]^n=V_1 S R_1$, where $V \equiv V_0 V_1$, $R=R_1 R_0$. Then $X \equiv V_0 V_1 S=V_0[C, D]^n R_1^{-1} \sim V_0 R_1^{-1}$ and $Y \equiv R_1 R_0 S$. After a conjugation by R_1 we have $R_1^{-1} X R_1 \sim R_1^{-1} V_0$ and $R_1^{-1} Y R_1 \sim R_0 V_1^{-1}$. Clearly $L([R_1^{-1} V_0, R_0 V_1^{-1}]) < L([X, Y])$. Thus $m=0$; or $m=1$ and $W_1=WK^{-1}S_1^{-1}$ where $K=[C, D]^k C$, $[C, D]^k CD$ or $[C, D]^k CDC^{-1}$ under the restriction

$L(K)<2n$ (if $L(K)\geqq 2n$, then $L([X, Y])\geqq L(W_1)$). The next paragraph shows that $m=0$.

Suppose $m=1$ and $W_1=WK^{-1}S_1^{-1}$ where K is as above. Then $VSRV^{-1}S^{-1}R^{-1}\equiv WK_1S_1^{-1}$ where $KK_1\equiv[C,D]^n$ so $L(K_1)>2n$. Again we may assume $L(V)>L(W)$. Now if $L(VS)\geqq L(WK_1)$ then $X\equiv VS\equiv WK_1S_0$. Hence we may replace X by $WK^{-1}S_0$. This gives us a contradiction to our choice of X. For $\alpha=WK^{-1}S_0$, $\beta=Y$ yields

$$L([\alpha,\beta])<L([X,Y]).$$

Thus we may assume $K_1\equiv V_1SR_1$, where $V\equiv V_0V_1$, $R\equiv R_1R_0$. Then again we let $\alpha=R_1V_0K^{-1}$, $\beta=R_0V_1^{-1}K^{-1}$ and obtain $(\alpha,\beta)\in M$ and $L([\alpha,\beta])<L[(X,Y)]$. So $m=0$.

Therefore $VSRV^{-1}S^{-1}R^{-1}\equiv W_1$. Since $G=\{A,B\,|\,[A,B]^n=1\}$ is a one defining relator group and W_1 is of finite order in G, we see by [7] $W_1\sim T[A,B]^PT^{-1}$, where $T\in G$. Allowing for cancelation we write $W_1\sim T_1[C,D]^PT_1^{-1}$ where $(C,D)=(A^{\pm1},B^{\pm1})$ or $(B^{\pm1},A^{\pm1})$ and no cancelation occurs between T_1 and $[C,D]$. By choosing T_1 so that $L(T_1)$ is minimal in its equivalence class and $[C,D]^P$ so that $L([C,D]^P)$ is minimal in its class we may assume $T_1[C,D]^PT_1^{-1}$ is of minimal length in its class. Thus since $G=\{X,Y\}$, $W_1\equiv T_1[C,D]^PT_1^{-1}$ or $T_1^{-1}[C,D]^PT_1$, where $1\leqq|P|\leqq n-1$. We may assume $W_1\equiv T_1[C,D]^PT_1^{-1}$. Then

$$VSRV^{-1}S^{-1}R^{-1}\equiv T_1[C,D]^PT_1^{-1}.$$

Immediately we observe $T_1\equiv 1$, for otherwise V and R begin with the same letter. So $VSRV^{-1}S^{-1}R^{-1}\equiv[C,D]^P$. By renaming C,D if necessary we may assume $P>0$. Thus V begins with C, hence S^{-1} begins with D^{-1}. Also R^{-1} ends in D^{-1}, therefore S ends in C. Noting that we have just shown S ends in both C and D we conclude $S\equiv 1$ and $VRV^{-1}R^{-1}\equiv[C,D]^P$.

We observe V begins and ends with C. Similarly R begins and ends in D. Therefore $V\equiv[C,D]^lC$ and $R\equiv[D,C^{-1}]^qD$. Computing

$$VRV^{-1}R^{-1}\equiv CD[C^{-1}DCD^{-1}]^lC^{-1}D^{-1}[C^{-1}DCD^{-1}]^q,$$

we see that $[V,R]\equiv[C,D]^P$ if and only if $l=q=0$, $p=1$ and $V=C$, $D=R$. q.e.d.

We have that if $\sigma([A,B])=-2\cos\pi/n$, then $\{A,B\}$ is disccrete. We now show this condition to be necessary.

Consider the possible standard presentations [see 6] for a Fuchsian group G which is generated by two hyperbolic transformations whose commutator is of finite order. We observe that of these presentations G must be presented by $\{X,Y,C\,|\,C^n=1,\ [X,Y]\,C=1\}$. By [3] we may construct a canonical polygon which gives this presentation. Moreover, from [3] we see that the angle of rotation for C is $2\pi/n$. Thus $\sigma(C)=\pm 2\cos\pi/n$ $(n\geqq 2)$. Now applying Lemma 3 we see that Theorem 1 is proved if we can show $\sigma(C)\neq 2\cos\pi/n$ $(n>2)$. Theorem 2 will follow immediately from Theorem 1 and Lemmas 3, 4, 5.

Lemma 4. *Let* $\sigma(A)>2$, $\sigma(B)>2$, $\Lambda(A,B)<0$. *Then there exists generators* X, Y *which are Nielsen equivalent to* A, B *such that*

$$\sigma(X)+\sigma(Y)+\sigma(XY)\leqq\sigma(\alpha)+\sigma(\beta)+\sigma(\alpha\beta) \quad \textit{for all } (\alpha,\beta)$$

which are Nielsen equivalent to A, B.

Proof. This lemma follows from Lemma 4 and Lemma 8 of [9].

Lemma 5. *Let* $\sigma(A)>2$, $\sigma(B)>2$, $\Lambda(A,B)<0$, *and* $2<\sigma(A)\leqq\sigma(B)\leqq\sigma(AB)$. *Then* $\sigma(A)+\sigma(B)+\sigma(AB)\leqq\sigma(X)+\sigma(Y)+\sigma(XY)$ *for all* X, Y *which are Nielsen equivalent to* A, B *if and only if* $\sigma(AB)\leqq\frac{1}{2}\sigma(A)\sigma(B)$. *Moreover, if* $\sigma([A,B])>-2$ *then* $\sigma(A)<3$.

Proof. $\sigma(AB)+\sigma(A^{-1}B)=\sigma(A)\sigma(B)$. So $\sigma(A^{-1}B)=\sigma(A)\sigma(B)-\sigma(AB)$. If A, B has the above minimal property then $\sigma(A^{-1}B)\geqq\sigma(AB)$. Thus $\sigma(AB)\leqq\frac{1}{2}\sigma(A)\sigma(B)$. If $\sigma(AB)\leqq\frac{1}{2}\sigma(A)\sigma(B)$ then $\sigma(A^{-1}B)\geqq\sigma(AB)$. Hence A, B satisfies the inequality $\sigma(A)+\sigma(B)+\sigma(AB)\leqq\sigma(X)+\sigma(Y)+\sigma(XY)$ for all X, Y which are Nielson equivalent to A, B.

To show $\sigma(A)<3$ we note

$$\sigma(A)^2+\sigma(B)^2+\sigma(AB)^2-\sigma(A)\sigma(B)\sigma(AB)=2+\sigma([A,B])>0.$$

Since $\sigma(A)\leqq\sigma(B)\leqq(AB)$ we have $\frac{1}{3}\sigma(A)\sigma(B)\leqq\sigma(AB)\leqq\frac{1}{2}\sigma(A)\sigma(B)$. Thus $\frac{1}{2}\sigma(A)\sigma(B)-\sigma(AB)\leqq\frac{1}{6}\sigma(A)\sigma(B)$. So we see

$$0<\sigma(A)^2+\sigma(B)^2+[\tfrac{1}{2}\sigma(A)\sigma(B)-\sigma(AB)]^2-[\tfrac{1}{2}\sigma(A)\sigma(B)]^2$$

implies $0<2\sigma(B)^2[1-\frac{1}{9}\sigma(A)^2]$. q.e.d.

Corollary. *Let* $G=\{A,B\,|\,[A,B]^n=1\}$ *be a Fuchsian group. Then* $\sigma([A,B])\neq +2\cos(\pi/n)$, $n\geqq 3$.

Proof. Let $\sigma([A,B])=2\cos\pi/n=\lambda_n$ $(n\geqq 3)$. We may assume

$$\sigma(A)+\sigma(B)+\sigma(AB)$$

is minimal. Let $C=[A,B]$. Then

$$\begin{aligned}\sigma(AC)&=\sigma(A)\sigma(C)-\sigma(A^{-1}C)\\&=\sigma(A)(\lambda_n-1)\\&=\mu(\lambda_n-1)\geqq 0,\end{aligned}$$

where $\mu=\sigma(A)<3$. Since G is discrete and has the presentation

$$\{A,B\,|\,[A,B]^n=1\},$$

we see that $\lambda_n \geqq 1+\frac{2}{\mu}$. Similarly $\sigma([A^2, B]) = \mu^2(\lambda_n - 2)+2<2$. So $\mu^2(\lambda_n-2) \leqq -4$. Thus $\lambda_n \leqq 2-\frac{4}{\mu^2}$. Hence $2-4/\mu^2 \geqq \lambda_n \geqq 1+\frac{2}{\mu}$ implies $\mu \geqq \sqrt{5}+1>3$. Since $\mu<3$ we have $\sigma([A, B]) \neq 2\cos\pi/n$.

Bibliography

1. Fricke, R., Klein, F.: Vorlesungen über der Theorie der automorphen Funktionen. Vol. I, p. 285–315, 389–399, Vol. II, p. 300. Leipzig: Teubner 1897 and 1901.
2. Knapp, A. W.: Doubly generated Fuchsian groups. Michigan Math. J. **15**, 289–304 (1968).
3. Keen, L.: Canonical polygons for finitely generated Fuchsian groups. Acta Math. **115**, 1–16 (1966).
4. Lehner, J., Newman, M.: Real two-dimensional representations of the modular group and related groups. Amer. Math. J. **87**, 945–954 (1965).
5. Lehner, J., Newman, M.: Real two-dimensional representations of the free product of two finite cyclic groups. Cambridge Philos. Soc. **62**, 135–141 (1966).
6. Lehner, J.: Discontinuous groups and automorphic functions. Surveys No. 8, American Math. Soc., Providence, R. I., 1964.
7. Magnus, W., Karrass, A., Solitar, D.: Combinatorial group theory. New York: Interscience Publishers 1966, p. 269.
8. Purzitsky, N.: Two generator discrete free products. Math. Z. **126**, 209–223 (1972).
9. Purzitsky, N.: Real two-dimensional representations of two generator free groups. Math. Z. **127**, 95–104 (1972).
10. Rosenberger, G.: Fuchssche Gruppen, die freies Produkt zweier zyklischer Gruppen sind, und die Gleichung $x^2+y^2+z^2=x\,y\,z$. Math. Ann. (To appear).

Dr. Norman Purzitsky
York University
Toronto, Ontario
Canada

G. Rosenberger
D-2104 Hamburg 92
Haferacker 1
Federal Republic of Germany

(Received May 23, 1972)

Math. Z. 128, 253–269 (1972)

Maximum Principles for Minimal Surfaces and for Surfaces of Continuous Mean Curvature

Stefan Hildebrandt

It is well known that maximum principles for elliptic equations have striking and sometimes quite surprising geometric consequences. In the bibliography we have included a selection from the extensive literature which is relevant to the present paper. First, we shall use the maximum principle for subharmonic functions to prove necessary conditions for the existence of multiply connected minimal surfaces and, more generally, of surfaces with continuous mean curvature $\mathscr{H}$ ($\mathscr{H}$-surfaces) having prescribed disconnected boundaries. These conditions form a counterpart to the sufficient conditions of Douglas [6, 7], Courant [3], and Werner [31] for the solvability of the general Plateau-Douglas problem, or of Courant's free boundary value problems. Our results are new for $\mathscr{H}$-surfaces, and, even in the case of minimal surfaces, they apply to much more general situations than the previous results of Nitsche [21–26] which were obtained by entirely different and quite involved considerations. It is conceivable that our results can be modified and improved by other and, possibly, luckier choices of the test function t.

Secondly, we shall derive certain inclusion principles for $\mathscr{H}$-surfaces in Riemannian manifolds involving the mean curvature Λ of the boundary S of the including set J. Related results for $\mathscr{H}$-surfaces in Euclidean space were found independently by Gulliver, Spruck [10], and by Gulliver, Osserman, Royden [9] who have used a completely different technique. There are many applications of this "$\mathscr{H}, \Lambda$-maximum principle". It yields, in particular, a new existence theorem for Plateau's problem. In a forthcoming paper, Gulliver and Spruck have worked out details. While their approach seems to be restricted to the Euclidean space, our method will also apply to Riemannian manifolds.

I am indebted to Klaus Steffen for pointing out to me the problem which is treated in part II. Also, I wish to thank Enrico Giusti for many discussions concerning the topic of this paper, and Helmut Kaul for an improvement of Theorems 3 and 4.

1. Necessary Conditions for the Solvability of Plateau's Problem by Connected Minimal Surfaces and H-Surfaces with Disconnected Boundary Values

In this section we consider surfaces

$$\mathfrak{x}=\mathfrak{x}(u,v)=\big(x(u,v), y(u,v), z(u,v)\big)$$

in the three-dimensional Euclidean space E^3, which are defined on a bounded open set Ω of $\mathbb{R}^2$. We shall identify $\mathbb{R}^2$ with $\mathbb{C}$, and (u,v) with $w=u+iv$, and use the notation $\mathfrak{x}=\mathfrak{x}(u,v)=\mathfrak{x}(w)$.

A surface $\mathfrak{x}\in C^1(\Omega, E^3)$ is said to be represented by conformal parameters u and v if it satisfies

$$\mathfrak{x}_u^2=\mathfrak{x}_v^2, \quad \mathfrak{x}_u\cdot\mathfrak{x}_v=0 \quad \text{in } \Omega. \tag{1}$$

These relations are equivalent with

$$\mathfrak{x}_w^2=x_w^2+y_w^2+z_w^2=0 \quad \text{in } \Omega, \tag{2}$$

where $\mathfrak{x}_w$ stands for the complex derivative $\frac{1}{2}(\mathfrak{x}_u-i\,\mathfrak{x}_v)$. Relation (2) implies

$$|z_w|^2\leqq|x_w|^2+|y_w|^2,$$

hence

$$|\nabla z|^2\leqq|\nabla x|^2+|\nabla y|^2. \tag{3}$$

Here f denotes the gradient (f_u, f_v) of a function $f=f(u,v)$, and $|\nabla f|^2=|f_u|^2+|f_v|^2$. Furthermore, we shall write $\Delta f=f_{uu}+f_{vv}$.

Let $t=t(x,y,z)$ be a function of class $C^2(E^3,\mathbb{R})$ with the gradient

$$t_{\mathfrak{x}}=(t_x, t_y, t_z)$$

and the Hessian matrix

$$t_{\mathfrak{x}\mathfrak{x}}=\begin{pmatrix} t_{xx} & t_{xy} & t_{xz}\\ t_{yx} & t_{yy} & t_{yz}\\ t_{zx} & t_{zy} & t_{zz}\end{pmatrix},$$

and consider a mapping $\mathfrak{x}\in C^2(\Omega, E^3)$. Then $\tau=t\circ\mathfrak{x}=t(\mathfrak{x})$ is of class $C^2(\Omega,\mathbb{R})$ and

$$\Delta\tau=\nabla\mathfrak{x}\cdot t_{\mathfrak{x}\mathfrak{x}}(\mathfrak{x})\,\nabla\mathfrak{x}+t_{\mathfrak{x}}(\mathfrak{x})\cdot\Delta\mathfrak{x}. \tag{4}$$

If $\mathfrak{x}=\mathfrak{x}(u,v)$ is harmonic in Ω, that is, $\Delta\mathfrak{x}=0$, we obtain

$$\Delta\tau=\nabla\mathfrak{x}\cdot t_{\mathfrak{x}\mathfrak{x}}(\mathfrak{x})\,\nabla\mathfrak{x}. \tag{5}$$

Choosing suitable functions t, we can derive interesting informations from (4), (5) about a surface $\mathfrak{x}=\mathfrak{x}(u,v)$. For instance, τ is harmonic in Ω for each linear function $t: E^3\to\mathbb{R}$ and each harmonic surface $\mathfrak{x}: \Omega\to E^3$, and the maximum principle gives the following well known result:

If $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ is harmonic in Ω, then it is contained in the convex hull of its boundary values $\mathfrak{x}(\partial\Omega)$.

A mapping $\mathfrak{x}=\mathfrak{x}(u,v)$ of class $C^2(\Omega, E^3)$ is said to be a minimal surface on Ω if $\mathfrak{x} \not\equiv \text{const}$ on Ω, and if it satisfies

$$\Delta\mathfrak{x}=0, \quad \mathfrak{x}_u^2=\mathfrak{x}_v^2, \quad \mathfrak{x}_u \cdot \mathfrak{x}_v=0 \quad \text{in } \Omega.$$

The following remarkable characterization of minimal surfaces is due to Beckenbach and Radó [1]:

A mapping $\mathfrak{x} \in C^2(\Omega, E^3)$ is a minimal surface on Ω if and only if $\tau=\log|\mathfrak{x}-\mathfrak{a}|$ is subharmonic in Ω for each constant vector $\mathfrak{a} \in E^3$.

New interesting results for minimal surfaces will be obtained if we choose

$$t(x,y,z)=x^2+y^2-f^2(z),$$

where $f=f(z)$ is a function of class $C^2(\mathbb{R},\mathbb{R})$. Then, for $\mathfrak{x} \in C^2(\Omega, E^3)$ and $\tau=t\circ\mathfrak{x}$, we have

$$\begin{aligned}\tfrac{1}{2}\Delta\tau &= |\nabla x|^2+|\nabla y|^2-[|f'(z)|^2+f(z)f''(z)]\,|\nabla z|^2 \\ &\quad +x\,\Delta x+y\,\Delta y-f(z)f'(z)\,\Delta z.\end{aligned} \tag{6}$$

Thus we obtain

Theorem 1. *If $\mathfrak{x} \in C^2(\Omega, E^3)$ is a minimal surface on Ω and $t(x,y,z)=x^2+y^2-z^2$, then $\tau=t\circ\mathfrak{x}$ is subharmonic in Ω. Moreover, if $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$, then*

$$\sup_{\Omega} t(\mathfrak{x}) \leqq \sup_{\partial\Omega} t(\mathfrak{x}).$$

Proof. Combining (6) and (3), we have

$$\tfrac{1}{2}\Delta\tau=|\nabla x|^2+|\nabla y|^2-|\nabla z|^2 \geqq 0.$$

The second statement follows from the maximum principle.

Consequently, if $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ is a minimal surface on Ω whose boundary values $\mathfrak{x}(\partial\Omega)$ are contained in a solid hyperboloid $\mathfrak{H}$ which is congruent to $\{(x,y,z)\colon x^2+y^2-z^2 \leqq \varepsilon^2\}$, $\varepsilon>0$, then $\mathfrak{x}(\bar{\Omega}) \subset \mathfrak{H}$.

Obviously, this is an improvement over the convex hull theorem for harmonic surfaces as the example of the catenoid shows.

Moreover, we easily obtain interesting informations concerning non-existence of multiply connected minimal surfaces whose boundary values lie in disconnected sets $\mathfrak{K}_1$ and $\mathfrak{K}_2$. It seems to be quite evident that there cannot be a connected minimal surface with boundary values in $\mathfrak{K}_1$ as well as in $\mathfrak{K}_2$ if $\mathfrak{K}_1$ and $\mathfrak{K}_2$ are too far apart. For instance, there exists no catenoid bounded by two coaxial circles in parallel planes of distance r if r is too large. The experimental values of Plateau[1] agree

[1] Cf. [26], p. 170, where also further references are given.

very well with the theoretical results of Lindelöf (1870) which are nicely presented by Bliss [2]. The case of the catenoid with one fixed and one free boundary circle was theoretically and experimentally investigated by Sinclair [28]. Apparently Nitsche [21-26] was the first to point out the general problem[2]. In the special case of doubly connected minimal surfaces bounded by two Jordan curves he derived various necessary conditions for the existence, some of which will be quoted later.

Our results are quite general. They apply to the general Plateau problem[3], which is concerned with the existence of connected minimal surfaces with n preassigned Jordan curves $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$, as well as to free or partially free boundary value problems[4].

Let $\mathfrak{H}_{-\varepsilon^2}$ be a two-sheeted solid hyperboloid congruent to

$$\{(x, y, z)\colon x^2+y^2-z^2 \leqq -\varepsilon^2\}, \qquad \varepsilon>0.$$

Then $\mathfrak{H}_{-\varepsilon^2}$ can be decomposed in two closed, connected components $\mathfrak{H}^+_{-\varepsilon^2}$ and $\mathfrak{H}^-_{-\varepsilon^2}$ which are congruent to $\{(x, y, z)\colon x^2+y^2-z^2 \leqq -\varepsilon^2, \pm z>0\}$. In virtue of Theorem 1 there cannot exist a minimal surface $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ on a *connected* bounded open set Ω, with $\mathfrak{x}(\partial\Omega) \subset \mathfrak{H}_{-\varepsilon^2}$, having boundary values in $\mathfrak{H}^+_{-\varepsilon^2}$ as well as in $\mathfrak{H}^-_{-\varepsilon^2}$. In fact, the same is true for the limit case $\varepsilon=0$, as the following result shows.

Theorem 2. *Let $\mathfrak{H}_0=\mathfrak{H}_0^- + \{P_0\} + \mathfrak{H}_0^+$ be a solid cone with vertex P_0 which is congruent to $\{(x, y, z)\colon x^2+y^2-z^2 \leqq 0\}$, and let $\mathfrak{H}_0^\pm$ be the two disjoint parts of $\mathfrak{H}_0$ which are congruent to*

$$\{(x, y, z)\colon x^2+y^2-z^2 \leqq 0, \ \pm z>0\}.$$

Then there exists no minimal surface $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ on a connected bounded open set Ω with $\mathfrak{x}(\partial\Omega) \subset \mathfrak{H}_0$ such that both $\mathfrak{x}(\partial\Omega) \cap \mathfrak{H}_0^-$ and $\mathfrak{x}(\partial\Omega) \cap \mathfrak{H}_0^+$ are nonempty.

Proof. Let $\mathfrak{x}$ be such a minimal surface. From Theorem 1 we have $\mathfrak{x}(\bar{\Omega}) \subset \mathfrak{H}_0$. Since $\mathfrak{x}(\Omega)$ is a connected set, there is a point $w_0 \in \Omega$ such that $\mathfrak{x}(w_0)=P_0$. Introducing appropriate Cartesian coordinates x, y, z around P_0 as origin, it is well known[5] that $\mathfrak{x}(w)$ is of the form

$$\left.\begin{aligned} x+iy &= a(w-w_0)^\nu + o(|w-w_0|^\nu) \\ z &= \phantom{a(w-w_0)^\nu +{}} o(|w-w_0|^\nu) \end{aligned}\right\} \quad \text{for } w \to w_0$$

where ν is an integer $\geqq 1$, and a is a positive real number (if w_0 is a branch point of $\mathfrak{x}$, then $\nu \geqq 2$). Evidently, this contradicts $\mathfrak{x}(\bar{\Omega}) \subset \mathfrak{H}_0$.

---

[2] Cf. [21], Chapter IV, and Problem 34, p. 259.

[3] Cf. [3, 6, 7].

[4] Cf. [3].

[5] Chen, cf. [15] or [21].

The following numerical results can be derived from Theorem 2 by means of elementary geometrical considerations.

Corollary. *Let* $\mathfrak{K}_1, \mathfrak{K}_2$ *be sets in* E^3*, and suppose that there exists a minimal surface* $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ *on an open bounded connected set* Ω *such that* $\mathfrak{x}(\partial\Omega) \subset \mathfrak{K}_1 \cup \mathfrak{K}_2$*, and that both* $\mathfrak{x}(\partial\Omega) \cap \mathfrak{K}_1$ *and* $\mathfrak{x}(\partial\Omega) \cap \mathfrak{K}_2$ *are non-empty.*

i) *If* $\mathfrak{K}_l$ *is a closed ball of radius* δ_l *and center* $\mathfrak{x}_l$*,* $(l=1,2)$*, and* $R = |\mathfrak{x}_1 - \mathfrak{x}_2|$ *then*

$$R \leqq \sqrt{2}(\delta_1 + \delta_2).$$

(ii) *Let* $\mathfrak{K}_1$ *and* $\mathfrak{K}_2$ *be closed discs* $\mathfrak{D}_1, \mathfrak{D}_2$*, with centers* $\mathfrak{x}_1$*,* $\mathfrak{x}_2$ *and radius* δ_1, δ_2*, resp. Suppose that* $\mathfrak{D}_1, \mathfrak{D}_2$ *are contained in parallel planes which have the distance* r*. Set* $R = |\mathfrak{x}_1 - \mathfrak{x}_2|$*,* $d = \sqrt{R^2 - r^2}$*. Then, by* (i),

$$\sqrt{r^2 + d^2} \leqq \sqrt{2}(\delta_1 + \delta_2).$$

Kaul has pointed out that this formula can be improved to

$$\sqrt{r^2 + \tfrac{1}{2} d^2} \leqq \delta_1 + \delta_2.$$

Nitsche [22] *has proved the same conclusion for a minimal surface of the type of the circular annulus bounded by two Jordan curves* Γ_1 *and* Γ_2 *which are contained in the discs* $\mathfrak{D}_1$ *and* $\mathfrak{D}_2$*, respectively.*

(iii) *If* $\mathfrak{K}_1$ *and* $\mathfrak{K}_2$ *are compact sets of diameters* d_1 *and* d_2*, resp., which are separated by a slab of width* $r > 0$*, then (by Jung's theorem)*

$$r \leqq \frac{\sqrt{3}}{2}(d_1 + d_2).$$

Nitsche [23] has proved: If $\mathfrak{x}$ is a minimal surface of the type of the circular annulus bounded by two Jordan curves Γ_1, Γ_2 of diameters d_1, d_2, resp., which are separated by a slab of width $r > 0$, then

$$r \leqq \tfrac{3}{2} \operatorname{Max}(d_1, d_2).$$

Further results of Nitsche [24–26] deal with the special situation described in (ii).

Now we shall generalize our considerations to H-surfaces, that is, to surfaces of prescribed bounded mean curvature H. For these the situation is quite different. Namely, a circular cylinder provides an example of a surface of constant mean curvature bounded by coaxial circles in parallel planes whose distance r can be arbitrarily large, although the configuration is no longer stable for large r, as Plateau has observed. Other examples of this type are given by some of the surfaces of revolution with constant mean curvature, discovered by Delauney

[4] (cf. also Thompson [30]), or by surfaces which are obtained by cutting off two caps from a sphere.

In what follows, let us assume that $H = H(x, y, z)$ is a function of class $C^0(E^3, \mathbb{R})$. Then an H-surface $\mathfrak{x} = \mathfrak{x}(u, v)$ on Ω is defined as a map of class $C^2(\Omega, E^3)$, which is not identically constant on Ω and satisfies

$$\begin{aligned} &\Delta \mathfrak{x} = 2H(\mathfrak{x})\, \mathfrak{x}_u \wedge \mathfrak{x}_v \\ &\mathfrak{x}_u^2 = \mathfrak{x}_v^2, \qquad \mathfrak{x}_u \cdot \mathfrak{x}_v = 0 \end{aligned} \quad \text{in } \Omega. \tag{7}$$

Here $\mathfrak{a} \wedge \mathfrak{b}$ is the exterior product of $\mathfrak{a}, \mathfrak{b} \in E^3$. Clearly, if $\mathfrak{x}$ is an H-surface and T is a congruence mapping ($=$translation$+$rotation), then $T\mathfrak{x}$ is an $\tilde{H}$-surface, where $\tilde{H}(\mathfrak{y}) = H(T^{-1}\mathfrak{y})$ for $\mathfrak{y} \in E^3$.

From (3) we obtain

$$|\mathfrak{x}_u \wedge \mathfrak{x}_v| \leqq |\nabla x|^2 + |\nabla y|^2 \tag{8}$$

and

$$\begin{aligned} &|\nabla x|^2 + |\nabla y|^2 - [|f'(z)|^2 + f(z) f''(z)]\, |\nabla z|^2 \\ &\qquad \geqq (|\nabla x|^2 + |\nabla y|^2)\, \{1 - \big||f'(z)|^2 + f(z) f''(z)\big|\}. \end{aligned} \tag{9}$$

In virtue of (6)–(9), the function $\tau = t \circ \mathfrak{x}$ satisfies

$$\tfrac{1}{2} \Delta \tau \geqq c(\mathfrak{x})(|\nabla x|^2 + |\nabla y|^3) \quad \text{ on } \Omega, \tag{10}$$

where

$$\begin{aligned} &c(\mathfrak{x}) = 1 - |f'(z)|^2 - |f(z) f''(z)| - 2\,|H(\mathfrak{x})|\, |\mathfrak{z}|, \\ &\mathfrak{z} = (x, y, -f(z) f'(z)). \end{aligned} \tag{11}$$

Choosing $f(z) = \sqrt{b}\, z$, $0 \leqq b < 1$, and

$$t = x^2 + y^2 - b z^2,$$

we find

$$\tfrac{1}{2} \Delta \tau \geqq (|\nabla x|^2 + |\nabla y|^2)\, [1 - b - 2\,|H(\mathfrak{x})| \sqrt{x^2 + y^2 + b^2 z^2}].$$

Thus we have proved:

Theorem 3. *Let $\mathfrak{x} = \mathfrak{x}(u, v)$ be an H-surface on Ω,*

$$0 \leqq b < 1, \qquad t = x^2 + y^2 - b z^2, \quad \textit{and} \quad \tau = t \circ \mathfrak{x}.$$

Then τ is subharmonic if

$$b + 2\,|H(\mathfrak{x})| \sqrt{x^2 + y^2 + b^2 z^2} \leqq 1 \quad \textit{on } \Omega. \tag{12}$$

A consequence of this result is

Theorem 4. *Suppose that $\mathfrak{x} \in C^0(\bar{\Omega}, E^3)$ is an H-surface on Ω which satisfies*

$$\sup_{w \in \Omega} |\mathfrak{x}(w)|\, |H(\mathfrak{x}(w))| = q < \tfrac{1}{2}. \tag{13}$$

Then, for $b=1-2q\in(0,1]$, *the function* $\tau=x^2+y^2-bz^2$ *is subharmonic on* Ω, *and therefore*

$$\sup_{w\in\Omega}\tau(w)\leqq\sup_{w\in\partial\Omega}\tau(w).$$

Moreover, if $\mathfrak{H}_0=\mathfrak{H}_0^-+\{0\}+\mathfrak{H}_0^+$, $\mathfrak{H}_0^\pm=\{(x,y,z)\colon x^2+y^2-bz^2\leqq 0,\ \pm z>0\}$, $\mathfrak{x}(\partial\Omega)\subset\mathfrak{H}_0$, *and if both* $\mathfrak{x}(\partial\Omega)\cap\mathfrak{H}_0^-$ *and* $\mathfrak{x}(\partial\Omega)\cap\mathfrak{H}_0^+$ *are nonempty, then* Ω *cannot be connected.*

Proof. The first part of the assertions immediately follows from Theorem 3, and the second part is proved as Theorem 2. We can find in [15], Satz 3, the asymptotic expansion of $\mathfrak{x}=\mathfrak{x}(w)$ which has to be used in the reasoning.

Remarks. (i) Condition (13) lies in the range where one can expect to find multiply connected surfaces of prescribed mean curvature H, as Werner [31] has shown.

(ii) From Theorem 4 we easily derive numerical results which are quite useful for isoperimetric inequalities and estimates of the number of branch points (cf. Kaul [18], and a forthcoming note).

2. The $\mathscr{H}$, Λ-Maximum Principle

We shall use the following notations[6]: If $f\colon M'\to M$ is a differentiable mapping between two differentiable manifolds M' and M, then its differential $f_*\colon TM'\to TM$ is the induced mapping of the tangent bundles. Moreover, T_pM' denotes the tangent space of $p\in M'$, and $f_{*p}\colon T_pM'\to T_{f(p)}M$ is the restriction of f_* to T_pM'.

In what follows, we shall always assume that M is a three-dimensional, connected, orientable, and complete Riemannian manifold of class C^2 with scalar product $\langle V,W\rangle$, and norm $\|V\|=\langle V,V\rangle^{\frac{1}{2}}$ for $V,W\in T_pM$, $p\in M$. The Levi-Civita connection, that is, the canonical covariant derivative of M, is denoted by D.

Consider a real-valued function σ on M which is of class $C^2(M,\mathbb{R})$. Then the gradient vector field grad σ is defined in $p\in M$ by

$$\langle \operatorname{grad}_p\sigma, V\rangle = V\sigma \qquad \text{for all } V\in T_pM. \tag{14}$$

The Hessian tensor Hess σ, the Hessian bilinear form hess σ, and the Laplacian Lap σ of σ are defined in $p\in M$ by

$$\operatorname{Hess}_p\sigma\, V = D_V \operatorname{grad}\sigma, \qquad V\in T_pM, \tag{15}$$

$$\operatorname{hess}_p\sigma(V,W)=\langle \operatorname{Hess}_p\sigma\, V, W\rangle_p, \qquad V,W\in T_pM, \tag{16}$$

$$\operatorname{Lap}_p\sigma=\operatorname{Lap}\sigma(p)=\operatorname{trace}(\operatorname{Hess}_p\sigma). \tag{17}$$

[6] For details we refer the reader to Dombrowski [5].

Consider the surface

$$S_c = \{p \in M : \sigma(p) = c\} \tag{18}$$

for some $c \in \mathbb{R}$, with its "interior"

$$J_c = \{p \in M : \sigma(p) < c\}, \tag{19}$$

and let us suppose that

$$\operatorname{grad}_p \sigma \neq 0 \quad \text{for all } p \in S_c.$$

Then the mean curvature $\Lambda_c(p)$ of S_c in $p \in S_c$ with respect to the "interior normal" $-\|\operatorname{grad}_p \sigma\|^{-1} \operatorname{grad}_p \sigma$ is defined by

$$\Lambda_c(p) = \frac{1}{2\,\|\operatorname{grad}_p \sigma\|} \left\{ \operatorname{Lap} \sigma(p) - \frac{1}{\|\operatorname{grad}_p \sigma\|^2} \operatorname{hess}_p \sigma(\operatorname{grad}_p \sigma, \operatorname{grad}_p \sigma) \right\}. \tag{20}$$

Consider a chart $\varphi: U \to \mathbb{R}^3$ of an open set $U \subset M$. Let $x = (x^1, x^2, x^3)$ be the corresponding local coordinates, and ∂_k $(k = 1, 2, 3)$ their basis fields. Set

$$\begin{aligned} &g_{jk}(x) = \langle \partial_j(p), \partial_k(p) \rangle, \qquad D_{\partial_j} \partial_k = \Gamma^l_{jk} \partial_l, \\ &\Gamma^l_{jk} = \Gamma^l_{jk}(x), \qquad G = (g_{jk}), \qquad G^{-1} = (g^{jk}), \qquad g = \det G. \end{aligned} \tag{21}$$

Then, for $V = V^k \partial_k$, $W = W^l \partial_l$, we have

$$\langle V, W \rangle_p = g_{jk}(x)\, V^j(x)\, W^k(x), \quad \|V\|_p = \{g_{jk}(x)\, V^j(x)\, V^k(x)\}^{\frac{1}{2}}. \tag{22}$$

Set $\psi = \varphi^{-1}$, and $t = \sigma \circ \psi = t(x)$. Then

$$\begin{aligned} &\operatorname{grad}_p \sigma = g^{jk}(x)\, t_{x^j}(x)\, \partial_k, \\ &\operatorname{hess}_p \sigma(\partial_j, \partial_k) = t_{x^j x^k}(x) - \Gamma^l_{jk}(x)\, t_{x^l}(x). \end{aligned} \tag{23}$$

Define

$$\begin{aligned} &\nabla_x t = G^{-1} t_x = \operatorname{grad}_{\psi(x)} \sigma, \qquad t_x = (t_{x^1}, t_{x^2}, t_{x^3}), \\ &\xi = \xi(x) = -\|\nabla_x t\|^{-1} \nabla_x t, \\ &T = (t_{jk}), \qquad t_{jk} = t_{x^j x^k} - \Gamma^l_{jk} t_{x^l}, \\ &\Delta_x t = \operatorname{trace} T = t_{x^j x^j} - \Gamma^l_{jj} t_{x^l} = \operatorname{Lap}_{\psi(x)} \sigma. \end{aligned} \tag{24}$$

Then, for $p = \psi(x)$, $x = \varphi(p)$,

$$\Lambda_c(p) = \frac{1}{2\,\|\nabla_x t(x)\|} \{\Delta_x t(x) - \xi(x) \cdot T(x)\, \xi(x)\} \tag{25}$$

where $a \cdot b = a^k b^k$ for $a = a^k \partial_k$, $b = b^k \partial_k$. Moreover, the vector product $V \times W$ of two vector fields $V = V^k \partial_k$, $W = W^k \partial_k$ is defined by

$$V \times W = \sqrt{g}\, g^{jk} (V \wedge W)_j\, \partial_k. \tag{26}$$

Let Ω be an open set in $\mathbb{C}$ with the standard Euclidean metric,

$$w=u+iv\in\Omega, \qquad u=u^1, \qquad v=u^2,$$

and let U_1, U_2 be the basis fields with respect to u^1, u^2. Let us agree that repeated Latin indices $j, k, l, \dots$ are to be summed from 1 to 3, and Greek indices $\alpha, \beta, \dots$ from 1 to 2.

Consider a mapping $f\in C^2(\Omega, M)$, and define

$$D_\alpha V = D_{U_\alpha} V \tag{27}$$

for a C^1-vector field V along f. Let $\mathscr{H}=\mathscr{H}(w)$ be a function of class $C^0(\Omega, \mathbb{R})$. Then f describes a surface with prescribed mean curvature $\mathscr{H}$ in M on Ω (in short: an $\mathscr{H}$-surface) if $f\not\equiv$const on Ω, and if it satisfies the equations

$$D_\alpha f_*(U_\alpha)=2\mathscr{H} f_*(U_1)\times f_*(U_2), \tag{28a}$$

$$\|f_*(U_1)\|=\|f_*(U_2)\|, \quad \langle f_*(U_1), f_*(U_2)\rangle=0. \tag{28b}$$

Pick a point $w_0\in\Omega$, a chart $\varphi\colon U\to\mathbb{R}^3$ of a neighborhood $U\subset M$ of $f(w_0)$, a sufficiently small neighborhood $\Omega_1\subset\Omega_0$ of w_0 such that $f(\Omega_1)\subset U$, and set $X=\varphi\circ f$. Then (28) implies

$$\begin{aligned} &\Delta X^l+\Gamma^l_{jk}(X)\,X^j_{u^\alpha}X^k_{u^\alpha}=2\mathscr{H}(X_u\times X_v)^l, \quad l=1,2,3,\\ &g_{jk}(X)\,X^j_u X^k_u=g_{jk}(X)\,X^j_v X^k_v, \quad g_{jk}(X)\,X^j_u X^k_v=0 \end{aligned} \tag{29}$$

on Ω_1, $\Delta=\dfrac{\partial^2}{\partial u^2}+\dfrac{\partial^2}{\partial v^2}$.

Theorem 5. *Let Ω be a bounded, open, and connected set in $\mathbb{C}$, M a three-dimensional, connected, orientable, and complete Riemannian manifold of class C^2, and σ a real-valued function on M of class $C^2(M,\mathbb{R})$. For some $c\in\mathbb{R}$, define the surface S_c and its interior by* (18), (19), *and suppose that* $\operatorname{grad}_p\sigma\neq 0$ *for all $p\in S_c$. Furthermore, let f be an $\mathscr{H}$-surface on Ω of class $C^0(\bar\Omega, M)\cap C^2(\Omega, M)$, for some $\mathscr{H}\in C^0(\Omega,\mathbb{R})$, which has the following properties:*

(i) $f(\Omega)\subset J_c\cup S_c$,

(ii) $f(\partial\Omega)\subset J_c$,

(iii) *if $f(w)\in S_c$ for some $w\in\Omega$, then the mean curvature Λ_c of S_c satisfies $\Lambda_c[f(w)]>|\mathscr{H}(w)|$.*

Then, in fact, $f(\bar\Omega)\subset J_c$.

The next result is a quantitative sharpening of Theorem 5.

Theorem 6. *Let Ω, M, σ satisfy the assumptions of Theorem* 5, *and suppose that* $\operatorname{grad}_p\sigma\neq 0$ *for all $p\in\bigcup_{c_1\leqq\gamma\leqq c} S_\gamma$ and some $c_1\leqq c$. Furthermore,* *we assume that H is a function of class* $C^0(M, \mathbb{R})$ *which satisfies*

$$|H(p)| < \Lambda_{\sigma(p)}(p) \quad \text{for all } p \text{ in } \bigcup_{c_1 \leqq \gamma \leqq c} S_\gamma .$$

Finally, let f *be an* $\mathscr{H}$*-surface on* Ω *of class* $C^0(\bar{\Omega}, M) \cap C^2(\Omega, M)$ *with* $\mathscr{H} = H \circ f$ *which satisfies* $f(\bar{\Omega}) \subset J_c + S_c$, *and* $f(\partial\Omega) \subset J_{c_1}$. *Then we can conclude that* $f(\bar{\Omega}) \subset J_{c_1}$.

Theorem 7. *Let J be a star-sharped open set in the three-dimensional Euclidean space* E^3, *with boundary* $S = \partial J$ *being a regular surface of class* C^2 *whose mean curvature* Λ_S *with respect to the inner normal is strictly positive on S. Suppose that* $X = X(u, v)$ *is a minimal surface in* E^3 *on an open, connected, and bounded set* $\Omega \subset \mathbb{C}$ *which is of class* $C^0(\bar{\Omega}, E^3)$ *and satisfies* $X(\partial\Omega) \subset J + S$. *Then* $X(\Omega) \subset J$.

Remarks. **1.** In the special case $M = E^3$, a related result to Theorem 5 was proved by Gulliver and Spruck [10], and a result similar to Theorem 7 was independently obtained by Gulliver, Osserman, and Royden [9]. The method of these authors lies on E. Hopf's maximum principle for non-parametric solutions of elliptic differential inequalities. Hence, their proofs are entirely different from ours.

2. In Theorem 7 it is essential to assume that J is star-shaped. Indeed, if we choose J to be a solid torus, the assertion of Theorem 7 is certainly not true. Moreover, as E. Giusti has pointed out, the statement can even be false for domains J in which each closed curve is contractible to one point. An example is given by the unstable catenoid bounded by two coaxial circles which is closed at both sides by appropriate caps of positive mean curvature, and then is approximated from the interior by domains J whose boundary S has strictly positive mean curvature.

3. Obviously, a suitable generalization of Theorem 7 holds for all domains J which can be approximated from the interior by star-shaped domains J_n whose boundary surfaces have strictly positive mean curvature.

4. It is not difficult to prove that each regular compact two-dimensional surface S in E^3 without selfintersections and of class C^3 can be represented in the form

$$S = \{x \in E^3 : \sigma(x) = 0\},$$

where σ is a function of class $C^2(E^3, \mathbb{R})$ such that $\operatorname{grad}_x \sigma \neq 0$ on S. First, using normal coordinates on S, one defines σ in an ε-neighborhood N_ε of S, and then, by appropriate extension, on the whole E^3. Of course, $\operatorname{grad}_x \sigma(x)$ can be zero for points $x \notin N_\varepsilon$.

5. Our result of Theorem 5 can be generalized to solutions X of certain variational inequalities

$$\delta E(X, \Phi) \leqq 0,$$

as introduced in [17]. Hence we can improve the results of [17] using some recent ideas of Gulliver and Spruck. We shall prove

Theorem 8. *Let $\Omega, M, \sigma, J_c, S_c$ be defined as in Theorem 5, and suppose that* $\operatorname{grad}_p \sigma \neq 0$ *for all* $p \in S_c$. *Let f be a surface of class*

$$C^0(\bar{\Omega}, M) \cap C^1(\Omega, M) \cap H_2^{2,\mathrm{loc}}(\Omega, M),$$

$f \not\equiv \mathrm{const}$ *on Ω, which has the properties* (i), (ii), (iii) *of Theorem* 5. *Moreover, set*

$$\Omega^* = f^{-1}(S_c) = \{w \in \Omega \colon \sigma(f(w)) = c\}, \qquad \Omega' = \Omega - \Omega^*,$$

and suppose that

$$D_\alpha f_*(U_\alpha) = 2\mathscr{H} f_*(U_1) \times f_*(U_2) \quad a.e.\ on\ \Omega', \tag{28c}$$

$$\|f_*(U_1)\| = \|f_*(U_2)\|, \quad \langle f_*(U_1), f_*(U_2)\rangle = 0 \quad on\ \Omega, \tag{28d}$$

where $\mathscr{H} = \mathscr{H}(w)$ denotes some function $\mathscr{H} \in C^0(\Omega, \mathbb{R})$. Then $f(\bar{\Omega}) \subset J_c$, that is, Ω^* is empty, and f is an $\mathscr{H}$-surface on Ω.

Remark. The application of Theorem 8 to solutions of $\delta E \leqq 0$ is possible because of certain regularity theorems proved by Hildebrandt [16] (cf. also Tomi [29]).

Proof of Theorem 5. On account of the assumptions (i) and (ii), the function $\tau = \sigma \circ f \in C^0(\bar{\Omega}, \mathbb{R}) \cap C^2(\Omega, \mathbb{R})$ has the properties

$$\tau(w) \leqq c \quad \text{for } w \in \bar{\Omega}, \tag{30}$$

$$\tau(w) < c \quad \text{for } w \in \partial\Omega. \tag{31}$$

Let us assume, that the assertion of Theorem 5 is not true. Then the set $\Omega^* = \{w \in \Omega \colon \tau(w) = c\}$ is non-empty. Moreover, (31) implies $\partial\Omega^* \subset \Omega$. Thus $\partial\Omega^*$ is non-empty. Take any point $w_0 \in \partial\Omega^*$. Then $\tau(w_0) = c$, and in any neighborhood Ω^+ of w_0 there is a point w such that $\tau(w) < c$. Introduce local coordinates $x = \varphi(p)$ by a chart $\varphi \colon U \to \mathbb{R}^3$ of a neighborhood $U \subset M$ of $f(w_0)$, and choose a neighborhood $\Omega_1 \subset \Omega$ of w_0 such that $f(\Omega_1) \subset U$, and set $X = \varphi \circ f|_{\Omega_1}$. First of all, since $f \not\equiv \mathrm{const}$ on Ω, we get $X \not\equiv \mathrm{const}$ on Ω_1 (cf. [15]), and, by an observation due to Heinz (cf. [13–15]), we have the asymptotic expansion

$$2X_w(w) = X_u(w) - iX_v(w) = (a - ib)(w - w_0)^\nu + o(|w - w_0|^\nu), \tag{32}$$

for $w \in \Omega_1$, provided that Ω_1 is a sufficiently small neighborhood of w_0. In (32) ν stands for a non-negative integer, and a, b are vectors in $\mathbb{R}^3$ with $|a|^2+|b|^2 \neq 0$, and $\|a\|=\|b\|$, $\langle a, b\rangle=0$. By (32) we obtain

$$\lambda(w)=\|X_u(w)\|_{X(w)}=\|X_v(w)\|_{X(w)}>0 \quad \text{for } w \in \Omega_1-\{w_0\}, \tag{33}$$

and $\lambda(w_0)=0$ if and only if w_0 is a branch point. For $w \in \Omega_1-\{w_0\}$, we define

$$\xi_1(w)=\lambda^{-1}(w)\, X_u(w), \qquad \xi_2(w)=\lambda^{-1}(w)\, X_v(w). \tag{34}$$

In virtue of (32),

$$\xi_3(w)=\lim_{w^* \to w} \xi_1(w^*) \times \xi_2(w^*) \tag{35}$$

exists for all $w \in \Omega_1$, and $\xi_3 \in C^0(\Omega_1, \mathbb{R}^3)$. Hence, for each $w \in \Omega'-\{w_0\}$, $\xi_1(w), \xi_2(w), \xi_3(w)$ form an orthonormal basis of $\mathbb{R}^3$ with respect to the scalar product

$$\langle \eta, \zeta\rangle_{X(w)}=\eta \cdot G[X(w)]\, \zeta, \qquad (\eta, \zeta \in \mathbb{R}^3).$$

Moreover, if we set $t=\sigma \circ \varphi^{-1}$, then $\tau(w)=t(X(w))$, and $\nabla_x t(X(w)) \neq 0$ for $w \in \Omega_1$, provided that Ω_1 is a sufficiently small neighborhood of w_0. Then we can define the unit vector

$$\xi=-\|\nabla_x t(X)\|^{-1}\, \nabla_x t(X) \tag{36}$$

on Ω_1. On account of (24), (29), (33)–(36), and

$$\Delta \tau=X_u \cdot t_{xx}(X)\, X_u+X_v \cdot t_{xx}(X)\, X_v+t_x(X) \cdot \Delta X,$$

we derive on $\Omega_1-\{w_0\}$

$$\lambda^{-2} \Delta \tau=\xi_1 \cdot T(X)\, \xi_1+\xi_2 \cdot T(X)\, \xi_2-2\mathscr{H}\, \|\nabla_x t(X)\|\, \langle \xi, \xi_3\rangle. \tag{37}$$

Furthermore, for any orthonormal system η_1, η_2, η_3 in $\mathbb{R}^3$ with respect to $\langle \eta, \zeta\rangle_x=\eta \cdot G(x)\, \zeta$, we obtain

$$\Delta_x t(x)=\eta_1 \cdot T(x)\, \eta_1+\eta_2 \cdot T(x)\, \eta_2+\eta_3 \cdot T(x)\, \eta_3. \tag{38}$$

Combining (25), (37), and (38), we arrive at

$$\Delta \tau=\lambda^2 g \quad \text{on } \Omega_1, \tag{39}$$

where

$$g=2\, \|\nabla_x t(X)\|\, \{\Lambda_\tau(f)-\mathscr{H}\langle \xi, \xi_3\rangle\}+\xi \cdot T(X)\, \xi-\xi_3 \cdot T(X)\, \xi_3, \tag{40}$$

since $\Delta \tau, \lambda, g \in C^0(\Omega_1, \mathbb{R})$.

Moreover, since τ assumes in w_0 a maximum, we may conclude that

$$\xi_3(w_0)=\pm \xi(w_0). \tag{41}$$

Otherwise, it follows from (32) that at least one of the scalar products $\langle \xi(w_0), a\rangle_{X(w_0)}$, $\langle \xi(w_0), b\rangle_{X(w_0)}$ is different from zero. For instance, let us assume that

$$\langle \xi(w_0), a\rangle_{X(w_0)} < 0,$$

or

$$\langle \nabla_x t[X(w_0)], a\rangle_{X(w_0)} > 0.$$

Then we have for a disc $B_\rho(w_0) = \{w: |w - w_0| < \rho\} \subset \Omega_1$ with sufficiently small radius $\rho > 0$

$$\langle \nabla_x t[X(w)], a\rangle_{X(w)} > 0 \qquad \text{for all } w \in B_\rho(w_0).$$

Hence, for $w = w_0 + s$, $0 < s < \rho$, we obtain

$$0 < t_x[X(w_0+s)] \cdot X_u(w_0+s) = \frac{\partial}{\partial s} t[X(w_0+s)] = \frac{\partial}{\partial s} \tau(w_0+s),$$

since

$$X_u(w_0+s) = a s^\nu + o(s^\nu) \qquad \text{for } 0 \leqq s < \rho.$$

Therefore

$$\tau(w_0+s) = \tau(w_0) + \int_0^s \frac{\partial}{\partial s} \tau(w_0 + \underline{s})\, d\underline{s} > \tau(w_0) \qquad \text{for } 0 < s < \rho$$

which contradicts the maximum property of w_0.

The other three cases can be dealt with in a similar way, and (41) is proved. Thus,

$$\{\xi \cdot T(X)\,\xi - \xi_3 \cdot T(X)\,\xi_3\}_{w=w_0} = 0. \tag{42}$$

In virtue of (40), (42), and assumption (iii), we arrive at

$$g(w) \geqq \varepsilon > 0 \qquad \text{for all } w \in \Omega_2$$

for a sufficiently small neighborhood $\Omega_2 \subset \Omega_1$ of w_0. Together with (39) this relation implies

$$\Delta\tau \geqq 0 \qquad \text{on } \Omega_2.$$

Then, relation $\tau(w_0) = c$ yields $\tau(w) \equiv c$ on Ω_2. But this is a contradiction to $w_0 \in \partial\Omega^*$, and Theorem 5 is proved.

Proof of Theorem 6. Set

$$c^* = \sup_{w \in \Omega} \tau(w).$$

Then, by the assumptions of the Theorem, $c^* \leqq c$. We have to show that $c^* < c_1$. Suppose, that $c^* \geqq c_1$. Then, applying the conclusions of the proof of Theorem 5 to S_{c^*} and $\Omega^* = \{w \in \Omega: \tau(w) = c^*\}$ we derive in an analogous way a contradiction. Consequently, $c^* < c_1$, which is equivalent to $f(\bar{\Omega}) \subset J_{c_1}$.

Proof of Theorem 7. We may assume that $x=0$ is the star-point of J. Then, there exists a function $t \in C^0(E^3, \mathbb{R}) \cap C^2(E^3 - \{0\}, \mathbb{R})$ with $t(0)=0$, and

$$t(\rho x) = \rho t(x) \quad \text{for any } x \in E^3 \text{ and } \rho > 0,$$

such that

$$J = \{x \in E^3 \colon t(x) < 1\}, \qquad S = \{x \in E^3 \colon t(x) = 1\},$$

$t(x) \geqq 0$ as well as $\operatorname{grad}_x t(x) \neq 0$ if $x \neq 0$. Then, there exists a number $r > 0$ such that $X(\bar{\Omega}) \subset J_r \cup S_r$ where J_ρ and S_ρ are defined by

$$J_\rho = \{x \in E^3 \colon t(x) < \rho\}, \qquad S_\rho = \{x \in E^3 \colon t(x) = \rho\}.$$

We consider a new minimal surface $X^{(r)} = r^{-1} X$ which satisfies

$$X^{(r)}(\bar{\Omega}) \subset J_1 \cup S_1 \quad \text{and} \quad X^{(r)}(\partial \Omega) \subset J_{1/r} \cup S_{1/r}.$$

Theorem 6 yields $X^{(r)}(\bar{\Omega}) \subset J_{1/r} \cup S_{1/r}$, therefore, $X(\bar{\Omega}) \subset J_1 \cup S_1$. Finally, we can prove $X(\Omega) \cap S_1 = \varnothing$ by using similar considerations as in the proof of Theorem 5, which shows that $X(\Omega) \subset J_1 = J$.

Proof of Theorem 8. The function

$$\mathscr{H}^*(w) = \begin{cases} \mathscr{H}(w) \\ \Lambda_c(f(w)) \end{cases} \text{for} \quad \begin{cases} w \in \Omega' \\ w \in \Omega^* \end{cases}$$

is of class $L_\infty^{\text{loc}}(\Omega)$. We claim that

$$D_\alpha f_*(U_\alpha) = 2 \mathscr{H}^* f_*(U_1) \times f_*(U_2) \quad \text{a.e. on } \Omega. \tag{43}$$

Obviously, it suffices to prove

$$D_\alpha f_*(U_\alpha) = 2 \Lambda_c(f(w)) f_*(U_1) \times f_*(U_2) \quad \text{a.e. on } \Omega^*. \tag{44}$$

First, we introduce the set

$$\Omega^{**} = \{w \in \Omega \colon \|f_*(U_1)\| = \|f_*(U_2)\| = 0\}.$$

Then we pick some $w_0 \in \Omega$, introduce local coordinates $x = \varphi(p)$ by a chart $\varphi \colon U \to \mathbb{R}^3$ of a neighborhood U of $f(w_0)$ in M, and choose a neighborhood $\Omega_1 \subset\subset \Omega$ of w_0 such that $f(\Omega_1) \subset U$. Set $X = \varphi \circ f|_{\Omega_1}$. Then we have to show that

$$\Delta X^l + \Gamma_{jk}^l(X) \{X_u^j X_u^k + X_v^j X_v^k\} = 2 \Lambda_c(f(w)) (X_u \times X_v)^l \tag{45}$$

holds a.e. on $\Omega_1 \cap \Omega^*$ $(l = 1, 2, 3)$.

On $\Omega_1 \cap \Omega^{**}$ we have $X_u = 0$, $X_v = 0$. Since $X \in H_2^2(\Omega_1, \mathbb{R}^3)$, a well known theorem due to Morrey implies $\nabla^2 X = 0$ and, in particular,

$\Delta X=0$ a.e. on $\Omega_1 \cap \Omega^{**}$. That is, (45) holds a.e. on $\Omega_1 \cap \Omega^{**}$. On the other hand,

$$\sigma(f(w))=t(X(w))=c \quad \text{for } w\in\Omega^* \cap \Omega_1.$$

In virtue of $t\in C^2(\varphi(U),\mathbb{R})$ and $X\in H_2^2(\Omega_1,\mathbb{R}^3)$, Morrey's Theorem yields

$$\left.\begin{aligned} 0&=t_x(X)\cdot X_u=\langle \nabla_x t(X), X_u\rangle \\ 0&=t_x(X)\cdot X_v=\langle \nabla_x t(X), X_v\rangle \end{aligned}\right. \quad \text{a.e. on } \Omega^*\cap\Omega_1.$$

Furthermore,

$$\|X_u(w)\|=\|X_v(w)\|\neq 0, \qquad \langle X_u(w), X_v(w)\rangle=0$$

for $w\in\Omega_1\cap(\Omega^*-\Omega^{**})$.

By standard calculations, we verify (45) also a.e. on $\Omega_1\cap(\Omega^*-\Omega^{**})$, and (43) is proved. Following an observation due to Heinz (cf. [15]) we may apply a device of Hartman-Wintner to obtain for f the asymptotic expansion (32) around any point $w_0\in\Omega$. Specifically, branch points of f are isolated, and Ω^{**} contains only denumerably many points.

Now we can reason as in the proof of Theorem 5. If Ω^* is nonempty, then there exists a point $w_0\in\partial\Omega^*\subset\Omega$. On a neighborhood of $f(w_0)$ we introduce local coordinates $x=\varphi(p)$ as before, and write $X=\varphi\circ f|_{\Omega_1}$ which is of class $C^1(\Omega_1,\mathbb{R}^3)\cap H_2^2(\Omega_1,\mathbb{R}^3)$ on some open neighborhood $\Omega_1\subset\subset\Omega$ of w_0. Then $\tau=t\circ X$ is of class $C^1(\Omega_1,\mathbb{R})\cap H_2^2(\Omega_1,\mathbb{R})$ and satisfies

$$\Delta\tau=\begin{cases} 0 & \text{a.e. on } \Omega_1\cap\Omega^* \\ \lambda^2 g & \text{a.e. on } \Omega_1\cap\Omega' \end{cases}$$

where

$$\lambda=\|X_u\|=\|X_v\|,$$

and

$$g=2\,\|\nabla_x t(X)\|\,\{A_\tau(f)-\mathscr{H}\langle\xi,\xi_3\rangle\}+\xi\cdot T(X)\,\xi-\xi_3\cdot T(X)\,\xi_3.$$

Observe that $\lambda, g\in C^0(\Omega_1,\mathbb{R})$, and

$$\xi_3(w_0)=\pm\xi(w_0),$$

because of the expansion (32). Therefore,

$$\Delta\tau\geqq 0 \qquad \text{a.e. on } \Omega_2$$

for some sufficiently small open neighborhood $\Omega_2\subset\Omega$ of w_0. Furthermore, $\tau\in C^0(\Omega_2)$. Thus we obtain by a simple approximation device that τ is subharmonic on Ω_2. Since $\tau(w_0)=c$, and $\tau(w)\leqq c$ for $w\in\Omega_2$, the maximum principle implies $\tau(w)\equiv c$ on Ω_2 which contradicts $w_0\in\partial\Omega^*$, and Theorem 8 is proved.

Bibliography

1. Beckenbach, E., Radó, T.: Subharmonic functions and minimal surfaces. Trans. Amer. Math. Soc. **35**, 648-661 (1933).
2. Bliss, G. A.: Calculus of Variations. Open Court Publ. Co. La Salle 1925.
3. Courant, R.: Dirichlet's principle, conformal mapping, and minimal surfaces. New York: Interscience Publ. 1950.
4. Delauney, C.: Sur la surface de revolution dont la courbure moyenne est constante. J. Math. Pures Appl. **6**, 309-315 (1841).
5. Dombrowski, P.: Krümmungsgrößen gleichungsdefinierter Untermannigfaltigkeiten Riemannscher Mannigfaltigkeiten. Math. Nachr. **38**, 133-180 (1968).
6. Douglas, J.: The problem of Plateau for two contours. J. Math. Phys. **10**, 315-359 (1931).
7. Douglas, J.: Minimal surfaces of higher topological structure. Ann. of Math. **40**, 205 298 (1939).
8. Gulliver, R. D.: The Plateau problem for surfaces of prescribed mean curvature in a Riemannian manifold. To appear.
9. Gulliver, R. D., Osserman, R., Royden, H. L.: A theory of branched immersions of surfaces. Preprint.
10. Gulliver, R. D., Spruck, J.: Surfaces of constant mean curvature which have a simple projection. Math. Z., to appear.
11. Heinz, E.: Über die Existenz einer Fläche konstanter mittlerer Krümmung bei vorgegebener Berandung. Math. Ann. **127**, 258-287 (1954).
12. Heinz, E.: On the nonexistence of a surface of constant mean curvature with finite area and prescribed rectifiable boundary. Arch. Rat. Mech. Analysis **35**, 249-252 (1969).
13. Heinz, E., Hildebrandt, S.: The number of branch points of surfaces of bounded mean curvature. J. Diff. Geometry **4**, 227-235 (1970).
14. Heinz, E., Hildebrandt, S.: Some remarks on minimal surfaces in Riemannian manifolds. Commun. Pure Appl. Math. **23**, 371 377 (1970).
15. Hildebrandt, S.: Einige Bemerkungen über Flächen beschränkter mittlerer Krümmung. Math. Z. **115**, 169-178 (1970).
16. Hildebrandt, S.: On the regularity of solutions of two-dimensional variational problems with obstructions. Commun. Pure Appl. Math., to appear.
17. Hildebrandt, S., Kaul, H.: Two-dimensional variational problems with obstructions and Plateau's problem for H-surfaces in a Riemannian manifold. Commun. Pure Appl. Math. **25**, 187-223 (1972).
18. Kaul, H.: Isoperimetrische Ungleichung und Gauß-Bonnet-Formel für H-Flächen in Riemannschen Mannigfaltigkeiten. Arch. Rat. Mech. Analysis **45**, 194-221 (1972).
19. Kaul, H.: Ein Einschließungssatz für H-Flächen in Riemannschen Mannigfaltigkeiten. Manuscripta Math. **5**, 103-112 (1971).
20. Lawson, B.: The global behavior of minimal surfaces in S^n. Ann. of Math. **92**, 224-237 (1970).
21. Nitsche, J. C. C.: On new results in the theory of minimal surfaces. Bull. Amer. Math. Soc. **71**, 195-270 (1965).
22. Nitsche, J. C. C.: A necessary criterion for the existence of certain minimal surfaces. J. Math. Mech. **13**, 659-665 (1964).
23. Nitsche, J. C. C.: A supplement to the condition of J. Douglas. Rend. Circ. Mat. Palermo (2) **13**, 192-198 (1964).
24. Nitsche, J. C. C.: Ein Einschließungssatz für Minimalflächen. Math. Ann. **165**, 71 75 (1966).
25. Nitsche, J. C. C.: Note on the non-existence of minimal surfaces. Proc. Amer. Math. Soc. **19**, 1303-1305 (1968).
26. Nitsche, J. C. C., Leavitt, J.: Numerical estimates for minimal surfaces. Math. Ann. **180**, 170-174 (1969).

27. Serrin, J.: On surfaces of constant mean curvature which span a given space curve. Math. Z. **112**, 77–88 (1969).
28. Sinclair, E.: On the minimum surface of revolution in the case of one variable end point. Ann. of Math. **8**, 177–188 (1906–1907).
29. Tomi, F.: Variationsprobleme vom Dirichlet-Typ mit einer Ungleichung als Nebenbedingung. Math. Z., to appear.
30. Thompson, D'Arcy W.: On growth and form. Cambridge University Press 1969.
31. Werner, H.: Das Problem von Douglas für Flächen konstanter mittlerer Krümmung. Math. Ann. **133**, 303–319 (1957).

Prof. Stefan Hildebrandt
Mathematisches Institut der Universität
D-53 Bonn, Wegelerstraße 10
Sonderforschungsbereich Theoretische Mathematik
Federal Republic of Germany

(Received May 21, 1972)

Math. Z. 128, 271–276 (1972)

Remarks on the Isoperimetric Inequality for Multiply-Connected H-Surfaces

Helmut Kaul

In the paper [3], the author has established an isoperimetric inequality for H-surfaces in the euclidean space E^3, which are bounded by finitely many closed Jordan curves $\Gamma_1, \ldots, \Gamma_n$ ([3], Satz 8). One disadvantage of this estimate is the appearance of the distances of $\Gamma_1, \ldots, \Gamma_n$ to some fixed point. This dependance shall be eliminated in this note in order to bring the inequality into the form $A \leqq \mathrm{const} \cdot L^2$. This will be achieved by taking into account results, which are obtained by Hildebrandt in the previous paper [1]. Using a maximum principle, Hildebrandt has proved in a simple manner, that the boundary of a connected H-surface can not be "separated" by certain cones. From this fact the distance of two components Γ_1, Γ_2 of the boundary of a connected H-surface can be easily estimated in terms of their diameter (and therefore also in terms of their length) by means of elementary geometric considerations.

Nitsche has established such estimates for minimal surfaces by different methods (see the bibliography in [1]). He applied them to prove an isoperimetric inequality for a multiply-connected minimal surface [4].

The problem of this note was suggested to me by Hildebrandt.

Throughout we use the following notations: Let $\Gamma_1, \Gamma_2 \subset E^3$ be disjoint closed C^2-Jordan curves and $\mathfrak{x}: \bar{\Omega} \to E^3$ an H-surface (for the definition, see [1] (7), for instance), bounded by $\Gamma_1 \cup \Gamma_2$, where Ω is an open circular annulus. (We restrict our considerations to Jordan curves, but we guess, that the results remain true, if Γ_1 and Γ_2 are replaced by general curves.) Denote by A the area of $\mathfrak{x}$, δ_i the diameter of Γ_i, L_i the length of Γ_i $(i=1,2)$, let $\mathfrak{x}_i$ be the center, r_i the radius of the unique smallest closed ball $\mathfrak{K}_i$ containing Γ_i, and set $L := L_1 + L_2$, $R := |\mathfrak{x}_1 - \mathfrak{x}_2|$, $\rho := \mathrm{dist}(\Gamma_1, \Gamma_2)$. Furthermore, we write

$$q(\mathfrak{a}) := \sup_{w \in \Omega} |H \circ \mathfrak{x}(w)| \cdot |\mathfrak{x}(w) - \mathfrak{a}| \tag{1}$$

for $\mathfrak{a} \in E^3$, and denote by

$$\mathfrak{x}_0 := \frac{1}{r_1 + r_2} \cdot (r_2 \mathfrak{x}_1 + r_1 \mathfrak{x}_2) \tag{2}$$

the "center" of $\Gamma_1 \cup \Gamma_2$.

In virtue of [3], Satz 8 (for $n=2$, $\theta=1$), by using

$$|B(\mathfrak{a})| = \Big|\int_\Omega \langle H \circ \mathfrak{x}, \mathfrak{x} - \mathfrak{a} \rangle \, \omega \Big| \leqq A \cdot q(\mathfrak{a}),$$

we obtain the following isoperimetric inequality

$$4\pi A \leqq \frac{\sum_{i=1}^{2} L_i(L_i + 2\pi \cdot \operatorname{dist}(\mathfrak{a}, \Gamma_i))}{1 - q(\mathfrak{a})} \tag{3}$$

for every $\mathfrak{a} \in E^3$ which satisfies $q(\mathfrak{a}) < 1$. This formula will be the starting point of our investigations.

First we deal with minimal surfaces.

Theorem 1. *For every minimal surface, bounded by Γ_1 and Γ_2, the isoperimetric inequality*

$$A \leqq c \cdot \frac{L^2}{4\pi} \tag{4}$$

holds, where

$$c \leqq \tfrac{1}{2}\left(1 + \tfrac{3}{4}\pi\sqrt{2}\right) \approx 2.17. \tag{5}$$

For special configurations of Γ_1 and Γ_2 we obtain better upper bounds for c:

(i) *If Γ_1, Γ_2 are contained in parallel planes which have distance $r \geqq 0$, then*

$$c \leqq \frac{1}{2}\left(1 + \pi\sqrt{\frac{3R^2 + r^2}{3(R^2 + r^2)}}\right). \tag{6}$$

(ii) *If Γ_1, Γ_2 are circles in the same position as in* (i), *then* (6) *can be improved to*

$$c \leqq \frac{1}{2}\left(1 + \sqrt{\frac{3R^2 + r^2}{R^2 + r^2}}\right). \tag{7}$$

Now, we state the corresponding results for H-surfaces.

Theorem 2. *Let $\mathfrak{x}: \bar{\Omega} \to E^3$ be an H-surface bounded by $\Gamma_1 \cup \Gamma_2$. If the condition*

$$q(\mathfrak{x}_0) < \tfrac{1}{2} \tag{8}$$

(q, $\mathfrak{x}_0$ *defined by* (1), (2)) *is satisfied, the isoperimetric inequality*

$$A \leqq \frac{c_*}{4\pi} \cdot \frac{L_1^2 + L_2^2}{1 - q(\mathfrak{x}_0)} \tag{9}$$

holds, where

$$c_* \leqq 1 + \pi \sqrt{\frac{3(1+2b)}{8b}} \tag{10}$$

and $b := 1 - 2q(\mathfrak{x}_0)$.

Furthermore, we have the estimates

$$R \leqq \sqrt{\frac{3(1+b)}{8b}} \cdot (\delta_1 + \delta_2) \tag{11}$$

and

$$\rho \leqq \sqrt{\frac{3(1+2b)}{8b}} \cdot (\delta_1 + \delta_2). \tag{12}$$

Better upper bounds for $R(\delta_1 + \delta_2)^{-1}$ *and* c_* *are obtained in the case of special configurations of* Γ_1 *and* Γ_2:

(i) *If* Γ_1, Γ_2 *are contained in parallel planes which have distance* $r \geqq 0$, *then*

$$\sqrt{3b\left(r^2 + \frac{d^2}{1+b}\right)} \leqq \delta_1 + \delta_2 \tag{13}$$

$(d := \sqrt{R^2 - r^2})$, *and*

$$c_* \leqq 1 + \pi \sqrt{\frac{(1+2b)R^2 + b^2 r^2}{3b(R^2 + b r^2)}}. \tag{14}$$

(ii) *If* Γ_1, Γ_2 *are circles in the same position as in* (i), *the extimate* (14) *can be improved to*

$$c_* \leqq 1 + \sqrt{\frac{(1+2b)R^2 + b^2 r^2}{b(R^2 + b r^2)}}. \tag{15}$$

Proof of Theorem 1 and 2. Let $\mathfrak{x}$ be a minimal surface. Assume $L_1 \leqq L_2$, choose in (3) $\mathfrak{a} \in \Gamma_2$ with $\operatorname{dist}(\mathfrak{a}, \Gamma_1) = \rho$ and set $\mu := \max\{1, \pi \rho L_2^{-1}\}$. Then, we obtain from (3)

$$4\pi A \leqq L_1(L_1 + 2\pi\rho) + L_2 L_2 \leqq L^2 + 2L_1 L_2(\mu - 1) \leqq L^2 + \tfrac{1}{2}L^2(\mu - 1) = c \cdot L^2,$$

where $c := \frac{1}{2}(1+\mu)$. If $\mathfrak{x}$ is an H-surface, we choose in (3) $\mathfrak{a} = \mathfrak{x}_0 = \dfrac{1}{r_1 + r_2}(r_2 \mathfrak{x}_1 + r_1 \mathfrak{x}_2)$ and obtain

$$\operatorname{dist}(\mathfrak{x}_0, \Gamma_i)^2 \leqq |\mathfrak{x}_0 - \mathfrak{x}_i|^2 + \operatorname{dist}(\mathfrak{x}_i, \Gamma_i)^2 \leqq \left(\frac{R r_i}{r_1 + r_2}\right)^2 + r_i^2.$$

Combining this estimate with (3), we arrive

$$4\pi A(1-q(\mathfrak{x}_0)) \leqq \sum_{i=1}^{2} L_i \left(L_i + 2\pi r_i \sqrt{1+\left(\frac{R}{r_1+r_2}\right)^2}\right).$$

By Jung's theorem [2], we have $r_i \leqq \sqrt{\frac{3}{8}}\,\delta_i \leqq \frac{1}{2}\sqrt{\frac{3}{8}}\,L_i$, and, if Γ_i is a plane curve, $r_i \leqq \frac{1}{\sqrt{3}}\,\delta_i \leqq \frac{1}{2\sqrt{3}}\,L_i$. Thus, we obtain the isoperimetric inequality

$$4\pi A(1-q(\mathfrak{x}_0)) \leqq c_*(L_1^2+L_2^2),$$

taking

$$c_* := \begin{cases} 1+\pi\sqrt{\dfrac{3}{8}\left(1+\left(\dfrac{R}{r_1+r_2}\right)^2\right)}, & \\ 1+\pi\sqrt{\dfrac{1}{3}\left(1+\left(\dfrac{R}{r_1+r_2}\right)^2\right)}, & \text{if } \Gamma_1, \Gamma_2 \text{ are plane curves,} \\ 1+\sqrt{1+\left(\dfrac{R}{r_1+r_2}\right)^2}, & \text{if } \Gamma_1, \Gamma_2 \text{ are circles.} \end{cases} \tag{16}$$

Now, it remains to estimate $R(r_1+r_2)^{-1}$ and ρL_2^{-1}. This will be done by using Hildebrandt's "cone separating theorem" ([1], Theorem 4). To simplify the considerations, we shift $\mathfrak{x}_0$ by a translation into the origin. One achieves by a rotation around $\mathfrak{x}_0=0$, that

$$\mathfrak{x}_1=(0,0,R r_1(r_1+r_2)^{-1}) \quad \text{and} \quad \mathfrak{x}_2=(0,0,-R r_2(r_1+r_2)^{-1}).$$

First, we prove

$$R \leqq \sqrt{\frac{1+b}{b}}\,(r_1+r_2). \tag{17}$$

Otherwise, each of the balls $\mathfrak{K}_i$ (defined as above) were contained in one of the half cones $\{(x,y,z)\in E^3:\ x^2+y^2-bz^2<0,\ z\gtrless 0\}$, contradicting Theorem 4 in [1].

By using $r_i \leqq \sqrt{\frac{3}{8}}\,\delta_i$ and the inequality

$$\rho^2 \leqq R^2+r_1^2+r_2^2, \tag{18}$$

we get the estimates (10), (11), and (12). For a minimal surface ($b=1$, $L_1 \leqq L_2$), (12) implies $\rho \leqq \sqrt{\frac{9}{8}}\cdot(\delta_1+\delta_2) \leqq \frac{3\sqrt{2}}{4}\,L_2$, hence $\mu \leqq \frac{3}{4}\sqrt{2}\,\pi$, such that (5) is proved.

In the case (i), let $\mathfrak{D}_i$ be the intersection of $\mathfrak{K}_i$ with the plane containing Γ_i. According to the following lemma, by using an appropriate rotation around $\mathfrak{x}_0=0$, and [1], Theorem 4, we find

$$\sqrt{b\left(r^2+\frac{d^2}{1+b}\right)} \leqq r_1+r_2 \leqq \frac{1}{\sqrt{3}}\,(\delta_1+\delta_2),$$

where $d := \sqrt{R^2 - r^2}$, or, equivalently

$$R \leqq \sqrt{\frac{(1+b) R^2}{b(R^2 + b r^2)}} (r_1 + r_2). \tag{19}$$

This, together with (16), implies (13) and (14).

If $\mathfrak{x}$ is a minimal surface ($b = 1, L_1 \leqq L_2$), we obtain (6) by using (18),

$$r_i \leqq \frac{1}{\sqrt{3}} \delta_i \leqq \frac{1}{2\sqrt{3}} L_i \leqq \frac{1}{2\sqrt{3}} L_2$$

and

$$\rho \leqq \sqrt{\frac{3R^2 + r^2}{R^2 + r^2}} (r_1 + r_2). \tag{20}$$

In the case (ii), one derives (15) from (16) and (19). The estimate (7) follows from (20) by using $r_i = L_i/2\pi \leqq L_2/2\pi$.

Now, we state the

Lemma. *Let b be a positive number, and suppose that $\mathfrak{D}_1, \mathfrak{D}_2$ are two closed discs with center $\mathfrak{x}_1, \mathfrak{x}_2$ and radius r_1, r_2, respectively, which are contained in parallel planes of distance $r \geqq 0$. Set $R := |\mathfrak{x}_1 - \mathfrak{x}_2|$ and $d := \sqrt{R^2 - r^2}$.*

Then $\mathfrak{D}_1, \mathfrak{D}_2$ are "separated" by a cone (i.e. each $\mathfrak{D}_i$ is contained in one of both half cones) congruent to $x^2 + y^2 - b z^2 < 0$ with vertex $\mathfrak{x}_0 = \frac{1}{r_1 + r_2}(r_2 \mathfrak{x}_1 + r_1 \mathfrak{x}_2)$ if and only if

$$R > \sqrt{\frac{1+b}{b(1 + b r^2 R^{-2})}} (r_1 + r_2),$$

or, equivalently,

$$\sqrt{b\left(r^2 + \frac{d^2}{1+b}\right)} > r_1 + r_2.$$

Proof. Set $\sin\alpha = r/R$ and consider the following problem: For any numbers b, r_0, α ($b, r_0 > 0, 0 \leqq \alpha \leqq \pi/2$), find an angle $\phi = \phi_0$, such that the circle

$$[0, 2\pi] \ni t \mapsto \begin{pmatrix} R_0 \sin\phi + r_0 \sin(\phi + \alpha) \cdot \cos t \\ r_0 \sin t \\ R_0 \cos\varphi + r_0 \cos(\varphi + \alpha) \cdot \cos t \end{pmatrix}$$

lies in the half cone $\{(x, y, z) \in E^3 : x^2 + y^2 - b z^2 \leqq 0, z > 0\}$, and that $R_0 > 0$ attains its minimum $R_{\min}$. One verifies, that the separability of the discs $\mathfrak{D}_1, \mathfrak{D}_2$ in the statement is equivalent to

$$\frac{R}{\delta_1 + \delta_2} > \frac{R_{\min}}{r_0}.$$

Now, we will calculate the quotient $R_{\min}/r_0$. In order that the circle lies in the half cone, the relation

$$0 \leqq \cos^2 t - 2P \cdot \cos t - Q = (\cos t - P)^2 - (P^2 + Q)$$

has to be satisfied for all t, provided that we assume $\cos(\phi + \alpha) \neq 0$. Here, we have used the notations

$$P := \frac{R_0}{r_0} \cdot \frac{\cos\alpha - (1+b)\cos\phi \cdot \cos(\phi+\alpha)}{(1+b)\cos^2(\phi+\alpha)},$$

$$Q := \frac{1 + R_0^2 r_0^{-2}(1-(1+b)\cos^2\phi)}{(1+b)\cos^2(\phi+\alpha)}.$$

From the condition $P^2 + Q = 0$ one gets

$$R_0^2 = r_0^2 \frac{(1+b)(1-\cos(\psi+2\alpha))}{2b\cos^2\alpha - (1+b)(1-\cos\psi)}.$$

where $\psi = 2\phi$. To find the minimum of R_0, we determine ψ_0 with $\frac{\partial R_0^2}{\partial \psi}(\psi_0) = 0$. After some calculations, we arrive at

$$\cos\psi_0 = 1 - \frac{2b^2 \sin^2\alpha \cdot \cos^2\alpha}{1 + b(2+b)\sin^2\alpha}, \qquad \sin\psi_0 = \frac{2b\sin\alpha\cdot\cos\alpha\cdot(1+b\sin^2\alpha)}{1+b(2+b)\sin^2\alpha},$$

which gives

$$R_{\min} = r_0 \sqrt{\frac{1+b}{b(1+b\sin^2\alpha)}}$$

as minimal value of R_0.

Bibliography

1. Hildebrandt, S.: Maximum principles for minimal surfaces and for surfaces of continuous mean curvature. Math. Z. **128**, 157–173 (1972).
2. Jung, H.: Über die kleinste Kugel, die eine räumliche Figur einschließt. J. für Math. **123**, 241–257 (1901).
3. Kaul, H.: Isoperimetrische Ungleichung und Gauss-Bonnet-Formel für H-Flächen in Riemannschen Mannigfaltigkeiten. Arch. Rat. Mech. Analysis **45**, 194–221 (1972).
4. Nitsche, J.C.C.: The isoperimetric inequality for multiply-connected minimal surfaces. Math. Ann. **160**, 370–375 (1965).

Dr. Helmut Kaul
Mathematisches Institut der Universität
D-5300 Bonn
Wegelerstr. 10
Federal Republic of Germany
Sonderforschungsbereich Theoretische Mathematik

(Eingegangen am 22. Juni 1972)

Math. Z. 128, 277–296 (1972)

Mappings Into Loop Spaces and Central Group Extensions

Lawrence L. Larmore and Emery Thomas

Introduction

There are several situations in algebraic topology where one considers the group formed by the homotopy classes of maps of one space into the loops on another – say, $[X, \Omega Y]$.

Examples are:

(i) the fundamental group of a function space; for

$$\pi_1(Y^X)=[S^1, Y^X]=[S^1 \wedge X, Y]=[X, Y^{S^1}]=[X, \Omega Y].$$

(ii) mappings into a topological group. If G denotes a topological group (or associative H-space), then G has a classifying space B and $[X, G]=[X, \Omega B]$.

(iii) generalized cohomology theories. If h^* is such a theory, then for each integer i there is a C_i such that

$$h^i(X)=[X, \Omega C_i].$$

Now any space Y can be factored via its Postnikov resolution into a tower of principal fibrations. The purpose of this paper is to compute the group $[X, \Omega E]$, when E is the total space of a principal fibration. We regard this as an extension problem (see § 1), with two of the three groups involved abelian. If E is a multiple loop space, then the extension itself is abelian. This is the situation in example (iii); and in two previous papers [6, 8] we have developed a theory for computing abelian group extensions arising in this way.

Our concern here is to compute $[X, \Omega E]$ even if this is a non-abelian group – this is the situation that obtains in general in examples (i) and (ii). We show in §1 that the extension involved is a central extension; Part I of the paper then shows how one can compute central group extensions using invariants given in terms of the groups involved. In Part II we study the problem of computing the group $[X, SO(n)]$, $n \geq 3$, where $SO(n)$ denotes the rotation group and where X is a complex with

$\dim X \leqq n$ if $n \equiv 3 \bmod 4$ and $\dim X \leqq n-1$ if $n \not\equiv 3 \bmod 4$. (Note that $[X, SO(n)]$ can be construed as the group of oriented n-plane bundles over the suspension of X.) We reduce this to an extension problem, and then show that this extension can be computed by means of K-theory cohomology. As a special case, in the last section, we compute the group $[X, SO(3)]$, $\dim X \leqq 3$, expressing the answer in terms of singular cohomology. (If X is an orientable 3-manifold then this group can be thought of as the group of vector 3-fields on M; cf. [8].) As a final example, illustrating the interesting results that can arise, we show that if S^1 denotes the circle and P^2 the real projective plane, then

$$[S^1 \times P^2, SO(3)] = \text{dihedral group of order eight.}$$

Part I

§ 1. The Extension Problem

Let K and L be pointed spaces, and θ: $K \to L$ a map. Consider the fibration sequence:

$$\ldots \Omega^2 K \xrightarrow{\Omega^2\theta} \Omega^2 L \xrightarrow{\Omega\lambda} \Omega E \xrightarrow{\Omega p} \Omega K \xrightarrow{\Omega\theta} \Omega L \xrightarrow{\lambda} E \xrightarrow{p} K \xrightarrow{\theta} L$$

where $E = \{(k, \sigma) \in K \times PL \mid \theta(k) = \sigma(1)\}$ is the fiber of θ, $p(k, \sigma) = k$, etc. If X is any pointed C.W. complex, we have an exact sequence of groups:

$$e\colon\ 0 \to \operatorname{Coker} \Omega^2 \theta_* \xrightarrow{\Omega\lambda_*} [X; \Omega E] \to \operatorname{Ker} \Omega\theta_* \to 1.$$

The purpose of our paper is to compute extension e.

We begin by observing that extension e is in fact a central extension.

Theorem 1.1. *$(\Omega\lambda)_*[X, \Omega^2 L]$ lies in the center of $[X, \Omega E]$.*

Proof. For simplicity assume that θ is an inclusion $K \subset L$. Then a point v in ΩE can be regarded as a map v: $I^2 \to L$ such that

$$v(s,0) = v(0,t) = v(1,t) = *,$$
$$v(s,1) \in K, \quad \text{all} \quad s, t \in I.$$

Given a point u in $\Omega^2 L$, i.e., a map u: $I^2 \to L$ with $u(\dot{I}^2) = *$, we define two additions for u and v:

$$(u +' v)(s,t) = \begin{cases} u(s, 2t), & 0 \leqq t \leqq \frac{1}{2}, \\ v(s, 2t-1), & \frac{1}{2} \leqq t \leqq 1, \end{cases}$$

$$(u + v)(s,t) = \begin{cases} u(2s, t), & 0 \leqq s \leqq \frac{1}{2} \\ v(2s-1, t), & \frac{1}{2} \leqq s \leqq 1. \end{cases}$$

Thus both $u+'v$ and $u+v$ are elements in ΩE. Now let $f\colon X\to\Omega^2 L$, $g\colon X\to\Omega E$ be maps. Define

$$\begin{aligned}(f+'g)(x)&=f(x)+'g(x),\\(f+g)(x)&=f(x)+g(x),\end{aligned}\qquad x\in X.$$

Passing to homotopy classes of maps we have defined two ways of adding elements of $(\Omega\lambda)_*[X,\Omega^2 L]$ with elements in $[X,\Omega E]$. It is immediate that these two additions satisfy properties (a) and (b) of [11, 1.6.8], (use [4, 1.5] for (b)). Thus by the conclusion of 1.6.8 [11], $(\Omega\lambda)_*[X,\Omega^2 L]$ is in the center of $[X,\Omega E]$.

Hereafter, we shall assume that K and L are H-spaces; thus $[X;\Omega K]$ is Abelian. Let $R=\ker\Omega\theta_*\subset[X;\Omega K]$, $C=\operatorname{Coker}\Omega^2\theta_*$.

We define a function $\Gamma\colon R\times R\to C$ as follows: if $x,y\in R$, pick $g,h\in[X;\Omega E]$ such that $p_* g=x$ and $p_* h=y$. Let $\Gamma(x,y)=\lambda_*^{-1}(g h g^{-1} h^{-1})$. Notice that $\Gamma(x,y)$ does not depend on the choice of g and h; for suppose $p_*\hat{g}=x$ and $p_*\hat{h}=y$. Then $\hat{g}=a g$ and $\hat{h}=b h$ for a, b in the center of $[X;\Omega E]$ (by Theorem 1.1); whence $\hat{g}\hat{h}\hat{g}^{-1}\hat{h}^{-1}=g h g^{-1} h^{-1}$.

Theorem 1.2. *Γ is bilinear and alternating.*

Proof. It is obvious that $\Gamma(x,x)=0$ and that $\Gamma(y,x)=-\Gamma(x,y)$. Suppose $\hat{x}\in R$ and $p_*\hat{g}=\hat{x}$. Then $\lambda_*\Gamma(x+\hat{x},y)=g\hat{g}h\hat{g}^{-1}g^{-1}h^{-1}=g h h^{-1}\hat{g}h\hat{g}^{-1}g^{-1}h^{-1}=g h g^{-1}h^{-1}\hat{g}h\hat{g}^{-1}h^{-1}=\lambda_*(\Gamma(x,y)+\Gamma(\hat{x},y))$, since. by Theorem 1.1, $h^{-1}\hat{g}h\hat{g}^{-1}=\lambda_*\Gamma(-y,\hat{x})$ and hence commutes with g^{-1}.

We may thus describe Γ as a homomorphism $\Lambda^2 R\to C$, where $\Lambda^2 R$ is the alternating 2-fold tensor product of R.

For any integer $n\geqq 2$, we let $R_n=\{x\in R\mid n x=0\}$. We define a function (not generally a homomorphism) $\Phi_n\colon R_n\to C/nC$ as follows: $\Phi_n(x)=\{y\in C\mid\lambda_* y=g^n$ for some $g\in p^{-1}x\}$. We leave it to the reader to verify that $\Phi_n(x)$ is always a coset of nC.

We finally have the theorem whose proof we postpone to Section 3:

Theorem 1.3. *To determine the extension $e\colon 0\to C\to[X;\Omega E]\to R\to 0$, it is sufficient to know Γ, and Φ_n for $n=p^k$, all primes p and all $k\geqq 1$.*

§ 2. Computation of Γ

We assume from here on that K and L are topological groups; note that by Milnor [10], this is a mild restriction. Now $\theta\colon K\to L$ may not be a homomorphism. We define Δ $(=\Delta_\theta)$, the deviation from additivity of θ, as follows: If $x,y\in K$, let $\Delta(x,y)=\theta(x y)\,\theta(y)^{-1}\,\theta(x)^{-1}$, and so

$$\Delta\colon (K\times K, K\vee K)\to(L,*).$$

(We assume that the identity element is taken as basepoint in K and L.) We define $''\Delta\colon\Omega K\times\Omega K\to\Omega^2 L$, the suspension of Δ, as follows: If

$\sigma, \tau \in \Omega K$, and $t, u \in I$, let $''\Delta(\sigma, \tau)(t, u) = \Delta(\sigma(t), \tau(u))$. Thus $''\Delta$ factors through $\Omega K \wedge \Omega K$. Now we have a function $''\Delta_*: [X; \Omega K] \times [X; \Omega K] \to [X; \Omega^2 L]$, and:

Theorem 2.1. *If* $x, y \in R$, *then* $''\Delta_*(x, y) \in -\Gamma(x, y)$.

Proof. Let $S = S^1$, and let $\gamma: S \to S \vee S$ be the map which generates the commutator of $\pi_1(S \vee S)$; i.e., if $S = I/\partial I$:

$$\gamma(t) = \begin{cases} (4t, 0) & \text{if } 0 \leqq t \leqq \frac{1}{4} \\ (0, 4t-1) & \text{if } \frac{1}{4} \leqq t \leqq \frac{1}{2} \\ (3-4t, 0) & \text{if } \frac{1}{2} \leqq t \leqq \frac{3}{4} \\ (0, 4-4t) & \text{if } \frac{3}{4} \leqq t \leqq 1. \end{cases}$$

Now we have a cofibration sequence:

$$S \xrightarrow{\gamma} S \vee S \xrightarrow{j} T \xrightarrow{q} S^2 \xrightarrow{S\gamma} S^2 \vee S^2$$

where $T = S \times S = (S \vee S) \cup_\gamma e^2$ is the torus. As in [6], we have a commutative diagram where all rows and columns are fibration sequences:

$$\begin{tikzcd}
 & & & & (\Omega K)^S \arrow[d, "\Omega\theta^S"] \\
 & & & (\Omega L)^{S \vee S} \arrow[r, "\Omega L^\gamma"] \arrow[d, "\lambda^{S \vee S}"] & (\Omega L)^S \arrow[d, "\lambda^S"] \\
 & & & E^{S \vee S} \arrow[r, "E^\gamma"] \arrow[d, "p^{S \vee S}"] & E^S \arrow[d, "p^S"] \\
 & K^{S^2} \arrow[r, "K^q"] \arrow[d, "\theta^{S^2}"] & K^T \arrow[r, "K^j"] \arrow[d, "\theta^T"] & K^{S \vee S} \arrow[r, "K^\gamma"] \arrow[d, "\theta^{S \vee S}"] & K^S \\
L^{S^2 \vee S^2} \arrow[r, "L^{S\gamma}"] & L^{S^2} \arrow[r, "L^q"] & L^T \arrow[r, "L^j"] & L^{S \vee S} &
\end{tikzcd}$$

Suppose $x, y \in [X; \Omega K]$, and $\Omega\theta_* x = \Omega\theta_* y = 0$. Then, since $\Omega K = K^S$ and $K^S \times K^S = K^{S \vee S}$, we can consider $(x, y) \in [X; K^{S \vee S}]$. We define

$$\Gamma(x, y) = (\lambda_*^S)^{-1} E_*^\gamma (p_*^{S \vee S})^{-1}(x, y) \in [X; (\Omega L)^S] / \Omega\theta_*^S [X; (\Omega K)^S]$$

and

$$\Gamma'(x, y) = (L_*^q)^{-1} \theta_* (K_*^j)^{-1}(x, y) \in [X; L^{S^2}] / \theta^{S^2} [X; K^{S^2}].$$

We can identify $\Omega^2 L = (\Omega L)^S = L^{S^2}$, and $\Omega^2 K = (\Omega K)^S = K^{S^2}$, and we see that the definition of Γ, above, agrees with that given in § 2, while by

Theorem 2.5 of [6] we have that $\Gamma = -\Gamma'$. Now consider the diagram:

$$\begin{tikzcd}
 & {[X; K^T]} \arrow[r, "K^j_*"] \arrow[d, "\theta^T_*"] & {[X; K^{S\vee S}]} \arrow[d, "\theta^{S\vee S}_*"] \arrow[dll, "''\Delta_*"'] \\
{[X; L^{S^2}]} \arrow[r, "L^q_*"] & {[X; L^T]} \arrow[r, "L^j_*"] & {[X; L^{S\vee S}]}
\end{tikzcd}$$

Let $f, g\colon X \to K^S$ represent x and y, respectively, and let $\phi_r, \psi_r\colon X \to L^S$, $0 \leqq r \leqq 1$, be homotopies where $\phi_0 = \psi_0 = 0$ and $\phi_1 = \theta \circ f$, $\psi_1 = \theta \circ g$. Let $h\colon X \to K^T$ be given by: $h(x)(t, u) = f(x)(t)\, g(x)(u)$ for all $x \in X$, $(t, u) \in T$, and let $\alpha_r\colon X \to L^T$ be the homotopy where

$$\alpha_r(x)(t, u) = \theta\,[f(x)(t)\, g(x)(u)]\,(\phi_r(x)(t))^{-1}(\psi_r(x)(u))^{-1}$$

for all $0 \leqq r \leqq 1$, $x \in X$, and $(t, u) \in T$. We see that $K^j \circ h = (f, g)$, $\theta^T \circ h = \alpha_0$, and $L^q \circ {}''\Delta \circ (f, g) = \alpha_1$; thus we are done.

§ 3. Central Extensions

Suppose we have an exact sequence of groups:

$$e\colon\ 1 \to C \xrightarrow{j} B \xrightarrow{p} R \to 1$$

where C and R are commutative, R is finitely generated, and $j(C) \subset$ Center B. For each integer $n \geqq 2$, let $R_n = \{x \in R \mid x^n = 1\}$, and let $\Phi_n\colon R_n \to C/C^n$ (where $C^n = \{x^n \mid x \in C\}$) be defined as follows: $\Phi_n(x) = \{y \in C \mid j\,y = b^n$ for some $b \in p^{-1}x\}$. Φ_n is not necessarily a homomorphism. Let $\Gamma\colon R \times R \to C$ be defined as follows: If $x, y \in R$, pick $g, h \in B$ such that $p\,g = x$ and $p\,y = y$; define $\Gamma(x, y) = j^{-1}(g\,h\,g^{-1}h^{-1})$.

Theorem 3.1. *The extension e is determined by Γ and Φ_n for values of n equal to powers of primes.*

Proof. Consider R to be the direct sum of cyclic groups of prime power or infinite order. Let n be the number of such summands; for each $1 \leqq i \leqq n$, let r_i be the generator of the i-th summand. Without loss of generality, we may assume that for all $1 \leqq i \leqq m$, r_i has finite order $k(i)$, while for $m < i \leqq n$, r_i has infinite order, for some $0 \leqq m \leqq n$. For each $1 \leqq i \leqq m$, choose $a_i \in \Phi_{k(i)}(r_i) \subset C$. Let F^n be a free group on n generators, $x_1, \ldots, x_n$, and let $C \times F^n$ be the direct product of C with F^n. Let G be the quotient group of $C \times F^n$ obtained by imposing the following relations:

1. $x_i^{k(i)} = a_i$ for all $1 \leqq i \leqq m$.
2. $x_j x_i = \Gamma(r_j, r_i)\, x_i x_j$, $1 \leqq i < j \leqq n$.

For each $1 \leqq i \leqq n$, choose $b_i \in B$ such that $p\,b_i = r_i$; if $i \leqq m$, we require that $b_i^{k(i)} = j\,a_i$. Now define $f\colon C \times F^n \to B$ by $f(a, x_i) = j(a)\, b_i$. Define

$g\colon C\times F^n\to R$ by $g=pf$. By checking relations, one sees that f and g both factor through G, and so we obtain a commutative diagram:

$$\begin{tikzcd}
e'\colon C \arrow[r, "i"] \arrow[dd, equal] & G \arrow[rr, dashed, "\bar g"] \arrow[dd, "f"'] & & R \arrow[dd, equal] \\
& & A\times F^n \arrow[ul] \arrow[ur, dashed, "g"'] \arrow[dl, "f"] & \\
e\colon C \arrow[r, "j"'] & B \arrow[rr, "p"'] & & R
\end{tikzcd}$$

Claim: e' is a short exact sequence and so is an extension equivalent to e.

Proof. i is injective, since j is, and $\bar g$ is surjective since g is. Also, $\bar g\, i\equiv 0$, and so all we need do is show that Kernel $\bar g\subset$ Image i. Let $(a,x)\in G$ be a class such that $\bar g(a,x)=0$. Using relations 1 and 2 we can write $(a,x)=(a',x')$, where $x'=x_1^{d_1}\cdot\dots\cdot x_n^{d_n}$, with $0\leqq d_i<k(i)$, for $1\leqq i\leqq m$. Thus,

$$\bar g(a,x)=\bar g(a',x')=\sum d_i r_i=0.$$

Hence, $d_i=0$ for $1\leqq i\leqq n$ and so $(a,x)=(a',x')=(a',1)=i(a')$, as claimed.

To prove Theorem 3.1 suppose that e_1 is a second extension with the same morphisms Φ_{p^k} and Γ. Then e_1 is also equivalent to e' and hence to e. This completes the proof.

For completeness we discuss in more detail the morphisms Φ_{p^k}.

Theorem 3.2. *If p is an odd prime, Φ_{p^k} is a homomorphism $R_{p^k}\to C/p^k C$. If $p=2$, Φ_2 is a quadratic function $R_2\to C/2C$. In fact, for $x,y\in R_2$, $\Phi_2(x+y)-\Phi_2(x)-\Phi_2(y)\equiv\Gamma(y,x)$ mod $2C$.*

Proof. Let $a,b\in B$ and let n be a positive integer. Notice that we can transform $a^n b^n$ into $(a\,b)^n$ by making $\binom{n}{2}$ transpositions $a\,b\to b\,a$. Since

$$(a\,b)(b^{-1}a^{-1}b\,a)=b\,a,$$

and since $b^{-1}a^{-1}b\,a$ is in the center of B, we see that

$$(a^n b^n)(b^{-1}a^{-1}b\,a)^{\binom{n}{2}}=(a\,b)^n.$$

Thus, if $x,y\in R_n$, then

$$\Phi_n(x)+\Phi_n(y)+\Gamma(y,x)^{\binom{n}{2}}\equiv\Phi_n(x+y).$$

If $n=p^k$, p odd prime, then $\binom{p^k}{2}\equiv 0$ mod p^k and so

$$\Phi_n(x+y)=\Phi_n(x)+\Phi_n(y)\quad\text{in}\quad C/p^k C,$$

while if $p=2$, then $\binom{n}{2}=1$, which shows that Φ_2 has the quadratic property given in 3.2. This completes the proof.

Let $\mathrm{Ext}_c(R, C)$ denote the group of central extensions of C by R (R and C abelian). Classically, $\mathrm{Ext}_c(R; C) \approx H^2(R; C)$ [9]. Suppose that $pR=0$, for some prime p. With this hypothesis we present a different computation of Ext_c, more useful for our purposes.

For any abelian group Π denote by $\Gamma_2(\Pi)$ the group defined by Whitehead [12, 3]. There is a canonical homogeneous quadratic function γ: $\Pi \to \Gamma_2(\Pi)$, and given any homogeneous quadratic function ϕ: $\Pi \to A$, there is a unique homomorphism $\hat{\phi}$: $\Gamma_2(\Pi) \to A$ such that $\phi = \hat{\phi} \circ \gamma$.

Now let C, R be abelian groups with $pR=0$, and set $\hat{R} = \Gamma_2 R$ if $p=2$, $\hat{R}=R$ if $p>2$. Define a homomorphism

$$\zeta_p\colon \mathrm{Ext}_c(R, C) \to \mathrm{Hom}(\hat{R}, C/pC) \oplus \mathrm{Hom}(\Lambda^2 R, C)$$

by $e \to (\hat{\Phi}_p, \Gamma)$, where $\hat{\Phi}_p$ is the homomorphism determined by the quadratic map Φ_p, if $p=2$, and $\hat{\Phi}_p = \Phi_p$, if $p>2$.

When $p=2$ define a subgroup $G(R, C)$ of $\mathrm{Hom}(\Gamma_2(R), C/2C) \oplus \mathrm{Hom}(\Lambda^2(R), C)$, by setting

$$G(R, C) = \text{all pairs } (\phi, \psi)$$

such that for all $x, y \in R$,

$$\phi[\gamma(x+y) - \gamma(x) - \gamma(y)] \equiv \psi(x, y) \bmod 2C.$$

Theorem 3.3. *If $p>2$, ζ_p is an isomorphism. If $p=2$, ζ_2 is an isomorphism of* $\mathrm{Ext}_c(R, C)$ *onto* $G(R, C)$.

The proof follows from Theorems 3.1 and 3.2; we omit the details.

Part II

§ 4. Bundles Over Suspensions

As an application of the preceding material we consider the problem of enumerating oriented vector bundles over a suspension. Specifically, let X be a complex and let SX denote the reduced suspension of X. We then have the set $[SX, BSO(n)]$, $n \geqq 2$; with its natural group structure this represents the group of isomorphism classes of oriented n-plane bundles over SX. Equivalently we have the group $[X, SO(n)]$, since $SO(n) \equiv \Omega BSO(n)$; it is this latter group (with dimensional restrictions on X) that we in fact compute. We view this as an extension problem as follows.

For $n \geqq 3$ let $J_n = Z_2$ if n is odd, $J_n = Z$, if n is even. Define

(4.1) $$\theta_n\colon BSO \to K(J_n, n+1)$$

by setting

$$\theta_n = w_{n+1}, \qquad n \text{ odd}$$
$$= \delta w_n, \qquad n \text{ even}.$$

Let $p_n\colon E_n \to BSO$ be the principal fibration with θ_n as classifying map. There is a map $q_n\colon BSO(n) \to E_n$ so that $p_n \circ q_n$ is the inclusion $BSO(n) \subset BSO$ and q_n is an $(r(n)+2)$-equivalence, where $r(n) = n$, if $n \equiv 3 \bmod 4$, $r(n) = n-1$, if $n \not\equiv 3 \bmod 4$. Thus if X is a complex with $\dim X \leqq r(n)$, then $[X, SO(n)] \approx [X, \Omega E_n]$. For this latter group we have the exact sequence

$$\widetilde{KO}^{-2}(X) \xrightarrow{\Omega^2\theta_n} H^{n-1}(X; J_n) \to [X, \Omega E_n] \xrightarrow{\Omega p_n} \widetilde{KO}^{-1}(X) \xrightarrow{\Omega\theta_n} H^n(X; J_n).$$

Setting $'\theta_n = \Omega\theta_n$, $''\theta_n = \Omega^2\theta_n$, we then have the extension (assuming $\dim X \leqq r(n)$),

(4.2) $$e\colon 0 \to \text{Cokernel } ''\theta_{n*} \to [X, SO(n)] \to \text{Kernel } '\theta_{n*} \to 0.$$

In general the group $[X, SO(n)]$ is non-abelian, so to compute e we will need the morphism Γ defined in § 1, as well as the morphisms $\Phi_p k$. We begin with a discussion of the latter.

The functor P

Let p be a fixed prime and k a fixed positive integer. Consider the cofibration

(4.3) $$S \xrightarrow{\gamma} S \xrightarrow{\delta} P \xrightarrow{\rho} S^2 \to \cdots,$$

where $S = S^1$, γ is a map of degree p^k and $P(=P(p,k))$ is the cofiber of γ – i.e., $P(p,k) = P = S^1 \bigcup_{p^k} e^2$.

Suppose we are given C.W.-complexes B and C and a map $\theta\colon B \to C$. One then has the map $\theta^P\colon B^P \to C^P$, which plays a central role in the theory of extensions as given in [6]. We discuss here several properties of the functor P.

Let K_q denote the Eilenberg-MacLane space $K(Z_p, q)$, with universal class ι_q, $q \geqq 1$, [11, p. 425]. Suppose we have mod p cohomology classes $u \in H^r(B)$, $v \in H^s(B)$, $u \cdot v \in H^{r+s}(B)$, $r, s \geqq 1$. We regard $u \cdot v$ as a map

$$u \cdot v\colon B \to K_t, \qquad t = r+s,$$

and our problem is: compute $(u \cdot v)^P\colon B^P \to K_t^P$. As shown in [6], $K_t^P \equiv K_{t-2} \times K_{t-1}$, and so $(u \cdot v)^P$ is a (non-homogeneous) mod p class with two components, $(u \cdot v)_{t-2}^P$ and $(u \cdot v)_{t-1}^P$.

The map δ induces a map (which we denote by the same symbol), $\delta\colon B^P \to B^S$. Now $B^S = B^{S^1} = \Omega B$; let $\sigma\colon H^*(B) \to H^*(B^S)$ denote the loop suspension (of degree -1). Our result is:

Theorem 4.4. *Let $u \in H^r(B)$, $v \in H^s(B)$ be* mod p *cohomology classes, and set $t = r+s$. Then*

$$
\begin{aligned}
&(u \cdot v)^P_{t-1} = 0, \\
&(u \cdot v)^P_{t-2} = 0, \quad \text{if } p>2 \quad \text{or} \quad k>1, \\
&\qquad\qquad = \delta^*(\sigma u \cdot \sigma v), \quad \text{if } p=2 \quad \text{and} \quad k=1.
\end{aligned}
$$

Proof. For any space X, let $e\colon X^P \wedge P \to X$ denote the canonical evaluation map

$$(u, x) \to u(x).$$

Suppose that $\theta\colon K_r \times K_s \to K_t$ is any map ($t = r+s$). The following diagram then commutes, where $T(u, v, x) = ((u, x), (v, x))$.

(4.5)
$$
\begin{tikzcd}
 & (K_r^P \times K_s^P) \wedge P \arrow[dl, "T"'] \arrow[dr, "\theta^P \wedge 1"] & \\
(K_r^P \wedge P) \times (K_s^P \wedge P) \arrow[d, "e \times e"] & & K_t^P \wedge P \arrow[d, "e"] \\
K_r \times K_s \arrow[rr, "\theta"] & & K_t
\end{tikzcd}
$$

We have shown in [6] that for $n \geqq 3$, $K_n^P \equiv K_{n-2} \times K_{n-1}$, with the evaluation map e given by

$$e^* \iota_n = \iota_{n-2} \otimes 1 \otimes a_2 + 1 \otimes \iota_{n-1} \otimes a_1.$$

Here $a_i \in H^i(P)$, $i = 1, 2$, is the mod p generator defined by

$$\delta^* a_1 = s_1, \qquad \rho^* s_2 = a_2,$$

where $s_i \in H^i(S^i)$, $i = 1, 2$, is the canonical mod p generator, and δ, ρ are given in (4.3).

For simplicity, let us define Case 1 to mean either $p>2$ or $k>1$, and Case 2 to mean $p=2$ and $k=1$. Then $a_1 \cdot a_1 = 0$ in Case 1, while $a_1 \cdot a_1 = a_2$ in Case 2 (see [2, p. 252]). Thus, given $(u \otimes a_1 \otimes v \otimes a_1)$ in $H^*((K_r^P \wedge P) \times (K_s^P \wedge P))$, we have

$$
\begin{aligned}
T^*(u \otimes a_1 \otimes v \otimes a_1) &= 0, \ \text{in Case 1} \\
&= u \otimes v \otimes a_2, \ \text{in Case 2}.
\end{aligned}
$$

Moreover, trivially, $a_i \cdot a_j = 0$, if $i+j>2$. Now take θ, in (4.5), to be given by

$$\theta^* \iota_t = \iota_r \otimes \iota_s.$$

Then,

$$T^*(e \times e)^* \theta^* \iota_t = 0, \text{ in Case 1}$$
$$= 1 \otimes \iota_{r-1} \otimes 1 \otimes \iota_{s-1} \otimes a_2, \text{ in Case 2}.$$

On the other hand,

$$(\theta^P \wedge 1)^* e^* \iota_t = (\theta^P \wedge 1)^* [\iota_{t-2} \otimes 1 \otimes a_2 + 1 \otimes \iota_{t-1} \otimes a_1)]$$
$$= \theta^{P*}(\iota_{t-2} \otimes 1) \otimes a_2 + \theta^{P*}(1 \otimes \iota_{t-1}) \otimes a_1,$$

and so by commutativity of (4.5),

$$\theta^{P*}(1 \otimes \iota_{t-1}) = 0,$$
$$\theta^{P*}(\iota_{t-2} \otimes 1) = 0, \text{ in Case 1},$$
$$= 1 \otimes \iota_{r-1} \otimes 1 \otimes \iota_{s-1}, \text{ in Case 2}.$$

We have the map

$$\delta \colon (K_r \times K_s)^P \to (K_r \times K_s)^S = K_r^S \times K_s^S,$$

and $\delta^*(\sigma\, \iota_r \otimes \sigma\, \iota_s) = 1 \otimes \iota_{r-1} \otimes 1 \otimes \iota_{s-1}$. Thus, Theorem 4.4 is true in the universal example and hence in general.

We note one other fact regarding $P(p,k)$ for arbitrary p. Let

$$u \in H^s(X; Z_q),$$

where q is a prime with $q \neq p$, and $s \geqq 3$.

Lemma 4.6. *Let* $P = P(p,k)$. *Thus,*

$$u^P = (0,0) \in H^{s-2}(X^P; Z_p) \oplus H^{s-1}(X^P; Z_p).$$

This is immediate, since $K(Z_q, s)^P \equiv *$.

We now take $p=2$ and state some results from [6]. Let X be a space and $u \colon X \to K_n$ a mod 2 cohomology class of degree $n \geqq 3$. Then u^P has two components; but $u_{n-1}^P = \delta^* \sigma(u)$, and so if we set $\hat{u} = u_{n-2}^P$ we have

$$u^P = (\hat{u}, \delta^* \sigma(u)).$$

Let Sq^t denote the Steenrod square of degree t, $t \geqq 1$. For each $q \geqq 1$ we think of Sq^t as a map

$$Sq^t \colon K_q \to K_{q+t}.$$

Since $(Sq^t(u))^P = (Sq^t)^P \circ (u^P)$, we then have

$$(Sq^t(u))^P = (Sq^t \hat{u} + \lambda_k Sq^{t-1} \delta^* \sigma u, Sq^t \delta^* \sigma u), \tag{4.7}$$
$$\text{where } \lambda_k = 0 \text{ if } k > 1, \lambda_1 = 1.$$

On the other hand, let δ_l denote the Bockstein coboundary associated to the exact sequence
$$Z \xrightarrow{2^l} Z \xrightarrow{\rho_l} Z_{2^l}, \qquad l \geqq 1.$$
Given a mod 2 class u in $H^n(X)$, we have shown in [6], that $(\delta_1 u)^P$ is a class in $H^{n-1}(X^P; Z_2 k)$, given by:

(4.8) $$(\delta_1 u)^P = \rho_k \delta_1(\hat{u}) + s_k \delta^*(\sigma u),$$

where s_k is the cohomology operation induced by the homomorphism $Z_2 \to Z_{2^k}$ given by
$$(a \bmod 2) \to 2^{k-1} a \bmod 2^k.$$

§ 5. The Space B^P

We now take the prime p to be 2 (cf. Lemma 4.6), and set
$$P = P(2, k) = S^1 \bigcup_{2^k} e^2, \qquad k \text{ fixed}.$$
In order to compute the operation Φ_{2^k}, we need to study the space BSO^P. Note that we have a principal fibration sequence
$$(BSO)^{S^2} \xrightarrow{\rho} BSO^P \xrightarrow{\delta} BSO^S \xrightarrow{\gamma} BSO^S,$$
where $\gamma(=\gamma_k)$ is induced by a map of degree 2^k on S. Let X be any (pointed) space, and set $(X^P)_0$ = component of the constant map. Set
$$B^P = (BSO)^P_0.$$

Claim: There is a fibration sequence

(5.1) $$\Omega\,\mathrm{Spin} \xrightarrow{\rho'} B^P \xrightarrow{\delta'} SO \xrightarrow{\gamma'_k} \mathrm{Spin}.$$

Recall that $BSO^S \equiv SO$; γ'_k is chosen so that $\pi \circ \gamma'_k = \gamma_k$, where π: Spin $\to SO$ is the covering map.

Since $\pi_1(SO) = Z_2$, it is clear γ_k does lift to such a map γ'_k. We leave it to the reader to check that the fiber of γ'_k is simply $(BSO^P)_0 = B^P$, with $\delta' = \delta \circ j$, where j: $B^P \subset (BSO)^P$ denotes the inclusion.

By an abuse of notation, if f: $BSO \to Y$, we continue to denote by f^P the composite $B^P \xrightarrow{j} (BSO)^P \xrightarrow{f^P} Y^P$.

Our immediate goal is to compute $H^*(B^P)$, mod 2 coefficients. For each $l \geqq 1$ define classes $g_{4l-2} \in H^{4l-2}(B^P)$ by

(5.2) $$(w^P_{4l})_{4l-2} = g_{4l-2},$$

where the Stiefel-Whitney class $w_j \in H^j(BSO)$, $j \geqq 2$, is regarded as a map $BSO \to K_j$. For $i \geqq 1$ define

(5.3) $$h_{2i-1} = \delta^* \sigma w_{2i} \in H^{2i-1}(B^P).$$

Theorem 5.4. $H^*(B^P) = \bigotimes_{l \geqq 1} Z_2[g_{4l-2}] \bigotimes_{i \geqq 1} Z_2[h_{2i-1}]$.

Proof. By (5.1) we have a fibration

$$\Omega\,\mathrm{Spin} \xrightarrow{\rho'} B^P \xrightarrow{\delta'} SO.$$

Set $f_{4l-2} = \rho'^* g_{4l-2} \in H^{4l-2}(\Omega\,\mathrm{Spin})$. Then, $f_{4l-2} = \sigma\pi^*\sigma w_{4l}$, and it follows from Cartan [1] that $H^*(\Omega\,\mathrm{Spin}) = \bigotimes_{l \geqq 1} Z_2[f_{4l-2}]$. Also, $H^*(SO) = \bigotimes_{i \geqq 1} Z_2[\sigma w_{2i}]$. Thus a simple spectral sequence argument shows that

$$H^*(B^P) = \bigotimes_{l \geqq 1} Z_2[g_{4l-2}] \bigotimes_{i \geqq 1} Z_2[h_{2i-1}],$$

as claimed.

We need also to know the Steenrod squares in $H^*(B^P)$. By Wu, if we set $h_{2i} = (h_i)^2$, $i \geqq 1$, then

$$Sq^i h_j = \binom{j}{i} h_{i+j}. \tag{5.5}$$

Before computing the other squares, it is convenient to introduce one other set of characteristic classes in $H^*(B^P)$. Namely, for $l \geqq 1$, define

$$B_{4l} = (w^P_{4l+2})_{4l} \in H^{4l}(B^P). \tag{5.6}$$

Of course by Theorem 5.4 we know that the B's can be expressed in terms of the g's and h's.

Theorem 5.7. *For* $l \geqq 1$,

$$B_{8l-4} = (g_{4l-2})^2 + \lambda_k \left[(h_{2l-1})^4 + \sum_{i=1}^{2l-1} h_{2i-1} \cdot h_{8l-2i-3}\right].$$

$$B_{8l} = (B_{4l})^2 + \lambda_k \sum_{i=1}^{2l} h_{2i-1} \cdot h_{8l-2i+1}.$$

Here $\lambda_k = 0$ *if* $k > 1$, $\lambda_1 = 1$.

Proof. For the first equation we use the fact ([14]) that

$$w_{8l-2} = Sq^{4l-2} w_{4l} + \sum_{i=1}^{2l-1} w_{2i} \cdot w_{8l-2-2i}.$$

Thus, by Theorem 4.4,

$$B_{8l-4} = (w_{8l-2})^P_{8l-4} = (Sq^{4l-2})^P (w^P_{4l})_{8l-4} + \sum_{i=1}^{2l-1} \lambda_k (h_{2i-1} \cdot h_{8l-3-2i}).$$

But by (4.7),

$$\begin{aligned}(Sq^{4l-2})^P (w^P_{4l}) &= (Sq^{4l-2})^P (g_{4l-2}, h_{4l-1}) \\ &= (Sq^{4l-2} g_{4l-2} + \lambda_k Sq^{4l-3} h_{4l-1}, Sq^{4l-2} h_{4l-1}).\end{aligned}$$

Since $Sq^{4l-2}g_{4l-2}=(g_{4l-2})^2$ and by (5.5), $Sq^{4l-3}h_{4l-1}=h_{8l-4}=(h_{2l-1})^4$, the result now follows. The proof for the second equation is entirely similar.

We proceed to compute Sq^i in $H^*(B^P)$. For simplicity we do only the case $k=1$ – i.e., $P=P(2,1)=S^1\bigcup_2 e^2$. The first result is:

$$\begin{aligned}(5.8)\qquad Sq^{4i}g_{4l-2}&=\binom{l-1}{i}g_{4(i+l)-2}+\binom{l-1}{i-1}h_{4(i+l)-2}\\&\quad+\sum_{j=1}^{i}\binom{l-j-1}{i-j}h_{4j-1}\cdot h_{4(l+i-j)-1}.\end{aligned}$$

Proof. We compute $(Sq^{4i}w_{4l})^P$ in two ways. On the one hand, by Wu [14],

$$Sq^{4i}w_{4l}=\sum_{0}^{i}\binom{l-j-1}{i-j}w_{4j}\cdot w_{4(l+i-j)},$$

and so by (5.2) and Theorem 4.4,

$$\begin{aligned}(5.9)\qquad &(Sq^{4i}w_{4l})^P_{4(i+l)-2}\\&=\binom{l-1}{i}g_{4(i+l)-2}+\sum_{j=1}^{i}\binom{l-j-1}{i-j}h_{4j-1}\cdot h_{4(l+i-j)-1}.\end{aligned}$$

On the other hand, by (4.7),

$$\begin{aligned}(Sq^{4i}w_{4l})^P&=(Sq^{4i})^P(w_{4l})^P=(Sq^{4i})^P(g_{4l-2},h_{4l-1})\\&=(Sq^{4i}g_{4l-2}+Sq^{4i-1}h_{4l-1},Sq^{4i}h_{4l-1}),\end{aligned}$$

and so,

$$(5.10)\qquad (Sq^{4i}w_{4l})^P_{4l-2}=Sq^{4i}g_{4l-2}+\binom{l-1}{i-1}h_{4(i+l)-2}.$$

Comparing (5.9) and (5.10) we obtain (5.8).

Next we compute Sq^1 and Sq^2 in $H^*(B^2)$ – for by the Adem relations we then can compute all Sq^t.

$$\begin{aligned}(5.11)\qquad Sq^2g_{4l-2}&=B_{4l}+h_{4l}+h_1\cdot h_{4l-1}.\\Sq^2B_{4l}&=h_{4l+2}+h_1\cdot h_{4l+1}.\end{aligned}$$

These follow from the fact that

$$Sq^2w_{2n}=(n-1)w_{2n+2}+w_2w_{2n}.$$

We are left with determining Sq^1 in $H^*(B^P)$.

For $l\geqq 1$,

$$\begin{aligned}(5.12)\qquad Sq^1g_{4l-2}&=\sum_{i=2}^{2l}h_{i-1}\cdot h_{4l-i}+h_{4l-1}.\\Sq^1B_{4l}&=\sum_{i=2}^{2l+1}h_{i-1}\cdot h_{4l+2-i}.\end{aligned}$$

The proof is similar to those given above, and so is omitted.

Remark 1. Since $f_n = \rho'^* g_n$, $n \equiv 2 \bmod 4$, the above formulae enable one to compute Sq^i in $H^*(\Omega\,\mathrm{Spin})$. In particular, let l be a positive integer, and write $l = 2^{r-1}(2s-1)$, where $r, s \geqq 1$. Then, by Theorem 5.7 and (5.11),

$$Sq^2 f_{4l-2} = (f_{4s-2})^{2^r}. \tag{5.13}$$

Remark 2. By Theorem 5.4, $H^*(SO)$ is embedded in $H^*(B^P)$ as a subalgebra, using δ^*. Thus, with no confusion we may write

$$h_i = \sigma\, w_{i+1} \in H^i(SO).$$

Given an element $\eta \in \widetilde{KO}^{-1}(X)$, we set

$$h_i(\eta) = \eta^* h_i \in H^i(X). \tag{5.14}$$

§ 6. The Calculation of $[X, SO(n)]$

We return to the problem discussed in §4 — the calculation of $[X, SO(n)]$. As before we regard this as an extension problem:

$$e\colon\ 0 \to \mathrm{Cokernel}\ {}''\theta_{n*} \to [X, SO(n)] \to \mathrm{Kernel}\ {}'\theta_{n*} \to 0,$$

for complexes X of $\dim \leqq r(n)$, where $r(n) = n$ if $n \equiv 3 \bmod 4$, $r(n) = n-1$, if $n \not\equiv 3 \bmod 4$. Here

$${}'\theta_{n*}\colon\ \widetilde{KO}^{-1}(X) \to H^n(X; J_n),$$

$${}''\theta_{n*}\colon\ \widetilde{KO}^{-2}(X) \to H^{n-1}(X; J_n),$$

with $J_n = Z$, if n even, $J_n = Z_2$, n odd.

Extension e is determined by the morphisms Γ and $\Phi_p k$ as described in §3. We proceed to compute these.

Theorem 6.1. *Let $\eta, \zeta \in \widetilde{KO}^{-1}(X)$. Then,*

$$\Gamma(\eta, \zeta) = \sum_{i=1}^{n-2} h_i(\eta)\, h_{n-1-i}(\zeta), \qquad n \text{ odd}$$

$$\Gamma(\eta, \zeta) = \sum_{i=1}^{n-3} \delta\big(h_i(\eta)\, h_{n-2-i}(\zeta)\big), \qquad n \text{ even}.$$

Proof. We compute Γ using Theorem 2.1. Let $B = BSO$ and consider $W_q\colon B \to K_q$, $q \geqq 2$. By the Whitney formula, if $m\colon B \times B \to B$ denotes the multiplication, then

$$m^* w_q = \sum_0^q w_i \otimes w_{q-i},$$

and hence, setting $\theta=w_q$,

$$\Delta_\theta=\sum_{2}^{q-2} w_i \otimes w_{q-i}.$$

Applying (twice) (5.1) and (5.2) of [5], we have

$$''\Delta_\theta=\sum_{1}^{q-3} h_i \otimes h_{q-2-i},$$

since $\sigma w_i=h_{i-1}$, $\sigma w_1=\sigma w_0=0$. This proves Theorem 6.1 for n odd by taking $\theta=w_{n+1}$. Similarly, Theorem 6.1 follows for n even, since then $\theta_n=\delta w_n$.

We consider now the evaluation of the morphism Φ_{p^k}. Since θ_n is of order two, it follows from Lemma 4.6 that $\Phi_{p^k}\equiv 0$ if p is odd, and so we consider only $p=2$. Also, when n is odd, then $J_n=Z_2$, and so in this case we need only the morphism Φ_2.

Theorem 6.2. *Let n (in extension e) be an odd integer ($\geqq 3$) and let X be a complex of dimension $\leqq r(n)$. Let $\eta\in\operatorname{Kernel}{}'\theta^n_*$ be a class such that $2\eta=0$, and let $\xi\in[X, B^P]$, $P=P(2,1)$, be a class such that $\delta^*\xi=\eta$. Then,*

$$\Phi_2(\eta)=\begin{cases} g_{n-1}(\xi), & \text{if } n\equiv 3 \bmod 4\\ Sq^2 g_{n-3}(\xi)+h_{n-1}(\eta)+h_1(\eta)\,h_{n-2}(\eta), & \text{if } n\equiv 1 \bmod 4.\end{cases}$$

The proof follows at once from Theorem 2.5 of [6], using (5.2), (5.6) and (5.11) above.

Example (1). Let M be a n-manifold, n odd, and suppose that M is parallelizable. Then, the principle tangent bundle of M is trivial and so the homotopy classes of oriented n-frames on M are given by $[M, SO(n)]$. We define this group to be $\mathrm{Vec}^n(M)$ (cf. [8]). By a simple extension of the Wu formulae, one has:

Suppose that M is a parallelizable n-manifold, and let q be an even integer. Then

$$Sq^q H^{n-q-1}(M; Z_2)=0.$$

Now take $n\equiv 3 \bmod 4$, n not a power of two. Write $n+1=4q=2^r(2m+1)$, where $m\geqq 1$, $r\geqq 2$. Set $i=2^{r-2}$. We then have:

Let M be a parallelizable n-manifold, $n\equiv 3 \bmod 4$, $n+1$ not a power of two. Then the group $\mathrm{Vec}^n(M)$ is given by an extension

$$0\to H^{n-1}(M; Z_2)\to \mathrm{Vec}^n(M)\to \widetilde{KO}^{-1}(M)\to 0,$$

where, if $\eta\in KO^{-1}(M)$ with $2\eta=0$, then

$$\Phi_2(\eta)=\sum_{j=1}^{i}\binom{q-i-j-1}{i-j} h_{4j-1}(\eta)\cdot h_{n-4j}(\eta).$$

The proof follows from Theorem 5.8 – we omit the details.

Example (2). Suppose that $n \equiv 1 \bmod 4$. Then $r(n) = n-1$, and so Kernel $'\theta_{n*} = \widetilde{KO}^{-1}(M)$. Also, if M is an $(n-1)$-dimensional spin manifold, with $n \equiv 1 \bmod 4$, then $Sq^2 H^{n-3}(M; Z_2) = 0$, and so for $\zeta \in \widetilde{KO}^{-2}(M)$, $\eta \in \widetilde{KO}^{-1}(M)$, and $\xi \in [M, B^P]$, we have

$$Sq^2 f_{n-3}(\zeta) = 0, \qquad Sq^2 g_{n-3}(\xi) = 0,$$

$h_{n-1}(\eta) = Sq^2 h_{n-3}(\eta) = 0$. Thus extension e becomes

$$0 \to Z_2 \to [M, SO(n)] \to \widetilde{KO}^{-1}(M) \to 0,$$

with Γ given by Theorem 6.1 and with

$$\Phi_2(\eta) = [h_1(\eta) \cdot h_{n-2}(\eta)] \cdot [M]_2.$$

We now suppose that n is an even integer.

Theorem 6.3. *Let n (in extension e) be an even integer ($\geqq 4$), and let X be a complex of dimension $\leqq n-1$. Suppose that $\eta \in \widetilde{KO}^{-1}(X)$ is a class such that $2^k \eta = 0$, for some $k \geqq 1$. Choose $\xi \in [X, B^P]$, $P = P(2, k)$, such that $\delta_*(\xi) = \eta$. Finally, choose $u \in H^{n-1}(X; Z)$ such that $\rho(u) = h_{n-1}(\eta)$. Then:*

(a) *If $n \equiv 0 \bmod 4$,*

$$\Phi_{2^k}(\eta) = \delta_2 g_{n-2}(\xi) + 2^{k-1}(u).$$

(b) *If $n \equiv 2 \bmod 4$,*

$$\Phi_{2^k}(\eta) = \delta_2 B_{n-2}(\xi) + 2^{k-1}(u).$$

The proof follows from Theorem 2.5 and 3.6 of [6], using (5.2) and (5.6) above.

Corollary 6.4. *Let M be an $(n-1)$-dimensional oriented manifold, with n even, $\geqq 6$. Then $[M, SO(n)]$ is an* abelian *group given by an extension*

$$e\colon\ 0 \to Z \to [M, SO(n)] \to \widetilde{KO}^{-1}(M) \to 0,$$

determined as follows.

(i) *$\Phi_{p^k} \equiv 0$, for p odd, $k \geqq 1$.*

(ii) *If $\eta \in \widetilde{KO}^{-1}(M)$ with $2^k \eta = 0$, $k \geqq 1$, then*

$$\Phi_{2^k}(\eta) = 0, \qquad \text{if } h_{n-1}(\eta) = 0,$$
$$\Phi_{2^k}(\eta) = 2^{k-1} \in Z/2^k Z, \qquad \text{if } h_{n-1}(\eta) \neq 0.$$

Example (3). $[P^{2m-1}, SO(2m)]$ is a free abelian group, of rank 1 if m is odd, and rank 2 if m is even.

Remark. For an interesting interpretation of the groups $[X, SO(n)]$, note §§ 5, 6 of Chapter XIV of H. Bass: Algebraic K-theory, New York: Benjamin Press 1968.

§ 7. Maps into $SO(3)$

Extension e, in § 6, is given in terms of singular cohomology theory and two general cohomology theories – $\widetilde{KO}^*$ and $\widetilde{KO}^*$ with coefficients: for if $P = P(2, k)$, then

$$[X, B^P] = \widetilde{KO}^{-2}(X; Z_{2^k}).$$

We show in this section that for $n = 3$ (n given in extension e), these two general cohomology theories can be replaced by singular cohomology. For this we use a different resolution of $SO(3)$ from that given in § 4.

Recall that $\Pi_i(BSO(3)) = 0$ for $i = 1, 3$ with $\Pi_2(BSO(3)) = Z_2$, $\Pi_4(BSO(3)) = Z$. Moreover, the first k invariant is

$$\delta_2 \mathfrak{P}_2;\ K_2 \to K_5^*.$$

Here $K_i = K(Z_2, i)$, $K_i^* = K(Z, i)$, $\mathfrak{P}_2$ is the Pontrjagin square [13, 15] and δ_2 is the coboundary given in § 4. Denote by $\pi: E \to K_2$ the principal fibration with $\delta_2 \mathfrak{P}_2$ as classifying map. Then if X is a complex with $\dim X \leqq 3$, $[X, SO(3)] \approx [X, \Omega E]$, and to compute this group we have the exact sequence

$$e:\ 0 \to H^3(X; Z) \to [X, SO(3)] \to H^1(X; Z_2) \to 0.$$

We now use the theory of §§ 1–3 to compute the extension; i.e., we need to calculate θ^P and $''\Delta_\theta$, where $\theta = \delta_2 \mathfrak{P}_2$ and $P = P(2, 1)$. Set $\phi = \mathfrak{P}_2$, so that $\theta = \delta_2 \phi$. We first calculate ϕ^P and $''\Delta_\phi$.

Given classes u and v in $H^2(X; Z_2)$, $\mathfrak{P}_2(u+v) = \mathfrak{P}_2(u) + \mathfrak{P}_2(v) + s_2(u \cdot v)$, where s_2: $Z_2 \to Z_4$ is the inclusion [13]. Therefore, given classes $x, y \in H^1(X; Z_2)$,

$$''\Delta_\phi(x, y) = s_2(x \cdot y).$$

Since δ_2 is a homomorphism, this means that

$$''\Delta_\theta(x, y) = \delta_2 s_2(x \cdot y) = \delta(x \cdot y) \in H^3(X; Z). \tag{7.1}$$

To calculate θ^P we use the theory developed in [6] – since θ is not stable, we need the mild extension of the theory given in the Remark in § 3 of [6].

For any space X let

$$\varepsilon:\ X^P \wedge P \to X$$

denote the evaluation map. We then have the commutative square

$$\begin{array}{ccc} K_2^P \wedge P & \xrightarrow{\theta^P \wedge 1} & K_5^{*P} \wedge P \\ \downarrow{\scriptstyle \varepsilon} & & \downarrow{\scriptstyle \varepsilon} \\ K_2 & \xrightarrow{\theta} & K_5^*. \end{array}$$

Now by § 3 of [6],

$$K_2^P \equiv K_0 \times K_1, \qquad K_5^{*P} = K_3.$$

Thus to compute θ^P we can work with mod 2 coefficients. Given $x \in H^2(X; Z_2)$,

$$\delta_2 \mathfrak{P}_2(x) \bmod 2 = Sq^2 Sq^1(x) + x \cdot Sq^1(x)$$

(see [15]). Working mod 2, and using (3.3) of [6], we find:

$$\theta^P(\iota_1) = \iota_1^3.$$

Therefore, by (2.5) of [6] we have: given $x \in H^1(X; Z_2)$,

(7.2) $$\Phi_2(x) = \rho^{-1}(x^3) \in H^3(X; Z)/2H^3(X; Z),$$

where $\rho: H^3(X; Z) \to H^3(X; Z_2)$ is reduction mod 2.

Summing up, we have proved:

Theorem 7.3. *Let X be a complex of dimension $\leqq 3$. We then have the exact sequence*

$$0 \to H^3(X; Z) \to [X, SO(3)] \to H^1(X; Z_2) \to 0,$$

with the extension given as follows:

$$\Gamma(x, y) = \delta(x \cdot y), \qquad x, y \in H^1(X; Z_2),$$
$$\Phi_2(x) = \rho^{-1}(x^3).$$

Example (1). Take $X = M$, an oriented 3-manifold. Then, $H^3(M; Z) = Z$ and so $\Gamma \equiv 0$ – i.e. $[M, SO(3)]$ is abelian. Also ρ induces an isomorphism

$$H^3(M; Z)/2H^3(M; Z) \to H^3(M; Z_2) = Z_2.$$

Thus we can think of Φ_2 as a homomorphism

$$H^1(M; Z_2) \to Z_2, \quad \text{given by} \quad u \to u^3 \cdot [M]_2.$$

(Φ_2 is a homomorphism, since $\Gamma \equiv 0$.) Thus we have:

Theorem 7.4. *Let M be an oriented 3-manifold. Then*

$$\mathrm{Vec}^3(M) = [M, SO(3)] = \mathrm{Kernel}\ \Phi_2 \oplus Z.$$

Example (2). Take $X = S^1 \times P^2$. Let u generate $H^1(S^1; Z_2)$ and v generate $H^1(P^2; Z_2)$ — use the same letters to denote these classes pulled back to $S^1 \times P^2$ by the projections. Notice that $\delta(u\,v)$ generates $H^3(S^1 \times P^2; Z)$, and so extension e becomes:

$$e\colon\ 0 \to Z_2(\delta(u\,v)) \to [S^1 \times P^2, SO(3)] \to Z_2(u) \oplus Z_2(v) \to 0,$$

where we let $Z_2(a)$ denote Z_2 with generator a. Moreover,

$$\Gamma(u, v) = \delta(u\,v)$$
$$\Phi_2(u) = 0 = \Phi_2(v)$$
$$\Phi_2(u+v) = \rho^{-1}(u+v)^3 = \delta(u\,v).$$

Comparing this extension with the extension for the dihedral group D_8 that one obtains using, e.g., page 150 (IV) of [16], one sees that

(7.5) $$[S^1 \times P^2, SO(3)] \approx \text{dihedral group}, D_8.$$

(In [7] we stated incorrectly that the group is the quaternion group.)

Remark (1). It is easily seen that there are 7-equivalences

$$SO \equiv K_1 \times K_3^*, \qquad \mathrm{Spin} \equiv K_3^*,$$

and a 6-equivalence

$$B^P \equiv K_1 \times K(Z_2^k, Z), \qquad P = P(2, k).$$

Thus Theorems 6.2 and 6.3 can be rephrased entirely in terms of cohomology, when $n \leqq 5$. However, for SO and B^P the above equivalences are not H-space equivalences, which complicates the rephrasing of 6.2–6.3. For this reason we chose to do $SO(3)$ by a different approach, which yields a simpler result.

Remark (2). Combining Theorems 6.1 and 6.2 with some K-theory computations, one can extend (7.5) to show that:

$$[S^1 \times P^{4k+2}, SO(4k+3)] = \begin{cases} D_8, & \text{if } k \text{ even} \\ D_8 \oplus Z_2, & \text{if } k \text{ odd}. \end{cases}$$

Bibliography

1. Cartan, H.: Periodicité des groups d'homotopie stable des groups classiques, d'apres Bott. Séminaire Henri Cartan, Ecole N. Sup., 1959/60.
2. Cartan, H., Eilenberg, S.: Homological algebra. Princeton University Press 1956.
3. Eilenberg, S., MacLane, S.: On the groups $H(\pi, n)$, II. Ann. of Math. **70**, 49–139 (1954).
4. Hilton, P.: Homotopy theory and duality. New York: Gordon and Breach 1965.
5. James, I., Thomas, E.: An approach to the enumeration problem. J. Math. Mech. **14**, 485–506 (1965).
6. Larmore, L. L., Thomas, E.: Group extensions and principal fibrations. Math. Scandinav., to appear.

7. Larmore, L. L., Thomas, E.: Group extensions and principal fibrations. Proceedings of the Advanced Study Institute on Alg. Topology, Aarhus, 1970, pp. 588–598.
8. Larmore, L. L., Thomas, E.: Group extensions and twisted cohomology theories. To appear in Illinois J. Math.
9. MacLane, S.: Homology. Berlin: Springer 1963.
10. Milnor, J.: Construction of universal bundles, I. Ann. of Math. **63**, 272–284 (1956).
11. Spanier, E.: Algebraic topology. New York: McGraw-Hill 1966.
12. Whitehead, J. H. C.: A certain exact sequence. Annals of Math. **52**, 51–110 (1950).
13. Whitehead, J. H. C.: On simply-connected, 4-dimensional polyhedron. Commentarii Math. Helvet. **22**, 48–92 (1949).
14. Wu, W. T.: Les i-carrés dans une variété grassmannienne. C. R. Acad. Sci. Paris **230**, 918–920 (1950).
15. Yamanoshita, T.: On certain cohomological operations. J. Math. Soc. Japan **8**, 300–344 (1956).
16. Zassenhaus, H.: The theory of groups, 2nd Edition. New York: Chelsea Publishing Company 1958.

Prof. L. L. Larmore
Department of Mathematics
California State College at Dominguez Hills
Dominguez Hills, Calif. 90247
USA

Prof. E. Thomas
Department of Mathematics
University of California
Berkeley, California 94720
USA

(Received November 30, 1971)

Math. Z. 128, 297–304 (1972)

Groups with a Steinberg Character

Forrest A. Richen

Let G be a finite group and p a rational prime. If G is a simple Lie type group of characteristic p then G has the following property. G has an absolutely irreducible character χ whose degree is the order of a p-Sylow subgroup of G, and all other characters are in the principal p-block, $B_0(p)$. (χ is called the Steinberg character. See [4, 5, 9] and [12].) For simple groups G, this property seems to hold only if G has Lie type of characteristic p.

The most general conjecture that can be made from this observation, namely that the above property about the characters of G forces G to be of Lie type at characteristic p, seems too difficult at present. In fact to prove anything in this direction, I have had to make arithemtical assumptions about G as well as assumptions about a subgroup which ultimately corresponds to the normalizer of a torus. In Section 4 I will indicate why such assumptions may be inevitable.

Theorem. *Let G be a finite group and p a prime such that $|G|=pg'$, $(p,g')=1$. Suppose that G has an irreducible character χ of degree p and that all other characters of G are in the principal p-block, $B_0(p)$. Let P be a p-Sylow subgroup and let $N(P)=HP$ where $H\cap P=\{1\}$. Suppose that H is a $T1$ set and $|N(H)|=2|H|$. Then G is isomorphic to $PSL_2(p)$.*

(Recall a subgroup A of a group G is a TI set if $A\cap A^g\neq\{1\}$ implies $g\in N(A)$.)

The following corollary is related to a theorem of Ito [11].

Corollary. *Let G be a transitive permutation group of prime degree, p. Suppose G has a unique character χ whose degree is divisible by p and that $\chi(1)=p$. Suppose that if the normalizer of a p-Sylow subgroup P of G is HP where $H\cap P=\{1\}$ then H is a $T1$ set and $|N(H)|=2|H|$. Then $p=5$, 7 or 11 and $G\cong PSL_2(p)$.*

The proof of the theorem is carried out in Sections 1–3 and the corollary is proved and discussed in Section 4. I assume familiarity with elementary finite group theory and character theory including the theory of blocks. (Gorenstein's book [8], Huppert's book [10] and Curtis and Reiner's book [3] will serve as general references.) At a crucial point detailed information about B_0 (p) is needed for groups containing p to the first power (Brauer [1]), but that information is fully stated in the text.

§ 1. Some Subgroups of G

Let G be a group that satisfies the hypotheses of the theorem. Let P be a p-Sylow subgroup, $B=N(P)=HP$ where $H\cap P=1$, and $N=N(H)$. Since the characters of G are $\{\chi\}\cup B_0(p)$, G has exactly one block of defect 1. Brauer's first main theorem gives $C(P)\subseteq P$ (or see Theorem 3 in Brauer [2]). Thus $H\cong N(P)/C(P)$ is a cyclic group whose order e divides $p-1$. Since $|N|=2e$, $e=1$ would give $|G|=2$ and $p=2$ contrary to the assumption that G has a character of degree 2. Thus $e>1$. This together with Burnside's fusion theorem (7.1.1, [8]) gives the following fact.

(1.1) *B is a Frobenius group of order pe with cyclic complement H of order $e>1$. G has $t=(p-1)/e$ classes of p-elements.*

(1.2) *G is a non abelian simple group.*

Proof. Suppose $1\neq S\vartriangle G$ and $p\,\|\,|G:S|$. Then $S\leqq O_{p'}(G)$ which is just $\bigcap\{\ker\zeta\colon\ \zeta\in B_0(p)\}$. (Theorem 1, Brauer [2].) Thus the characters of $G/O_{p'}(G)$ are just the characters of $B_0(p)$. Then summing the squares of the degrees of the characters of G gives

$$|G|=\sum_{\zeta\in B_0(p)}\zeta(1)^2+p^2=|G\colon\ O_{p'}(G)|+p^2.$$

The only solution to this equation is $|G|=(p+1)\,p$. But $e\,|\,(p-1)$ and $e\,|\,(p+1)\,p$ forces $e=1$ against (1.1).

On the other hand suppose $p\,\|\,|S|$ and $S\vartriangle G$. Then $P\leqq S$, and the Frattini Lemma (1.3.7, [8]) gives $G=SB$. Hence $G/S\cong B/S\cap B$ is cyclic of order prime to p. Thus $G'\subseteq S$ and we may assume $p\,\|\,|G'|$. Now $\chi\varphi$ is an irreducible character of degree p if φ is a linear character of G. Since χ is the unique irreducible character of degree p, $\chi\varphi=\chi$. Thus χ vanishes on $G-G'$. Hence

$$\begin{aligned}(\chi|_{G'},\chi|_{G'})&=\frac{1}{|G'|}\sum_{x\in G'}|\chi(x)|^2\\&=\frac{1}{|G'|}\sum_{x\in G}|\chi(x)|^2\\&=|G\colon\ G'|.\end{aligned}$$

Now Clifford's Theorem (V, 17.3 in [10]) gives

$$\chi|_{G'}=e\sum_{i=1}^{n}\lambda_i$$

where e is some positive integer and the λ_i's are the distinct G-conjugates of some irreducible character λ_1 of G'. Taking degrees shows $p=e\,n\,\lambda_1(1)$.

Taking the inner product of $\chi|_{G'}$ with itself shows $|G: G'|=e^2 n$. Thus $e=n=1$ and $G'=G$. Hence $S=G$ proving (1.2).

To determine the structure of N we need a lemma.

Lemma. *If $x\in G$ and $\chi(x)\neq 0$, and u and v are involutions then $x=u_1 v_1$ for some u_1 and v_1 conjugate to u and v respectively. There is one class of involutions.*

Proof. Suppose not. Then if u and v are any involutions in G,

$$0=\sum \frac{\zeta(u)\,\zeta(v)\,\overline{\zeta(x)}}{\zeta(1)}$$

where the sum is over all irreducible characters ζ of G (p. 315, [8]). Since $\chi(1)=p$ and u and v are involutions $\chi(u)\,\chi(v)\neq 0$. Moreover (1.2) implies that χ is faithful so that $\chi(u)\,\chi(v)\,\overline{\chi(x)}$ is non zero and relatively prime to p. Also since χ is the unique character of G of degree p, χ is integer valued. This gives

$$-\ell\,(\chi(u)\,\chi(v)\,\overline{\chi(x)})=p\left(\ell \sum_{\zeta\in B_0(p)} \frac{\zeta(u)\,\zeta(v)\,\overline{\zeta(x)}}{\zeta(1)}\right)$$

where $\ell=l\,c\,m\,\{\zeta(1)\colon\ \zeta\in B_0(p)\}$ is relatively prime to p. The left side is an integer prime to p and the right side is an algebraic integer divisible by p. This contradiction proves the first sentence.

Since G is a simple group, G contains an element x of prime order q different than 2 and p (Burnside's Theorem 4.3.3, [8]). If q_0 is a prime ideal in the ring of algebraic integers containing q then $\chi(x)\equiv\chi(1)\equiv p \pmod{q_0}$. Thus $\chi(x)\neq 0$. Thus if u and v are involutions in G, they have conjugates u_1 and v_1 respectively such that $u_1 v_1=x$. Hence u_1 and v_1 are conjugate in the dihedral group $\langle u_1, x\rangle$, and so u and v are conjugate.

(1.3) *N is dihedral of order $2e$.*

Proof. Since B is a Frobenius group B has t non-linear characters $\varphi_1,\dots,\varphi_t$ of degree e and e linear characters $\mu_1,\dots,\mu_e$. Each $\varphi_i=(\xi_i)^B$ for some non-identity character ξ_i of P. (See 4.5.3, [8] for example.) Now χ is fixed under field automorphisms since it is the unique character of G of degree p. Since each ξ_i is algebraically conjugate to ξ_1, we have

$$a=(\chi|_P, \xi_1)_P=(\chi|_P, \xi_i)_P$$

for all i. Moreover since χ is faithful $a\neq 0$. Thus by Frobenius reciprocity $a=(\chi|_B, \varphi_i)_B$ for all i. Since $\chi(1)=p$ and $\varphi_i(1)=e$, $a=1$. Thus

$$\chi|_B-\sum_{i=1}^{t}\varphi_i$$

is some linear character μ of B.

Now let x be a generator of H.

$$\chi(x)=\mu(x)+\sum_{i=1}^{t}\varphi_i(x)=\mu(x)\neq 0,$$

since $\varphi_i=\xi_i^B$ implies that φ_i vanishes off of P. Thus the lemma implies that there is an involution $u\in G$ such that $x^u=x^{-1}$. Thus $\langle x, u\rangle$ is a dihedral subgroup of N of order $2e$ and so is equal to N.

§ 2. Some Characters of G

Recall that if A is a subgroup of G which is a TI set then the mapping from the set of generalized characters of $N(A)$ whose support is contained in $A-\{1\}$ to generalized characters of $G, \alpha\mapsto\alpha^G$, the induced character, is an isometry (4.4.6, [8]).

(2.1) *Suppose $e>4$. Then there are two irreducible characters, λ_1, λ_2 of N of degree 2 which are induced from linear characters of H. λ_1 and λ_2 are real valued. There exist irreducible characters $\gamma_1, \gamma_2, \alpha$ of G and integers ε and δ of absolute value 1 so that*

$$1_H^G-\lambda_i^G=1-\varepsilon\gamma_i+\delta\alpha.$$

$\gamma_1(x)=\gamma_2(x)$ unless x is conjugate to an element of $H-\{1\}$.

Proof. The first two sentences are elementary since N is a dihedral group. A proof of the last, which is well known, is included. Frobenius reciprocity gives $(\Gamma_i^G, 1_G)=1$ where $\Gamma_i=1_H^N-\lambda_i$. Applying the isometry to Γ_i gives $(\Gamma_i^G, \Gamma_j^G)_G=2+\delta_{ij}$. Thus

$$\Gamma_i^G=1+\delta\alpha+\varepsilon_i\gamma_i$$

where $\alpha, \gamma_1, \gamma_2$ are distinct irreducible characters and $\delta, \varepsilon_1, \varepsilon_2$ are integers of absolute value 1. Moreover $\varepsilon_1\gamma_1-\varepsilon_2\gamma_2=\Gamma_1^G-\Gamma_2^G$ has degree zero while $\gamma_i(1)>0$. Thus $\varepsilon_1=\varepsilon_2$ proving (2.1) if we set $\varepsilon=-\varepsilon_1$.

We now must recall some facts about the principal block of G, $B_0(p)$. No assumption about e is made. Brauer's paper [1] is the reference. (Alternatively see [13].)

(2.2) *There are $e+t$ characters in $B_0(p)$. The first e of them $\zeta_1, \ldots, \zeta_e$ have the property that ζ_i has constant value a_i on all p-elements and a_i is either 1 or -1. As a result $\zeta_i(1)\equiv a_i \pmod p$. The remaining t make up the so called exceptional class, $\theta_1, \ldots, \theta_t$. If $g\in G$ has order prime to p, then $\theta_i(g)=\theta_j(g)$. $\sum_{i=1}^{t}\theta_i$ has constant value a on the p-elements, $a=1$ or -1, and as a result $t\theta_i(1)\equiv\sum\theta_i(1)\equiv a \pmod p$. Moreover $a\theta_1+\sum_{i=1}^{e}a_i\zeta_i$ vanishes on p-regular elements. $B_0(p)$ gives a tree T in the following way.*

The vertices of T are the characters $\zeta_1, \ldots, \zeta_e$, $\theta = \sum \theta_i$, and two vertices are joined by an edge if and only if they have a modular constituent in common. Two vertices are joined by at most one edge. If two vertices are joined by an edge, they take different values on p-elements.

(2.3) *Suppose $e > 4$. Then $\alpha = \chi$; γ_1 and γ_2 are real valued characters of degree $p+1$.*

Proof. At least one γ_i is not in the exceptional family of $B_0(p)$ for if they both were, then (2.1) and (2.2) would say that they are equal everywhere contrary to $\gamma_1 \neq \gamma_2$. Let $x = \gamma_1(1)$ and $a = \alpha(1)$. Then (2.1) implies $0 = 1 - \varepsilon x + \delta a$, and (2.2) implies that $\varepsilon x \equiv \pm 1 \pmod p$. If $\varepsilon x \equiv 1 \pmod p$, then $p \mid a$ and so $\alpha \notin B_0(p)$. Thus $\alpha = \chi$. If $\varepsilon x \equiv -1 \pmod p$ then $\delta a \equiv -2 \pmod p$. Since $e > 4$ and $e \mid p-1$, $p \geqq 7$. Thus $\delta a \not\equiv \pm 1$ and so α must belong to the exceptional family of $B_0(p)$. But then $t \delta a \equiv \pm 1 \pmod p$ which gives $-2t \equiv \pm 1 \pmod p$. Hence $p \leqq 2t + 1$. But $e > 4$ and $e t = p - 1$ yield $4t < 2t$ a contradiction.

Since $\alpha = \chi$, (2.1) gives that

$$\gamma_i = \varepsilon(1 + \delta \chi - \Gamma_i^G).$$

χ is rational valued since it is the only character of degree p, and (2.1) says that λ_i and hence Γ_i^G are real valued. Thus γ_i is real valued. Its degree $\varepsilon(1 + \delta \chi(1)) = \varepsilon + \varepsilon \delta p$. Since this must be a positive number, $\varepsilon \delta = 1$. Thus $\gamma_i(1)$ is $p-1$ or $p+1$.

Suppose $\gamma_i(1) = p - 1$. Then $\gamma_i(g) = -1$ if g is a p-element, and so $\gamma_{i|P} + 1_P$ is the regular character of P. In the notation of the proof of (1.3) we then have $1 = (\gamma_{i|P}, \xi_j)_P = (\gamma_{i|B}, \varphi_j)_B$ by Frobenius reciprocity, and so $\gamma_{i|B} = \sum_{j=1}^{t} \varphi_j$ which vanishes on $H - \{1\}$. But if x is a generator of H we have $\gamma_i(x) = -1 + \chi(x) - 1_H^G(x) + \lambda_i^G(x) = -3 + \chi(x) + \lambda_i^G(x)$. But $\chi(x)$ is a rational root of unity from the proof of (1.3), and $\lambda_i^G(x)$ is easily computed to be a real number < 2. Thus $0 = \gamma_i(x) < -3 + 1 + 2 = 0$ a contradiction. Thus $\gamma_i(1) = p + 1$ proving (2.3).

For the remainder of this section assume $e > 4$. By a theorem of Tuan (Theorems A and B, [14]) the tree T described in (2.2) can be drawn so that it is symmetric with respect to a stem (an open polygonal subgraph) and so that mapping a character to its contragredient is the automorphism of the tree which is reflection through the stem. Thus a character is real valued if and only if it is a vertex on the stem. The trivial character is one end point of the stem. Both of γ_1 and γ_2 are on the stem and so least one of them, γ_1 say, has more than one modular constitutent. No modular constituent of γ_1 is the trivial modular representation since $\gamma_1(g) = 1$ for a p-element g by (2.2). Thus one of the modular constituents of γ_1 is a faithful irreducible representation of G in a field of characteristic

p of degree no larger than $(p+1)/2$. Feit's theorem (Theorem 1, [6]) implies that G is of type $L_2(p)$ and so (1.2) implies that $G \cong PSL_2(p)$.

§3. $e \leqq 4$

It is possible to reduce the cases $e=2, 3$ and 4 to high powered classification theorems, but with very little extra effort these cases can be disposed of by elementary methods.

By (2.2) there are $e+t$ characters in $B_0(p)$ and there is just one other, the Steinberg character χ. Hence G has exactly $e+t+1$ conjugacy classes. By (1.1) t of these classes are classes of p-elements and there is one class for the identity. Writing $|G|=(1+rp)ep$, where $1+rp=|G\colon N(P)|$, each p-element g is in a class of size $|G\colon C(g)|=[(1+rp)ep]/p=(1+rp)e$. Thus there are $(1+rp)et=(1+rp)(p-1)$ p-elements in G. Let $x_1, \ldots, x_e$ be the orders of the elements in each of the remaining e-classes and $c_1, \ldots, c_e$ the orders of the centralizers of these elements. Summing the orders of the conjugacy classes we get

$$(1+rp)ep=1+(1+rp)(p-1)+(1+rp)ep\left[\frac{1}{c_1}+\cdots+\frac{1}{c_e}\right]$$

or

$$(*) \qquad 1=\frac{1+r(p-1)}{1+rp}\cdot\frac{1}{e}+\frac{1}{c_1}+\cdots+\frac{1}{c_e}.$$

Suppose $e=2$. Then N is the centralizer of an involution and has order 4. Since G is simple, G has a prime q different from 2 and p in its order. The centralizer of a q-element is a power of q, since G has only 3 classes of p-regular elements. Thus we may write $c_1=q^n$ and $c_2=4$. Eq. (*) becomes

$$1=\frac{1+r(p-1)}{1+rp}\cdot\frac{1}{2}+\frac{1}{q^n}+\frac{1}{4}.$$

This forces $q^n<4$ and $q^n=3$ first of all and then forces $r=1$ and $p=5$. Thus $|G|=60$ and so $G \cong PSL_2(5)$. (See [7] for example.)

Suppose $e=3$. Since N has order 6 and is dihedral the three elements are self centralizing and conjugate to their inverses. This says that the 3-Sylow subgroups have order 3 and that there is one class of 3-elements. By the lemma of Section 2 there is one class of involutions. Thus we may write $c_3=x_3=3$, $x_2=2$ and $c_2 \geqq 4$. Eq. (*) yields that

$$1<\frac{1}{3}+\frac{1}{c_1}+\frac{1}{4}+\frac{1}{3}$$

which gives $c_1<12$. By the above $c_1 \neq 2, 3, 6$, or 9. c_2 cannot be 10 because then G would contain elements of order 5 and 10, and there are not enough classes for this. Neither can c_1 be 5, 7 or 11 because if $c_1=x_1=q$

for a prime q, then a q-Sylow subgroup Q has order q, and its normalizer induces the full automorphism group of Q since there is only one class of q-elements. This forces G to contain an element of order $q-1$ and there are not enough classes for such an element. Thus $c_2=2^a$ and $c_1=2^b$, a and $b \geqq 2$. The only solution consistent with $*$ is $c_1=4$ and $c_2=8$. Applying $(*)$ again yields $r=1$ and $p=7$ giving $|G|=168$. Thus $G \cong PSL_2(7)$. (See [7] for example.)

Finally suppose $e=4$ for purposes of obtaining a contradiction. Since H is characteristic in N, and N is dihedral of order 8, N is a 2-Sylow subgroup of G and H is self centralizing. Moreover there is just one class of elements of order 4 by Sylow's theorem. By the lemma of Section 2 there is just one class of involutions. Thus we may wirte $x_2=2$ and $c_2 \geqq 8$, $x_4=c_4=4$ and $c_3 \geqq c_1 \geqq 3$ with neither c_1 nor c_3 being 4. Eq. $(*)$ gives

$$1<\frac{1}{4}+\frac{2}{c_1}+\frac{1}{4}+\frac{1}{8}$$

which says that $c_1=3$ or 5. Hence $x_1=c_1=q$ where q is 3 or 5. Thus a q-Sylow subgroup has order q and $c_2=8$.

This gives just two possibilities for c_3 and x_3. Either $c_3=x_3=q$ or $x_3=s$ a prime different from p, 2 and q and $c_3=s^n$ for some n. If $c_3=q$ then Eq. $(*)$ becomes

$$1=\frac{1}{4}\left(\frac{1+r(p-1)}{1+rp}\right)+\frac{2}{q}+\frac{1}{4}+\frac{1}{8}$$

with q either 3 or 5. $q=3$ gives no positive integer solutions to r and p while $q=5$ implies $r=1$ and $p=9$ a contradiction.

On the other hand if $x_3=s$ then the normalizer of an s-Sylow subgroup is a Frobenius group of order $(s^n-1)s^n$. (This follows from the Burnside fusion theorem and the fact that $c_3=s^n$.) Inspecting the possible Frobenius complements (see 10.3.1, [8] for example) yields that $s^n-1=2$ or 4 and so $c_3=3$ or 5. Thus we may assume $c_1=3$, $c_2=8$, $c_3=5$, and $c_4=4$, and equation $(*)$ has no solution with these values for the c_i. This completes the proof of the theorem.

§ 4. The Corollary

The corollary follows directly from the theorem. Since G is a transitive permutation group of degree p where p is a prime, G is isomorphic to a subgroup of the symmetric group on p letters. Thus p divides $|G|$ only once and the p-Sylow groups of G are self centralizing. But a p-Sylow subgroup being self centralizing implies that every character whose degree is prime to p belongs to $B_0(p)$ (Theorem 3, [2]). Thus the hypotheses of the theorem hold for G and p, and so $G \cong PSL_2(p)$. Since G has a subgroup of index p, a theorem of Galois implies that $p=2, 3, 5, 7$ or 11 (II, 8.28, [10]), and $p=2$ and 3 are excluded since G is simple.

Ito proves the following related result (Theorem, [11]). Suppose G is a transitive permutation group of prime degree p. (A) Suppose $(p-1)/2=q$ is a prime. (B) Suppose exactly one character of G has degree divisible by p. Then $p=5, 7$, or 11 and $G \cong PSL_2(p)$.

Hypothesis (B) is weaker than the corresponding hypothesis about the Steinberg character in the corollary. However hypothesis (A) is certainly stronger than the corollary's hypothesis about H. In fact it is not hard to prove that if G is a simple transitive permutation group of degree $p=2q+1$ with p and q primes and if H is defined as in the corollary, then H is a TI set and $|N(H)|=2|H|$. Thus perhaps some strong assumption on H is inevitable.

References

1. Brauer, R.: On groups whose order contains a prime number to the first power I. Amer. J. Math. **64**, 401–420 (1942).
2. Brauer, R.: Some applications of the theory of blocks of characters of finite groups I. J. Algebra **1**, 152–167 (1964).
3. Curtis, C. W., Reiner, I.: Representation theory of finite groups and associative algebras. New York: Interscience 1962.
4. Curtis, C. W.: The Steinberg character of a finite group with a (B, N)-pair. J. Algebra **4**, 433–441 (1966).
5. Dagger, S. W.: On the blocks of the Chevalley groups. J. London Math. Soc. (2) **3**, 21–29 (1971).
6. Feit, W.: Groups with a cyclic Sylow subgroup. Nagoya Math. J. **27**, 571–584 (1966).
7. Frobenius, G.: Über Gruppen des Grades p oder $p+1$. Sitzber. Preuß. Akad. Wiss., (1902), 351–369.
8. Gorenstein, D.: Finite groups. New York: Harper and Row 1968.
9. Humphreys, J. E.: Defect groups for finite groups of Lie type. Math. Z. **119**, 149–152 (1971).
10. Huppert, B.: Endliche Gruppen I. Berlin-Heidelberg-New York: Springer 1967.
11. Ito, N.: A note on transitive permutation groups of degree $p=2q+1$, p and q being prime numbers. J. Math. Kyoto. Univ. **3-1**, 111–113 (1963).
12. Richen, F.: Blocks of defect 0 of a split (B, N)-pair. J. Algebra **21**, 275–279 (1972).
13. Thompson, J. G.: Vertices and sources. J. Algebra **6**, 1–6 (1967).
14. Tuan, H. F.: On groups whose orders contain a prime to the first power. Ann. of Math. **45**, 110–140 (1944).

Prof. F. A. Richen
Department of Mathematics
University of Michigan
Ann Arbor, Michigan 48104
USA

(Received November 23, 1971)

Math. Z. 128, 305–309 (1972)

On Some Inequalities of Series of Positive Terms

László Leindler

Prékopa [3] proved that

$$\int_{-\infty}^{\infty} \sup_{x+y=t} f(x)\, g(y)\, d t \geqq 2 \left[\int_{-\infty}^{\infty} f^{2}(x)\, d x\right]^{\frac{1}{2}}\left[\int_{-\infty}^{\infty} g^{2}(y)\, d y\right]^{\frac{1}{2}} \tag{1}$$

holds for arbitrary measurable non-negative functions $f(x)$ and $g(y)$.

It is easy to see that the formal analogue of (1) with non-negative sequences, i.e.

$$\sum_{n=-\infty}^{\infty} \sup_{k+l=n} a_{k} b_{l} \geqq 2\left(\sum_{k=-\infty}^{\infty} a_{k}^{2}\right)^{\frac{1}{2}}\left(\sum_{k=-\infty}^{\infty} b_{k}^{2}\right)^{\frac{1}{2}}$$

is not true generally. (See e.g. the sequences $a_0=b_0=1$ and $a_n=b_n=0$ if $n\neq 0$.) However the weaker inequality

$$\sum_{n=-\infty}^{\infty} \sup_{k} a_{k} b_{n-k} \geqq\left[\sum_{n=-\infty}^{\infty} a_{n}^{p}\right]^{1/p}\left[\sum_{n=-\infty}^{\infty} b_{n}^{q}\right]^{1/q}, \tag{2}$$

where $p, q \geqq 1$ and $\frac{1}{p}+\frac{1}{q}=1$, does hold true. Indeed, inequality (2) is a special case of our following result (see Theorem in [1]):

Suppose that $1 \leqq r, s \leqq \infty$ and $\frac{1}{r}+\frac{1}{s}=1+\frac{1}{\gamma}$, where $1 \leqq \gamma \leqq \infty$. Then for non-negative a_n, b_n

$$\sum_{n=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} a_{k}^{\gamma} b_{n-k}^{\gamma}\right)^{1/\gamma} \geqq\left(\sum_{k=-\infty}^{\infty} a_{k}^{r}\right)^{1/r}\left(\sum_{n=-\infty}^{\infty} b_{n}^{s}\right)^{1/s} {}^{1} \tag{3}$$

holds.

In [2] we generalized inequality (1) as follows:

$$\int_{-\infty}^{\infty} \sup_{x+y=t} f(x)\, g(y)\, d t \geqq p^{1/p} q^{1/q}\left[\int_{-\infty}^{\infty} f^{p}(x)\, d x\right]^{1/p}\left[\int_{-\infty}^{\infty} g^{q}(x)\, d x\right]^{1/q}, \tag{4}$$

where $p, q \geqq 1$ and $\frac{1}{p}+\frac{1}{q}=1$.

---

[1] If $\gamma=\infty$ and $c_n \geqq 0$, then $\left(\sum_{n=-\infty}^{\infty} c_n^{\gamma}\right)^{1/\gamma}$ means $\sup_n c_n$.

In the present paper we prove the following assertion which may be regarded as an analogue of (4) for sequences.

Theorem 1. *If* $1 \leqq p,\ q \leqq \infty$ *and* $\frac{1}{p}+\frac{1}{q}=1$, *then for any non-negative* a_n *and* b_n $(n=0, \pm 1, \pm 2, \ldots)$ *we have*

$$(\sup_i a_i)(\sup_i b_i) + \sum_{n=-\infty}^{\infty} \sup_k a_k b_{n-k} \geqq p^{1/p} q^{1/q} \left(\sum_{n=-\infty}^{\infty} a_n^p\right)^{1/p} \left(\sum_{n=-\infty}^{\infty} b_n^q\right)^{1/q}. \tag{5}$$

Theorem 1 can be extended from two to any finite number of sequences by a straightforward generalization of the proof of (5). Thus, if $\{a_n^{(i)}\}$ $(i=1, 2, \ldots, m;\ n=0, \pm 1, \pm 2, \ldots)$ are m sequences of non-negative numbers, then

$$\begin{aligned}(m-1)\prod_{i=1}^{m}(\sup_k a_k^{(i)}) &+ \sum_{n=-\infty}^{\infty} \sup_{k_1+k_2+\cdots+k_m=n} a_{k_1}^{(1)} a_{k_2}^{(2)} \ldots a_{k_m}^{(m)} \\ &\geqq \prod_{i=1}^{m} (p_i)^{1/p_i} \prod_{i=1}^{m}\left(\sum_{k=-\infty}^{\infty} (a_k^{(i)})^{p_i}\right)^{1/p_i},\end{aligned}$$

where $1 \leqq p_i \leqq \infty$ and $\sum_{i=1}^{m} \frac{1}{p_i} = 1$.

A similar generalization of (4) to m functions can also be found in [2].

We shall obtain Theorem 1 as a special case of the following

Theorem 2. *Suppose that* $1 \leqq \gamma \leqq \infty$, $1+\frac{1}{\gamma} \leqq p,\ q \leqq \infty$ *and* $\frac{1}{p}+\frac{1}{q}=1$. *Then we have*

$$\begin{aligned}(\sup_n a_n)(\sup_n b_n) &+ \sum_{n=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} a_k^\gamma b_{n-k}^\gamma\right)^{1/\gamma} \\ &\geqq p^{1/p} q^{1/q} \left(\sum_{n=-\infty}^{\infty} a_n^p\right)^{1/p} \left(\sum_{n=-\infty}^{\infty} b_n^q\right)^{1/q}\end{aligned} \tag{6}$$

for arbitrary non-negative a_n *and* b_n.

It is clear that in the case $\gamma=\infty$ Theorem 2 reduces to Theorem 1. Furthermore, it is also evident that if $\gamma=1$ then (3) states more than (6).

Proof of Theorem 2. We may assume that the left-hand side of (6) has finite value and that not all a_k and not all b_k are zero, furthermore that $p, q > 1$. Indeed in the other cases inequality (6) obvious. Let us then choose m such that both

$$A_m = \max_{|i| \leqq m} a_i \quad \text{and} \quad B_m = \max_{|i| \leqq m} b_i$$

are different from 0, and set

$$\alpha_k = \frac{a_k}{A_m} \quad \text{and} \quad \beta_k = \frac{b_k}{B_m} \qquad (k=0, \pm 1, \pm 2, \ldots).$$

It is clear that we can fix integers v and μ between $-m$ and m such that $\alpha_v = 1$ and $\beta_\mu = 1$. Furthermore, we define

$$\bar{\alpha}_k = \begin{cases} \alpha_k, & \text{if } k \neq v, \\ 2, & \text{if } k = v, \end{cases} \quad \text{and} \quad \bar{\beta}_k = \begin{cases} \beta_k, & \text{if } k \neq \mu, \\ 2, & \text{if } k = \mu. \end{cases}$$

First we show that

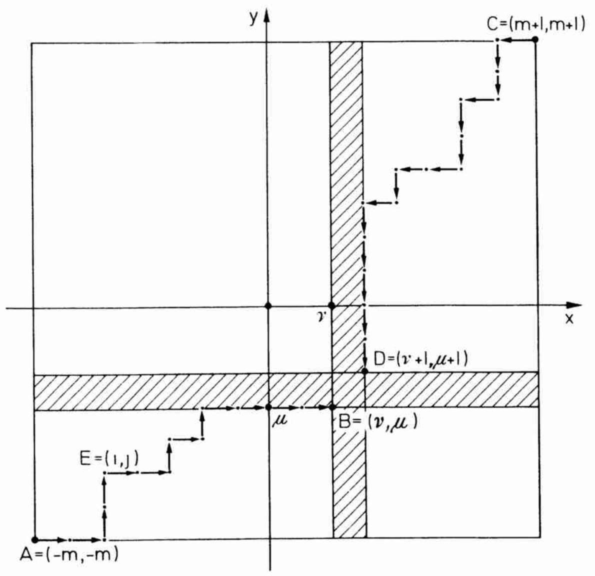
$$1 + \sum_{n=-2m}^{2m} c_n \geqq \sum_{k=-m}^{m} (\alpha_k^p + \beta_k^q) \tag{7}$$

where

$$c_n = \left\{ \sum_{i=-\infty}^{\infty} (\alpha_i \beta_{n-i})^\gamma \right\}^{1/\gamma}.$$

To prove (7) we first construct two directed graphs. The origin of the first graph will be the point $A=(-m, -m)$ and its destination the point $B=(v, \mu)$. If $(-m, -m) \equiv (v, \mu)$ then the first graph is an isolated point. Otherwise we assume that our graph from A has reached the point $E=(i, j)$ $(-m \leqq i \leqq v,\ -m \leqq j \leqq \mu$ and $i+j < v+\mu)$. From E we go one unit to the right or upward according as $\bar{\alpha}_i^p \leqq \bar{\beta}_j^q$ or $\bar{\alpha}_i^p > \bar{\beta}_j^q$.

If we continue this procedure then one of the points of the graph comes up to the point B. This ensues necessarily (see the following figure)



since one of the points E "knocks against the lined wall" and henceforth the points E go along the wall to the point (ν, μ) (in fact, the sequences $\{\bar{\alpha}_k\}$ and $\{\bar{\beta}_k\}$ $(k=-m, -m+1, \ldots, m)$ take their largest values on the wall). For similar reasons and by an analogous method we can construct the second graph coming back from the point C to the point D.

Now we can prove by means of these graphs that

$$\sum_{n=-2m}^{\nu+\mu-1} c_n \geqq \sum_{k=-m}^{\nu-1} \alpha_k^p + \sum_{k=-m}^{\mu-1} \beta_k^q \tag{8}$$

and

$$\sum_{n=\nu+\mu+1}^{2m} c_n \geqq \sum_{k=\nu+1}^{m} \alpha_k^p + \sum_{k=\mu+1}^{m} \beta_k^q. \tag{9}$$

(If $\nu+\mu=-2m$ or $\nu+\mu=2m$, then we omit the first or the second inequality, respectively.)

It is easy to see that there exists a one-to-one correspondence between the points of the graph $\overrightarrow{AB}$ and the terms of the sequence

$$\{c_n\} \quad (n=-2m, -2m+1, \ldots, \nu+\mu-1).$$

The one-to-one mapping suitable for us is

$$(i, j) \leftrightarrow c_{i+j}. \tag{10}$$

Using this correspondence we can prove (8). Indeed, it is evident that

$$c_{i+j} \geqq \alpha_i \beta_j \geqq \min(\alpha_i^p, \beta_j^q). \tag{11}$$

Furthermore our graph is constructed in such a way that if it goes from a point (i, j) to the right then $\bar{\alpha}_i^p \leqq \bar{\beta}_j^q$, and if upward then $\bar{\alpha}_i^p > \bar{\beta}_j^q$. As the graph steps $(\nu+m)$-times to the right and $(\mu+m)$-times upward it follows, hence and from (11), that there exist c_n's with $\nu+m$ different subscripts n such that for these c_n

$$c_n = c_{i+j} \geqq \alpha_i^p \qquad (i=-m, -m+1, \ldots, \nu-1), \tag{12}$$

and for other $(\mu+m)$ c_n

$$c_n = c_{i+j} \geqq \beta_j^q \qquad (j=-m, -m+1, \ldots, \mu-1) \tag{13}$$

hold, furthermore by (10) the same c_n does not appear in the inequalities (12) and (13) simultaneously.

Herewith we have proved (8).

The inequality (9) can be proved similarly, the only change is that then the one-to-one mapping has the following form:

$$(i+1, j+1) \leftrightarrow c_{i+j}.$$

Since $\alpha_\nu=\beta_\mu=1$, thus $c_{\nu+\mu}\geqq 1$, consequently

$$1+c_{\nu+\mu}\geqq\alpha_\nu^p+\beta_\mu^q \tag{14}$$

holds.

Taking into consideration the inequalities (8), (9) and (14) we have proved the inequality (7).

By (7) we have

$$1+\sum_{n=-\infty}^{\infty} c_n\geqq\sum_{k=-m}^{m}(\alpha_k^p+\beta_k^q),$$

whence, using the following well-known inequality

$$x\,y\leqq\frac{x^p}{p}+\frac{y^q}{q},$$

the inequality

$$1+\sum_{n=-\infty}^{\infty} c_n\geqq p^{1/p}q^{1/q}\left(\sum_{k=-m}^{m}\alpha^p\right)^{1/p}\left(\sum_{k=-m}^{m}\beta^q\right)^{1/q} \tag{15}$$

follows. Multiplying both sides of (15) by $A_m\cdot B_m$ we obtain

$$A_m B_m+\sum_{n=-\infty}^{\infty}\left\{\sum_{i=-\infty}^{\infty}(a_i b_{n-i})^\gamma\right\}^{1/\gamma}\geqq p^{1/p}q^{1/q}\left(\sum_{k=-m}^{m}a_k^p\right)^{1/p}\left(\sum_{k=-m}^{m}b^q\right)^{1/q}$$

which obviously implies the required inequality (6).

The proof is completed.

References

1. Leindler, L.: One some inequalities concerning series of positive terms. Acta Sci. Math. **33**, 11-14 (1972).
2. Leindler, L.: On a certain converse of Hölder's inequality. II, Acta Sci. Math. **33** (1972).
3. Prékopa, A.: Logarithmic concave measures with application to stochastic programming. Acta Sci. Math. **32**, 301-316 (1971).

Prof. László Leindler
Bolyai Institute
Attila József University
Szeged, Hungary

(Received June 6, 1972)

Math. Z. 128, 311–323 (1972)

Herleitung von Runge-Kutta-Verfahren der Ordnung vier mit übersichtlichen Fehlerdarstellungen

Rudolf Scherer

1. Einleitung

Bei der Herleitung und Untersuchung von Runge-Kutta-Verfahren benützt man meistens Taylor-Abgleich in Verbindung mit Differentialoperatoren oder Differentialausdrücken (vgl. etwa Kopal [5], Butcher [3]).

Auch bei der Herleitung von Restgliedern, die den Fehler bei Ausführung eines Schrittes angeben, verwendet man ähnliche Hilfsmittel; siehe etwa Bieberbach [2]. Albrecht [1] arbeitete mit der Einschaltung geschickt gewählter Größen und gelangte zu genauen Abschätzungen.

Eine andere Methode wurde von Stetter und Zeller [7] vorgeschlagen: Verwendung des rekursiven Aufbaus der Runge-Kutta-Verfahren im Zusammenhang mit Integrationsformeln. Beim gewöhnlichen Runge-Kutta-Verfahren führte dies zu einer übersichtlichen Fehlerdarstellung. Zeller [8] hat angedeutet, wie man mit dieser Methode Runge-Kutta-Verfahren herleiten kann.

Die vorliegende Arbeit bringt weitere Anwendungen dieser Methode, insbesondere zur allgemeinen Darstellung des Schrittfehlers und zur Herleitung von Runge-Kutta-Formeln mit übersichtlichen Fehlerschranken. Die allgemeine Fehlerdarstellung enthält sieben Glieder, die alle die Ordnung $O(s^5)$ haben und aus Kombinationen verschiedener Integrationsfehlerglieder bestehen. Es treten partielle Ableitungen der rechten Seite der Differentialgleichung bis zur Ordnung drei auf. Die Fehlerschranken werden übersichtlicher, falls die partiellen Ableitungen verschwinden. Wir stellen Bedingungen für die Parameter der Verfahren auf, mit deren Hilfe sich die partiellen Ableitungen dritter Ordnung vermeiden lassen, und geben Runge-Kutta-Formeln an, die diesen Bedingungen genügen.

2. Bezeichnungen

Wir betrachten die gewöhnliche Differentialgleichung (DGL) $y' = f(x, y)$. f sei in einem hinreichend großen konvexen Gebiet G, das die Lösung und alle benötigten Hilfspunkte enthält, definiert und „genügend glatt“ (Eindeutigkeit der Lösung und Existenz der höheren Ableitungen).

Ohne Beschränkung der Allgemeinheit wählen wir den Koordinatenursprung (0, 0) als Anfangspunkt. Die exakte Lösung durch diesen Punkt werde mit dem Funktionssymbol u bezeichnet.

Bei den expliziten Runge-Kutta-Verfahren (RKV) werden nun rekursiv Hilfsgrößen U_n, U'_k gebildet:

$$U_n = s\sum_{k=0}^{n-1} a_{nk}\, U'_k, \qquad U'_k = f(c_k s, U_k) \qquad (n=1,\dots,4) \qquad (c_0=0,\ c_4=1). \tag{2.1}$$

Beginnend mit $U_0=0$ erhält man also nacheinander Näherungen U_k für die entsprechenden Werte $u_k := u(c_k s)$ der wahren Lösung und ebenso Näherungen U'_k für die entsprechenden Werte $u'_k := u'(c_k s)$ der Ableitung (Steigung) der wahren Lösung. Der gesuchte Näherungswert U_4 für die wahre Lösung u_4 (nach einem Schritt mit der Schrittweite s) lautet:

$$U_4 = s\sum_{k=0}^{3} a_{4k}\, U'_k. \tag{2.2}$$

Damit ist der Schrittfehler E gegeben durch

$$E = U_4 - u_4. \tag{2.3}$$

Wir interessieren uns nur für den Schrittfehler (Formelfehler); auf Fortpflanzungs- und Rundefehler gehen wir nicht ein.

Zur Berechnung der U_n (vgl. (2.1)) verwendet man Integrationsformeln, in denen die schon verfügbaren Näherungen U'_k für die entsprechenden Werte u'_k auftreten. Der grundsätzliche Integrationsfehler ist

$$E_n = s\sum_{k=0}^{n-1} a_{nk}\, u'_k - u_n \qquad (n=1,\dots,4). \tag{2.4}$$

Die Abweichung der Näherung U'_k von u'_k führt zu

$$\begin{aligned} U'_k - u'_k = (U_k - u_k)\, p_k \quad \text{mit} \quad & p_k = f_y(c_k s, \eta_k) && (k=0,1,2,3), \\ & \eta_k = u_k + \theta_k (U_k - u_k) && (0<\theta_k<1). \end{aligned} \tag{2.5}$$

Hilfssatz 2.1. *Für die Differenzen $p_j - p_k$ $(j \neq k)$ erhält man (unter Anwendung des Mittelwertsatzes)*

$$\begin{aligned} p_j - p_k &= (c_j - c_k)\, s f_{yx}(\dots,\dots) + (\eta_j - \eta_k)\, f_{yy}(\dots,\dots), \\ \eta_j - \eta_k &= s\left((1-\theta_j)\, c_j f(\dots,\dots) + \theta_j \sum_{m=0}^{j-1} a_{jm} f(c_m s, U_m)\right) \\ &\quad - s\left((1-\theta_k)\, c_k f(\dots,\dots) + \theta_k \sum_{m=0}^{k-1} a_{km} f(c_m s, U_m)\right). \end{aligned}$$

(Die Punkte deuten hier und im folgenden an, daß jeweils ein geeigneter Zwischenwert einzusetzen ist.)

Falls $c_j = c_k$ $(j \neq k)$ *ist, gilt*

$$p_j - p_k = \big(\theta_j(U_j - u_j) - \theta_k(U_k - u_k)\big) f_{yy}(\ldots, \ldots) = O(s^2)$$

$\big(U_n - u_n = O(s^2)$; *man beachte Schema* 3.1 *und* (3.7)$\big)$.

3. Ablauf der Runge-Kutta-Verfahren

Mit unseren Festlegungen läßt sich der Ablauf eines Schrittes durch folgendes Schema beschreiben (man beachte (2.1), (2.4) und (2.5)):

Schema 3.1.

$$U_0 = \quad 0 \quad = u_0,$$

$$U_0' = f(c_0\, s, U_0) = u_0',$$

$$U_1 = \quad s\, a_{10}\, U_0' \quad = u_1 + E_1,$$

$$U_1' = f(c_1\, s, U_1) = u_1' + E_1\, p_1,$$

$$U_2 = s \sum_{k=0}^{1} a_{2k}\, U_k' = u_2 + E_2 + s\, a_{21}\, E_1\, p_1,$$

$$U_2' = f(c_2\, s, U_2) = u_2' + E_2\, p_2 + s\, a_{21}\, E_1\, p_1\, p_2,$$

$$U_3 = s \sum_{k=0}^{2} a_{3k}\, U_k' = u_3 + E_3 + s(a_{31}\, E_1\, p_1 + a_{32}\, E_2\, p_2) + s^2\, a_{32}\, a_{21}\, E_1\, p_1\, p_2,$$

$$\begin{aligned} U_3' &= f(c_3\, s, U_3) \\ &= u_3' + E_3\, p_3 + s(a_{31}\, E_1\, p_1 + a_{32}\, E_2\, p_2)\, p_3 + s^2\, a_{32}\, a_{21}\, E_1\, p_1\, p_2\, p_3 \end{aligned}$$

und

$$U_4 = s \sum_{k=0}^{3} a_{4k}\, U_k' = u_4 + E.$$

Der Schrittfehler (Runge-Kutta-Restglied) E ist gegeben durch

$$\begin{aligned} E &= E_4 + s\, F_3 + s^2\, F_2 + s^3\, F_1 \\ &= E_4 + s(a_{41}\, E_1\, p_1 + a_{42}\, E_2\, p_2 + a_{43}\, E_3\, p_3) \\ &\quad + s^2\big(a_{42}\, a_{21}\, E_1\, p_1\, p_2 + a_{43}(a_{31}\, E_1\, p_1 + a_{32}\, E_2\, p_2)\, p_3\big) \\ &\quad + s^3\, a_{43}\, a_{32}\, a_{21}\, E_1\, p_1\, p_2\, p_3. \end{aligned} \tag{3.2}$$

Falls die Bedingungen

$$s(a_{40}u'_0+a_{41}u'_1+a_{42}u'_2+a_{43}u'_3)-u_4=O(s^5), \qquad (3.3)$$

$$a_{41}E_1p_1+a_{42}E_2p_2+a_{43}E_3p_3=O(s^4), \qquad (3.4)$$

$$a_{42}a_{21}E_1p_1p_2+a_{43}(a_{31}E_1p_1+a_{32}E_2p_2)p_3=O(s^3), \qquad (3.5)$$

$$a_{43}a_{32}a_{21}E_1p_1p_2p_3=O(s^2) \qquad (3.6)$$

erfüllt sind, liegt ein Verfahren der Ordnung vier vor.

Aus den Bedingungen (3.3)−(3.6) *lassen sich Gleichungen für die Koeffizienten herleiten* (vgl. Zeller [8]):

Die Stützstellen wählt man so, daß

$$c_n=a_{n0}+\cdots+a_{n,n-1} \qquad (n=1,2,3) \qquad (3.7)$$

gilt. Dann sind die $E_n=O(s^2)$; insbesondere ist (3.6) erfüllt.

Bedingung (3.3) liefert durch Betrachtung der Fälle $u(x)=x, x^2, x^3, x^4$ folgende Gleichungen:

$$a_{40}+a_{41}+a_{42}+a_{43}=1,$$

$$a_{41}c_1+a_{42}c_2+a_{43}c_3=\tfrac{1}{2},$$

$$a_{41}c_1^2+a_{42}c_2^2+a_{43}c_3^2=\tfrac{1}{3},$$

$$a_{41}c_1^3+a_{42}c_2^3+a_{43}c_3^3=\tfrac{1}{4}.$$

Bedingung (3.4) führt durch Betrachtung der drei Fälle

$$u(x)=x^2 \quad \text{und} \quad p_k=\text{const} \quad \text{oder} \quad p_k=c_k,$$

$$u(x)=x^3 \quad \text{und} \quad p_k=\text{const}$$

auf die Gleichungen

$$a_{41}c_1^2+a_{42}c_2^2+a_{43}c_3^2=2\big(a_{42}a_{21}c_1+a_{43}(a_{31}c_1+a_{32}c_2)\big),$$

$$a_{41}c_1^3+a_{42}c_2^3+a_{43}c_3^3=3\big(a_{42}a_{21}c_1^2+a_{43}(a_{31}c_1^2+a_{32}c_2^2)\big),$$

$$a_{41}c_1^3+a_{42}c_2^3+a_{43}c_3^3=2\big(a_{42}a_{21}c_1c_2+a_{43}(a_{31}c_1+a_{32}c_2)c_3\big).$$

Entsprechend ergibt sich aus (3.5) mit $u(x)=x^2$ die Gleichung

$$a_{42}a_{21}c_1^2+a_{43}(a_{31}c_1^2+a_{32}c_2^2)=2a_{43}a_{32}a_{21}c_1.$$

Zusammen mit (3.7) erhält man damit die gewohnten 11 Gleichungen (vgl. Kopal [5], S. 201). (Diese 11 Gleichungen seien mit (3.8) bezeichnet.)

Anmerkungen. Wir haben 11 Gleichungen für 13 Unbekannte; zwei weitere konsistente Gleichungen können beliebig angenommen werden. Man kann auch zeigen, daß das Gleichungssystem (3.8) notwendig und hinreichend für die Bedingungen (3.3)–(3.6) ist; der Beweis wird über Taylor-Abgleich geführt (siehe [6]).

4. Allgemeine Fehlerdarstellung

Das Runge-Kutta-Restglied (3.2) kann umgeformt werden. Mit Hilfe der Taylor-Entwicklung für p_j $(j=1, 2, 3)$ an der Stelle $(0, 0)$ wird (vgl. [6])

$$\begin{aligned} F_3 = \bar{F}_3\, p_0 &+ s(a_{41}\, E_1\, c_1 + a_{42}\, E_2\, c_2 + a_{43}\, E_3\, c_3)(f_{yx} + f f_{yy})_{(0,0)} \\ &+ s^2 (a_{41}\, E_1\, A_1 + a_{42}\, E_2\, A_2 + a_{43}\, E_3\, A_3), \end{aligned} \tag{4.1}$$

wobei $\bar{F}_3 = a_{41}\, E_1 + a_{42}\, E_2 + a_{43}\, E_3$ ist; die A_k $(k=1, 2, 3)$ sind Ausdrücke, in denen partielle Ableitungen zweiter und dritter Ordnung von f und zweite Ableitungen von u vorkommen.

Das Glied F_2 läßt sich umformen zu

$$F_2 = \bar{F}_2\, p_2\, p_3 + a_{42}\, a_{21}\, E_1\, p_2 (p_1 - p_3) + a_{43}\, a_{31}\, E_1\, p_3 (p_1 - p_2), \tag{4.2}$$

wobei $\bar{F}_2 = a_{42}\, a_{21}\, E_1 + a_{43} (a_{31}\, E_1 + a_{32}\, E_2)$ ist.

Durch Kombination der Fehlerformeln E_n erhält man Fehlerglieder höherer Ordnung:

$$\bar{F}_3 = O(s^4), \qquad \bar{F}_2 = O(s^3), \qquad a_{41}\, E_1\, c_1 + a_{42}\, E_2\, c_2 + a_{43}\, E_3\, c_3 = O(s^3).$$

Dann hat jeder der drei Terme in (4.1) bzw. (4.2) die Ordnung $O(s^4)$ bzw. $O(s^3)$.

Mit Hilfe der Ausdrücke (4.1) für F_3 und (4.2) für F_2 ergibt sich folgender Satz:

Satz 4.1. *Für den Schrittfehler E der Runge-Kutta-Verfahren vierter Ordnung gilt:*

$$\begin{aligned} E = E_4 & \\ &+ s(a_{41}\, E_1 + a_{42}\, E_2 + a_{43}\, E_3)\, p_0 \\ &+ s^2 (a_{41}\, E_1\, c_1 + a_{42}\, E_2\, c_2 + a_{43}\, E_3\, c_3)(f_{yx} + f f_{yy})_{(0,0)} \\ &+ s^3 (a_{41}\, E_1\, A_1 + a_{42}\, E_2\, A_2 + a_{43}\, E_3\, A_3) \\ &+ s^2 \big(a_{42}\, a_{21}\, E_1 + a_{43} (a_{31}\, E_1 + a_{32}\, E_2)\big)\, p_2\, p_3 \\ &+ s^2 \big(a_{42}\, a_{21}\, p_2 (p_1 - p_3) + a_{43}\, a_{31}\, p_3 (p_1 - p_2)\big)\, E_1 \\ &+ s^3\, a_{43}\, a_{32}\, a_{21}\, E_1\, p_1\, p_2\, p_3; \end{aligned} \tag{4.3}$$

jedes der sieben Fehlerglieder hat die Ordnung $O(s^5)$.

Anmerkung. Die sieben Terme in (4.3) bestehen aus Kombinationen verschiedener Integrationsfehlerglieder E_n. Partielle Ableitungen der rechten Seite der Differentialgleichung treten nur bis zur Ordnung drei auf (a-posteriori-Schranke).

5. Spezielle Runge-Kutta-Formeln

Falls die zwei freien Parameter in dem Gleichungssystem (3.8) geeignet gewählt werden, lassen sich die partiellen Ableitungen dritter Ordnung von f in der allgemeinen Fehlerdarstellung (4.3) vermeiden und wir erreichen übersichtlichere Fehlerschranken (siehe Abschnitt 6).

Unter Vermeidung partieller Ableitungen dritter Ordnung spalten wir F_3 in drei Glieder auf:

$$F_3 = \bar{F}_3\, p_j + \Big(\sum_{\substack{k=1\\k\neq j}}^{3} a_{4k} E_k\Big)(p_\mu - p_j) + a_{4\nu} E_\nu (p_\nu - p_\mu) \qquad (\mu \neq j,\ \nu \neq j, \mu;\ j = 1, 2, 3). \tag{5.1}$$

Forderung. Jedes der drei Glieder in (5.1) soll die Ordnung $O(s^4)$ haben.

Wir stellen dafür hinreichende Bedingungen im Einklang mit dem Gleichungssystem (3.8) auf.

Folgende Aufspaltungen von F_3 im Sinne von (5.1) sind möglich.

1. Möglichkeit

$$\begin{aligned} F_3 &= \bar{F}_3\, p_3 + (a_{41} E_1 + a_{42} E_2)(p_1 - p_3) + a_{42} E_2 (p_2 - p_1),\\ F_3 &= \bar{F}_3\, p_3 + (a_{41} E_1 + a_{42} E_2)(p_2 - p_3) + a_{41} E_1 (p_1 - p_2). \end{aligned} \tag{5.2}$$

Bei dieser Zerlegung ist F_3 von vierter Ordnung in s, wenn eine der folgenden Bedingungen erfüllt ist:

a) $a_{41} E_1 + a_{42} E_2 = O(s^3)$ und $p_2 - p_1 = O(s^2)$, d.h.
$a_{41} c_1^2 + a_{42} c_2^2 = 2 a_{42} a_{21} c_1$ und $c_2 = c_1$.

b) $a_{41} E_1 + a_{42} E_2 = O(s^3)$ und $E_2 = O(s^3)$, d.h.
$a_{41} c_1^2 + a_{42} c_2^2 = 2 a_{42} a_{21} c_1$ und $a_{21} c_1 = \frac{1}{2} c_2^2$.

c) $p_1 - p_3 = O(s^2)$ und $E_2 = O(s^3)$, d.h.
$c_1 = c_3$ und $a_{21} c_1 = \frac{1}{2} c_2^2$.

2. Möglichkeit

$$\begin{aligned} F_3 &= \bar{F}_3\, p_2 + (a_{41} E_1 + a_{43} E_3)(p_1 - p_2) + a_{43} E_3 (p_3 - p_1),\\ F_3 &= \bar{F}_3\, p_2 + (a_{41} E_1 + a_{43} E_3)(p_3 - p_2) + a_{41} E_1 (p_1 - p_3). \end{aligned} \tag{5.3}$$

In diesem Fall ist F_3 von vierter Ordnung in s, wenn eine der folgenden Bedingungen erfüllt ist:

a) $a_{41} E_1 + a_{43} E_3 = O(s^3)$ und $p_3 - p_1 = O(s^2)$, d.h.
$a_{41} c_1^2 + a_{43} c_3^2 = 2a_{43}(a_{31} c_1 + a_{32} c_2)$ und $c_3 = c_1$.

b) $a_{41} E_1 + a_{43} E_3 = O(s^3)$ und $E_3 = O(s^3)$, d.h.
$a_{41} c_1^2 + a_{43} c_3^2 = 2a_{43}(a_{31} c_1 + a_{32} c_2)$ und $a_{31} c_1 + a_{32} c_2 = \frac{1}{2} c_3^2$.

c) $p_1 - p_2 = O(s^2)$ und $E_3 = O(s^3)$, d.h.
$c_1 = c_2$ und $a_{31} c_1 + a_{32} c_2 = \frac{1}{2} c_3^2$.

3. Möglichkeit

$$
\begin{aligned}
F_3 &= \bar{F}_3 p_1 + (a_{42} E_2 + a_{43} E_3)(p_2 - p_1) + a_{43} E_3 (p_3 - p_2), \\
F_3 &= \bar{F}_3 p_1 + (a_{42} E_2 + a_{43} E_3)(p_3 - p_1) + a_{42} E_2 (p_2 - p_3).
\end{aligned} \tag{5.4}
$$

F_3 ist von vierter Ordnung in s, wenn eine der nachstehenden Bedingungen erfüllt ist:

a) $a_{42} E_2 + a_{43} E_3 = O(s^3)$ und $E_3 = O(s^3)$, d.h.
$a_{42} c_2^2 + a_{43} c_3^2 = \frac{1}{3}$ und $a_{31} c_1 + a_{32} c_2 = \frac{1}{2} c_3^2$.

b) $a_{42} E_2 + a_{43} E_3 = O(s^3)$ und $E_2 = O(s^3)$, d.h.
$a_{42} c_2^2 + a_{43} c_3^2 = \frac{1}{3}$ und $a_{21} c_1 = \frac{1}{2} c_2^2$.

c) $p_2 - p_1 = O(s^2)$ und $E_3 = O(s^3)$, d.h.
$c_2 = c_1$ und $a_{31} c_1 + a_{32} c_2 = \frac{1}{2} c_3^2$.

d) $p_3 - p_1 = O(s^2)$ und $E_2 = O(s^3)$, d.h.
$c_3 = c_1$ und $a_{21} c_1 = \frac{1}{2} c_2^2$.

Kopal [5] diskutiert ausführlich die Auflösung des Gleichungssystems (3.8). Wir prüfen nach, welche RKV zusätzlich einer der oben aufgestellten Bedingungen genügen.

Ergebnis. *Genau folgende Runge-Kutta-Formeln vierter Ordnung genügen mindestens einer der obenstehenden Bedingungen:*

RKV 5.5.

$a_{40} = \frac{1}{6}$

$a_{41} = \frac{2}{3} - a_{42}$ $\quad c_1 = \frac{1}{2}$ $\quad a_{10} = \frac{1}{2}$

$a_{42} = a_{42}$ $\quad c_2 = \frac{1}{2}$ $\quad a_{20} = \frac{1}{2} - \dfrac{1}{6a_{42}}$ $\quad a_{21} = \dfrac{1}{6a_{42}}$

$a_{43} = \frac{1}{6}$ $\quad c_3 = 1$ $\quad a_{30} = 0$ $\quad a_{31} = 1 - 3a_{42}$ $\quad a_{32} = 3a_{42}$.

(a_{42} ist beliebig wählbar ($\neq 0$); für $a_{42}=\frac{1}{3}$ liegt das gewöhnliche (klassische) RKV vor.)

RKV 5.6. $a_{40}=\frac{1}{6}$

$$\begin{array}{llll} a_{41}=0 & c_1=c_1 & a_{10}=c_1 & \\ a_{42}=\frac{2}{3} & c_2=\frac{1}{2} & a_{20}=\frac{1}{2}-\dfrac{1}{8c_1} & a_{21}=\dfrac{1}{8c_1} \\ a_{43}=\frac{1}{6} & c_3=1 & a_{30}=\dfrac{1}{2c_1}-1 & a_{31}=-\dfrac{1}{2c_1} \quad a_{32}=2. \end{array}$$

(c_1 ist beliebig wählbar zwischen 0 und 1.)

RKV 5.7.

$$\begin{array}{llll} a_{40}=\frac{1}{6} & & & \\ a_{41}=\frac{1}{6}-a_{43} & c_1=1 & a_{10}=1 & \\ a_{42}=\frac{2}{3} & c_2=\frac{1}{2} & a_{20}=\frac{3}{8} & a_{21}=\frac{1}{8} \\ a_{43}=a_{43} & c_3=1 & a_{30}=1-\dfrac{1}{4a_{43}} & a_{31}=-\dfrac{1}{12a_{43}} \quad a_{32}=\dfrac{1}{3a_{43}}. \end{array}$$

(a_{43} ist beliebig wählbar ($\neq 0$).)

Anmerkung. Falls $f(x, y)$ für $y\to\pm\infty$ beschränkt bleibt, so ist auch das RKV 5.5 für $a_{42}=0$ und das RKV 5.7 für $a_{43}=0$ sinnvoll.

6. Restglieder

In der allgemeinen Fehlerdarstellung (4.3) ersetzen wir den Ausdruck für F_3 (vgl. (4.1)) durch die neu gewonnene Zerlegung (5.2). Dann erhalten wir für die RKV 5.5 und 5.6 ein Restglied, in dem die dritten partiellen Ableitungen von f nicht explizit auftreten.

A. Ergebnis für das RKV 5.5

$$\begin{aligned} E=E_4&+\frac{s}{6}E_{123}\,p_3+\frac{s^2}{3}E_{112}\,p_2\,p_3+\frac{s^3}{12}E_1\,p_1\,p_2\,p_3 \\ &+\tfrac{2}{3}s\,E_{12}(p_1-p_3)+\frac{s^2}{6}E_1\,p_2(p_1-p_3) \\ &+s\,a_{42}\,E_2(p_2-p_1)+\frac{s^2}{6}(1-3a_{42})\,E_1\,p_3(p_1-p_2); \end{aligned} \tag{6.1}$$

dabei treten folgende Integrationsfehlerglieder auf:

$$E_4 = \frac{s}{6}(u_0' + 4u_1' + u_2') - u_3 = \frac{s^5}{2880} u^{(5)}(w_1) \qquad \text{(Kepler-Simpson)},$$

$$E_{123} = 6(\tfrac{2}{3} - a_{42}) E_1 + 6a_{42} E_2 + E_3$$

$$= s(u_0' + 2u_1') - 4u_1 - u_3 = -\frac{s^4}{96} u^{(4)}(w_2) \qquad \text{(Spezielle Hermite-Interpolation)},$$

$$E_{112} = (1 - \tfrac{3}{2} a_{42}) E_1 + \tfrac{3}{2} a_{42} E_2$$

$$= \frac{s}{4}(u_0' + u_1') - u_2 \quad = \quad \frac{s^3}{96} u'''(w_3) \qquad \text{(Trapez)},$$

$$E_1 = \frac{s}{2} u_0' \qquad - u_1 \qquad = -\frac{s^2}{8} u''(w_4) \qquad \text{(Tangente links)}, \tag{6.2}$$

$$E_{12} = (1 - \tfrac{3}{2} a_{42}) E_1 + \tfrac{3}{2} a_{42} E_2 = E_{112},$$

$$E_2 = s\left(\left(\tfrac{1}{2} - \frac{1}{6a_{42}}\right) u_0' + \frac{1}{6a_{42}} u_1'\right) - u_2$$

$$= \frac{s^2}{8}\left(\frac{2}{3a_{42}} - 1\right) u_0'' + \frac{s^3}{48}\left(\frac{1}{a_{42}} - 1\right) u_0''' + O(s^4),$$

$$E_2 = \frac{s}{2} u_1' - u_2 = \frac{s^2}{8} u''(w_5) \qquad (\text{falls } a_{42} = \tfrac{1}{3}) \text{ (Tangente rechts)},$$

$$E_2 = E_{112} \qquad (\text{falls } a_{42} = \tfrac{2}{3});$$

$w_1, \ldots, w_5$ sind geeignete Werte zwischen 0 und s; Herleitungen dieser geläufigen Restglieder findet man etwa bei Hildebrand [4].

Für die Differenzen $p_1 - p_3$, $p_2 - p_1$ ergeben sich Abschätzungen aus Hilfssatz 2.1.

Anmerkungen. Die ersten vier Terme in (6.1) sind unabhängig von a_{42}; d.h. für eine lineare DGL mit konstanten Koeffizienten ist der Schrittfehler E unabhängig von a_{42}.

Falls $a_{42} = \frac{1}{3}$ ist, liegt das gewöhnliche RKV vor und man erhält aus (6.1) die von Stetter und Zeller [7] angegebene Fehlerschranke.

Falls $a_{42} = \frac{2}{3}$ ist, liegt ein Spezialfall des RKV 5.6 vor.

B. Ergebnis für das RKV 5.6

$$E = E_4 + \frac{s}{6} E_{23} p_3 + \frac{s^2}{3} \bar{E}_2 p_2 p_3 + \frac{s^3}{24c_1} \bar{E}_1 p_1 p_2 p_3$$
$$+ \tfrac{2}{3} s \bar{E}_2 (p_2 - p_3) + \frac{s^2}{12c_1} \bar{E}_1 p_1 (p_2 - p_3); \tag{6.3}$$

dabei treten folgende Integrationsfehlerglieder auf:

E_4 (Kepler-Simpson) (siehe (6.2)),

$E_{23} = 4\bar{E}_2 + \bar{E}_3 = E_{123}$ (siehe (6.2)),

$$\bar{E}_1 = s\,c_1\,u_0' - u_1 = -\frac{s^2}{2}\,c_1^2\,u''(w_6) \quad \text{(Tangente links)},$$

$$\begin{aligned} \bar{E}_2 &= s\left(\left(\tfrac{1}{2} - \frac{1}{8c_1}\right) u_0' + \frac{1}{8c_1}\,u_1'\right) - u_2 \\ &= \frac{s^3}{16}\,(c_1 - \tfrac{1}{3})\,u_0''' + O(s^4), \end{aligned} \tag{6.4}$$

$\bar{E}_2 = E_{112}$ (falls $c_1 = \frac{1}{2}$) (siehe (6.2)),

$$\bar{E}_2 = \frac{s}{2}\,u_1' \qquad - u_2 = -\frac{s^3}{192}\,u'''(w_7) \quad \text{(falls } c_1 = \tfrac{1}{4}\text{) (Mittelpunkt)},$$

$$\bar{E}_2 = \frac{s}{8}\,(u_0' + 3u_1') - u_2 = -\frac{s^4}{3456}\,u^{(4)}(w_8) \quad \text{(falls } c_1 = \tfrac{1}{3}\text{) (Radau)};$$

$w_6, \ldots, w_8$ sind geeignete Werte zwischen 0 und s. (Die Striche auf $\bar{E}_1$, $\bar{E}_2$ dienen nur der Unterscheidung zum RKV 5.5.)

Für die Differenz $p_2 - p_3$ ergibt sich eine Abschätzung aus Hilfssatz 2.1.

Bemerkungen zur Wahl von c_1 für das RKV 5.6

Die ersten zwei Terme in (6.3) sind unabhängig von c_1. Für eine lineare DGL mit konstanten Koeffizienten gilt:

$$|E| \leqq |E_4| + \frac{s}{6}\,|E_{23}| + \frac{s^2}{3}\,|\bar{E}_2| + \frac{s^3}{24c_1}\,|\bar{E}_1|.$$

Um die Schranke für $\frac{s^2}{3}\,\bar{E}_2$ und für $\frac{s^3}{24c_1}\,\bar{E}_1$ miteinander vergleichen zu können, nehmen wir an, daß die Schranke für u'' etwa so groß wie die Schranke für u''' sei. Wir erhalten dann eine minimale Schranke für E, falls

$$W(c_1) := |c_1 - \tfrac{1}{3}| + c_1$$

minimal wird. Es ist

$$W(c_1) = \begin{cases} \tfrac{1}{3}, & \text{für } 0 < c_1 \leqq \tfrac{1}{3} \\ 2c_1 - \tfrac{1}{3}, & \text{für } \tfrac{1}{3} \leqq c_1 \leqq 1. \end{cases}$$

Im allgemeinen Fall ist auch der Ausdruck

$$2\left(\frac{s}{3}\bar{E}_2+\frac{s^2}{24c_1}\bar{E}_1 p_1\right)(p_2-p_3)$$

in Abhängigkeit von c_1 zu betrachten. Falls p_1 die Größenordnung Eins hat, ist der erste Term auf den obigen Fall zurückgeführt. Für die Differenz p_2-p_3 folgt aus Hilfssatz 2.1 die Abschätzung

$$|p_2-p_3|\leqq\frac{s}{2}|f_{yx}(\ldots,\ldots)|+|\eta_2-\eta_3|\,|f_{yy}(\ldots,\ldots)|; \tag{6.5}$$

dabei ist

$$|\eta_2-\eta_3|\leqq\begin{cases}\frac{7}{2}s|f(\ldots,\ldots)|, & \text{für } c_1=\frac{1}{2}\\ \frac{9}{2}s|f(\ldots,\ldots)|, & \text{für } c_1=\frac{1}{3}\\ \frac{11}{2}s|f(\ldots,\ldots)|, & \text{für } c_1=\frac{1}{4}.\end{cases}$$

Anmerkung. Für das RKV 5.7 kann eine ähnliche Fehlerdarstellung angegeben werden wie für das RKV 5.5 bzw. 5.6 (diese Schranke ist allerdings ungünstiger). Der Fall $a_{41}=0$ ist in 5.6 enthalten.

7. Fehlerabschätzungen

Die rechte Seite der DGL möge in dem erwähnten Gebiet G nachstehenden Abschätzungen genügen:

$$|f|\leqq D_0,\qquad |f_y|\leqq L,\qquad |f_{yx}|\leqq G_1,\qquad |f_{yy}|\leqq G_2;$$

die exakte Lösung u genüge in dem betrachteten Intervall den Ungleichungen

$$|u^{(j+1)}|\leqq D_j\qquad (j=1,\ldots,4).$$

Für den Schrittfehler E des RKV 5.5 bzw. 5.6 erhalten wir aus (6.1) bzw. (6.3) folgende Abschätzungen ($K_1,\ldots,K_3$ bzw. $K_4,\ldots,K_7$ sind Konstanten, die von der Wahl des Parameters a_{42} bzw. c_1 abhängen).

A. Für das RKV 5.5

$$\begin{aligned}|E|\leqq\frac{s^5}{2880}\Big(&D_4+5D_3L+10D_2L^2+30D_1L^3\\&+10(D_2+3D_1L)(G_1+3K_1D_0G_2)+30D_1^2G_2K_2\left(1+\frac{s}{4}LK_3\right)\Big);\end{aligned}\tag{7.1}$$

für das gewöhnliche RKV ($a_{42}=\frac{1}{3}$) ist $K_1=K_2=K_3=1$.

B. Für das RKV 5.6

$$|E| \leqq \frac{s^5}{2880}\left(D_4+5D_3 L+5K_4 D_2 L^2+15K_5 D_1 L^3 \right.$$
$$\left. +5K_6(D_2+3D_1 L)(G_1+K_7 D_0 G_2)\right); \tag{7.2}$$

für $c_1=\frac{1}{2}$ ist $K_4=K_5=K_6=2$, $K_7=7$;
für $c_1=\frac{1}{4}$ ist $K_4=K_5=K_6=1$, $K_7=11$;
für $c_1=\frac{1}{3}$ läßt sich die Schranke D_2 für u''' vermeiden:

$$|E| \leqq \frac{s^5}{2880}\left(D_4+5D_3 L+\frac{5s}{18} D_3 L^2+20D_1 L^3\right.$$
$$\left.+5\left(\frac{s}{18} D_3+4D_1 L\right)(G_1+9D_0 G_2)\right). \tag{7.3}$$

Vergleich der Fehlerabschätzungen. In der Abschätzung (7.2) und (7.3) für das RKV 5.6 tritt das etwas störende Glied

$$30D_1^2 G_2 K_2\left(1+\frac{s}{4} LK_3\right)$$

nicht auf. Falls die Schranke $D_0 G_2$ für ff_{yy} nicht zu stark ins Gewicht fällt, sind die Abschätzungen (7.2) und (7.3) günstiger als (7.1) für das gewöhnliche RKV. In der Abschätzung (7.3) wird die Schranke D_2 für u''' vermieden. Insgesamt scheint die 3-Term-Runge-Kutta-Formel

$$U_4=\frac{s}{6}(U_0'+4U_2'+U_3')$$

sehr günstig zu sein.

Schlußbemerkung. Die hier gezeigte Methode der Herleitung von Runge-Kutta-Verfahren und der daraus resultierenden Fehlerdarstellungen läßt sich auch auf Verfahren höherer Ordnung anwenden. Ebenso können Systeme von Differentialgleichungen behandelt werden.

Literatur

1. Albrecht, J.: Beiträge zum Runge-Kutta-Verfahren. Z. angew. Math. Mech. **35**, 100–110 (1955).
2. Bieberbach, L.: On the remainder of the Runge-Kutta formula in the theory of ordinary differential equations. Z. angew. Math. Phys. **2**, 233–248 (1951).
3. Butcher, J. C.: Coefficients for the study of Runge-Kutta integration processes. J. Austral. Math. Soc. **3**, 185–201 (1963).
4. Hildebrand, F. B.: Introduction to numerical analysis. New York: McGraw-Hill Book Co. 1956.

5. Kopal, Z.: Numerical analysis. London: Wiley 1961.
6. Scherer, R.: Rekursiver Aufbau der Runge-Kutta-Verfahren in Verbindung mit Integrationsformeln. Diss. Tübingen 1972.
7. Stetter, F., Zeller, K.: Fehleruntersuchungen für das gewöhnliche Runge-Kutta-Verfahren. Math. Z. **98**, 179–184 (1967).
8. Zeller, K.: Runge-Kutta-Approximationen. Abstract Spaces Approx., Proc. Conf. Math. Res. Inst. Oberwolfach 1968. ISNM **10**, 365–366 (1969), Basel: Birkhäuser.

Dr. R. Scherer
Mathematisches Institut der Universität
D-74 Tübingen
Brunnenstr. 27
Bundesrepublik Deutschland

(Eingegangen am 3. Juli 1972)

Math. Z. 128, 325–341 (1972)

Zur Theorie Levitzkischer Radikale in Halbringen

Hanns Joachim Weinert

§ 1. Einleitung

Besitzt ein Halbring S ein annullierendes Nullelement 0, so lassen sich die Begriffe „nil", „lokal nilpotent", „nilpotent", „Annullatorideal" wie für Ringe definieren [1]. Für solche Halbringe S (mit kommutativer Addition) hat Barbut in [1] die Existenz eines größten lokal nilpotenten Ideals $L(S)$ nachgewiesen und als Levitzkisches Radikal von S definiert. Seine Arbeit zeigt, daß diese Begriffsbildung in der Tat gestattet, entsprechende ringtheoretische Aussagen auf Halbringe zu übertragen. Wir zitieren dafür zwei Ergebnisse aus [1] (Lemma 6 und Theorem):

A) *Ist $L(S)$ ein k-Ideal, gilt* $L(S/L(S))=(0)$ (vgl. 2.4, 2.5).

B) *Ist S ein Halbring mit aufsteigender Kettenbedingung für links- und rechtsseitige Annullatorideale, und ist $L(S)$ ein k-Ideal, so ist jeder Unterhalbring U von S mit $0 \in U$, der nil ist, auch schon nilpotent* [2].

Dies wird noch deutlicher durch die Feststellung (vgl. [19]), daß die einschränkende Voraussetzung „ist $L(S)$ ein k-Ideal" in A) und B) (und in einigen anderen Resultaten von [1]) entfallen kann, da das von Barbut eingeführte Levitzkische Radikal $L(S)$ stets k-abgeschlossen ist. Wir zeigen hier gleich allgemeiner (Satz 5.1), daß für jedes Ideal M von S der k-Abschluß $\bar{A}$ eines lokal M-potenten Ideals A von S (Definition in 3.1) lokal $\bar{M}$-potent ist, woraus für $M=(0)=\overline{(0)}$ sofort $\overline{L(S)}=L(S)$ folgt [3].

Nun besitzt ein Halbring mit annullierendem Nullelement 0 trivialerweise ein (allgemein als Kern bezeichnetes) kleinstes Ideal K, nämlich $K=(0)=\{0\}$. Es liegt daher nahe, die hier diskutierten Begriffsbildungen für Halbringe S mit Kern K zu verallgemeinern, indem man sie (statt auf

---

[1] Die in dieser Arbeit verwendeten Begriffsbildungen und einige elementare Aussagen über Halbringe sind im folgenden § 2 zusammengestellt.

[2] Durch $0 \in U$ wird gewährleistet, daß sich „nil" und „nilpotent" für U nicht auf ein Nullelement 0_U dieses Unterhalbrings mit $0_U \neq 0 = 0_S$ bezieht. Ohne diese Voraussetzung (sie fehlt in [1]) können leicht Gegenbeispiele für B) gegeben werden.

[3] Für diesen Spezialfall kann der Beweis von Satz 5.1 unabhängig vom übrigen Inhalt dieser Arbeit gelesen werden. Die möglichen Vereinfachungen für $M=(0)=\overline{(0)}$ sind übrigens unerheblich.

(0)) auf das Ideal K bezieht. Dies wurde bereits früher von Costa in [6] vorgeschlagen. Doch wird in [6] nur gezeigt, daß für einen solchen Halbring S (mit kommutativer Addition) die Summe endlich vieler lokal K-potenter Rechtsideale wieder lokal K-potent ist, woraus sich leicht die Existenz eines größten lokal K-potenten Ideals $L_K(S)$ von S ergibt. Die Behauptung in [6], daß daraus eine Theorie der Levitzkischen Radikale $L_K(S)$ für Halbringe S mit Kern K wie für Ringe folgt[4], trifft jedoch nur bedingt zu.

Versucht man nämlich wie in Barbut [1] weitergehende ringtheoretische Aussagen nun auf (zunächst additiv kommutative) Halbringe S mit Kern K zu übertragen, benötigt man sehr bald, daß $L_K(S)$ ein k-Ideal ist. Nach dem oben genannten Satz 5.1 ist dies gewährleistet, wenn K selbst $K=\bar{K}$ erfüllt, nicht aber für Halbringe mit $K\subset\bar{K}$. Diese Schwierigkeiten werden behoben, wenn man für Halbringe S mit Kern K von vornherein alle Begriffsbildungen auf $\bar{K}$ bezieht und das größte lokal $\bar{K}$-potente Ideal $L_{\bar{K}}(S)$ als Levitzkisches Radikal von S definiert. Die Untersuchung des so definierten Levitzkischen Radikals und der Vergleich mit $L_K(S)$ ist Gegenstand der vorliegenden Arbeit. In diesem Zusammenhang kommen wir auch auf das in Costa [6, 7] definierte klassische Radikal $N_K(S)$ zu sprechen, wobei sich ebenfalls $N_{\bar{K}}(S)$ als die zweckmäßigere Begriffsbildung herausstellt.

Im Hinblick auf diese Zielstellung beweisen wir in § 3 einige Aussagen über lokal M-potente und M-potente Ideale, bezogen auf ein beliebiges Ideal M eines Halbrings S. Es folgt die Existenz eines größten lokal M-potenten Ideals $L_M(S)$ und eines kleinsten Ideals $N_M(S)$, welches alle M-potenten Ideale von S enthält. Gewisse radikalähnliche Eigenschaften gelten bereits für alle diese Ideale $L_M(S)$ bzw. $N_M(S)$, wie wir in § 4 zeigen. Aus diesen Ergebnissen und dem bereits genannten Satz 5.1 über den k-Abschluß lokal M-potenter bzw. M-potenter Ideale erhalten wir in § 6 die in Analogie zur Ringtheorie angestrebten radikaltypischen Aussagen für das Levitzkische Radikal $L_{\bar{K}}(S)$ und ähnlich für $N_{\bar{K}}(S)$. Wesentliche Aussagen dieser Art, insbesondere die Radikalfreiheit des Restklassenhalbrings nach dem Radikal, gelten nicht mehr, wenn man $L_K(S)$ bzw. $N_K(S)$ als Radikale definiert[5]. Wir zeigen dies durch ein gemeinsames Gegenbeispiel in § 7.

Weiterhin gehen wir in § 8 auf den Zusammenhang zum Jacobsonschen Radikal $J(S)$ ein, welches Bourne [2] für additiv kommutative Halbringe S mit annullierendem Nullelement eingeführt hat. Wir stellen fest, daß diese Theorie (vgl. auch [3, 11]) nicht von der Existenz eines

---

[4] Diese Behauptung wird in [7] auch für Halbringe mit nicht notwendig kommutativer Addition wiederholt.

[5] Die Radikalfreiheit von $S/N_{\bar{K}}(S)$ bzw. $S/N_K(S)$ ist natürlich unter der Voraussetzung zu diskutieren, daß $N_{\bar{K}}(S)$ selbst $\bar{K}$-potent bzw. $N_K(S)$ selbst K-potent ist.

Nullelementes abhängt, wenn man $J(S)=\varnothing$ zuläßt. Es gilt dann

$$L_{\bar{K}}(S)\subseteqq J(S)$$

für alle Halbringe mit $J(S)\neq\varnothing$.

Eine wichtige, auf Levitzki [13, 14] zurückgehende Anwendung seines Radikals in der Ringtheorie besteht darin, aus gewissen Kettenbedingungen für Annullatorideale von „nil" auf „lokal nilpotent" bzw. „nilpotent" schließen zu können. Wir verweisen auf die schöne Darstellung in Herstein [9, Chap. 5], deren Übertragung auf Halbringe mit annullierendem Nullelement das Hauptanliegen von Barbut [1] ist (vgl. die vorn zitierte Aussage B)). In einer Fortsetzung dieser Arbeit werden wir zeigen, wie sich diese Überlegungen auf Halbringe mit Kern verallgemeinern lassen.

Schließlich bemerken wir: Alle hier skizzierten Gedankengänge (mit Ausnahme von § 8) sind von der Kommutativität der Addition der betrachteten Halbringe prinzipiell unabhängig. In der Tat führen wir die Überlegungen in den §§ 3 und 4 für beliebige Halbringe durch, ohne daß die hier gegebenen Beweise bei kommutativer Addition echt einfacher würden. Dagegen müßten für die folgenden Paragraphen zunächst die Begriffsbildungen des k-Abschlusses eines Ideals A und des Restklassenhalbrings S/A in geeigneter Weise für Halbringe S mit nicht notwendig koummutativer Addition verallgemeinert werden. Dies ist möglich, führt jedoch sofort auf andere damit zusammenhängende Fragen, so daß wir diese Überlegungen unabhängig von der vorliegenden Arbeit veröffentlichen wollen. Daher beschränken wir uns hier ab § 5 auf Halbringe mit kommutativer Addition.

§ 2. Verwendete Begriffsbildungen und elementare Aussagen

2.1. Allgemein versteht man unter einem *Halbring* eine Algebra $S=(S,+,\cdot)$ mit zwei zweistelligen Operationen, so daß $(S,+)$ und $(S,\cdot)$ Halbgruppen sind und die Distributivgesetze wie für Ringe gelten (vgl. Bourne [2], Weinert [17]). Besitzt der Halbmodul $(S,+)$ ein neutrales Element 0, so bezeichnen wir 0 als das *Nullelement* des Halbringes S. Gilt dabei insbesondere $0\,x=x\,0=0$ für alle $x\in S$, sprechen wir von einem *annullierendem Nullelement*.

2.2. Eine nichtleere Teilmenge $A\subseteqq S$ heißt *Rechtsideal* von S, wenn A Unterhalbmodul von $(S,+)$ ist und $a\,x\in A$ für alle $a\in A$, $x\in S$ gilt. Sind $A_\lambda(\lambda\in\Lambda)$ beliebige Teilmengen von S, definiert man als Summe $\sum A_\lambda(\lambda\in\Lambda)$ die Menge aller endlichen Summen

$$\sum_{i=1}^{n} a_i;\quad n=1,2,\ldots\qquad \text{mit } a_i\in\bigcup A_\lambda\ (\lambda\in\Lambda).$$

Für Rechtsideale A_λ ist dann $\sum A_\lambda$ wieder ein Rechtsideal, und zwar das kleinste Rechtsideal A von S mit $A_\lambda \subseteqq A$ für alle $\lambda \in \Lambda$. Entsprechendes gilt für *Linksideale* und (zweiseitige) *Ideale* von S.

2.3. Es sei K der Durchschnitt aller (zweiseitigen) Ideale eines Halbrings S. Genau für $K \neq \emptyset$ besitzt S ein kleinstes Ideal, nämlich K, welches dann der *Kern* von S genannt wird. Ein Halbring mit annullierendem Nullelement 0 hat das Ideal $(0)=\{0\}$ als Kern. Wir vermerken für später (vgl. Costa [6], Th. 2): Ist A ein Ideal eines Halbrings S mit Kern K, so ist K auch Kern des Unterhalbrings A.

2.4. Die folgenden Begriffsbildungen und Aussagen für additiv kommutative Halbringe gehen auf [3, 8, 10, 12] zurück, sind aber von der dort vorausgesetzten Existenz eines annullierenden Nullelementes unabhängig. Ist zunächst A Unterhalbmodul eines kommutativen Halbmoduls $(S, +)$, definiert man den *k-Abschluß* $\bar{A}$ *von* A gemäß

$$\bar{A}=\{x \in S \mid x+a_1=a_2 \text{ für geeignete } a_1, a_2 \in A\}.$$

$\bar{A}$ ist dann ebenfalls Unterhalbmodul von $(S, +)$ und es gilt:

$$A \subseteqq \bar{A}; \quad \bar{\bar{A}}=\bar{A};$$
$$A \subseteqq B \Rightarrow \bar{A} \subseteqq \bar{B}; \quad \overline{\bar{A}+\bar{B}}=\overline{A+B}.$$

Mit A ist auch $\bar{A}$ ein Rechtsideal des Halbrings $(S, +, \cdot)$. Insbesondere spricht man von einem *k-abgeschlossenen Rechtsideal* (oder *k-Rechtsideal*) A, wenn $A=\bar{A}$ gilt.

2.5. Ist A ein (zweiseitiges) Ideal eines additiv kommutativen Halbrings S, so wird durch

$$x \equiv y(A) \Leftrightarrow x+a_1=y+a_2 \text{ für geeignete } a_1, a_2 \in A$$

eine Kongruenz modulo A auf S definiert[6]. Die zugehörige Kongruenzklassenstruktur bezeichnen wir als *Restklassenhalbring* und schreiben S/A. Dabei bilden die Elemente des k-Abschlusses $\bar{A}$ von A genau eine Restklasse modulo A, welche annullierendes Nullelement von S/A ist. Natürlich gilt $S/A=S/\bar{A}$. Der folgende Satz darf als bekannt angesehen werden: *Ist φ ein Epimorphismus eines additiv kommutativen Halbrings S auf einen Halbring S' mit annullierendem Nullelement $0'$, so ist $\varphi^{-1}(0')=A$ ein k-abgeschlossenes Ideal von S und φ läßt sich zerlegen in den natürlichen Epimorphismus $\varphi_1: S \to S/A$ und einen 0-injektiven Epimorphismus* (bei Bourne [2]: Semi-Isomorphismus) φ_2: $S/A \to S'$.

[6] Für Begriffe wie Kongruenz, Kongruenzklassenstruktur, Homomorphismus usw. betrachten wir (beliebige) Halbringe $(S, +, \cdot)$ stets als Algebren bezüglich der Operationen $+$ und $\cdot$ (vgl. etwa Cohn [5]).

§ 3. M-potente und lokal M-potente Ideale in beliebigen Halbringen

3.1. Es sei M ein Ideal und X eine Teilmenge eines Halbrings S. X heißt *M-potent* (für die entsprechende Begriffsbildung in Halbgruppen vgl. [16], [4] Chap. 6), wenn für eine natürliche Zahl n

$$X^n \underset{\mathrm{Df}}{=} \{x_{i_1} \dots x_{i_n} \mid x_{i_\nu} \in X\} \subseteqq M$$

erfüllt ist. Weiterhin heißt X *M-nil* bzw. *lokal M-potent*, wenn jedes Element bzw. jede endliche Untermenge von X M-potent ist. Besitzt der Halbring S ein annullierendes Nullelement 0, so bezeichnen wir diese Begriffsbildungen für $M = K = (0)$ mit „nilpotent", „nil" bzw. „lokal nilpotent" wie in der Ringtheorie.

3.2. Lemma. *Sind A und B lokal M-potente (M-potente) Rechtsideale von S, so gilt das gleiche für $A+B$.*

Beweis. Es sei $X = \{x_1, \dots, x_k\}$ eine endliche Teilmenge von $A+B$. Jedes

$$x_i = \sum_j c_{ij} \in X \tag{1}$$

ist eine endliche Summe von Elementen $c_{ij} \in A \cup B$. Die Menge aller dieser Summanden $C = \{c_{ij}\}$ ist endlich, also auch $A_1 = A \cap C$ und $B_1 = B \cap C$. Da A und B lokal M-potent sind, gibt es eine natürliche Zahl n mit $A_1^n \subseteqq M$ und $B_1^n \subseteqq M$. Wir bilden die endliche Menge

$$A_2 = \{a_i, a_i b_{j_1}, \dots, a_i b_{j_1} \dots b_{j_{n-1}} \mid a_i \in A_1, b_{j_\nu} \in B_1\} \subseteqq A,$$

zu der also eine natürliche Zahl m mit $A_2^m \subseteqq M$ existiert. Dann gilt aber bereits $X^{n(m+1)} \subseteqq M$: Jedes Produkt aus $n(m+1)$ Elementen (1) ist nämlich eine Summe von Produkten der Form

$$c_1 c_2 \dots c_{n(m+1)} \quad \text{mit } c_i \in C. \tag{2}$$

Aus Anzahlgründen treten in (2) entweder n (oder mehr) Faktoren aus B_1 nacheinander auf, oder (2) ist ein Produkt von mindestens m Elementen aus A_2. Also liegt jeder Summand (2) in M.

Sind A und B M-potent, so folgt aus $A^n \subseteqq M$, $B^n \subseteqq M$ auf die gleiche Weise $(A+B)^{n(n+1)} \subseteqq M$.

3.3. Lemma. *Jedes lokal M-potente (M-potente) Rechtsideal A von S ist in einem ebensolchen zweiseitigen Ideal $A+SA$ enthalten.*

Beweis. Nach 3.2 genügt es zu zeigen, daß

$$SA = \left\{ \sum_{i=1}^{n} s_i a_i \mid n = 1, 2, \dots;\ s_i \in S,\ a_i \in A \right\}$$

lokal M-potent ist. Sei X eine endliche Teilmenge von SA. Die in den Elementen $x \in X$ insgesamt auftretenden Faktoren $s_i \in S$ bilden eine endliche Menge S_1, entsprechendes gilt für die Menge A_1 der auftretenden $a_i \in A$. Zu der endlichen Menge

$$A_2 = \{a_i s_j \mid a_i \in A_1, s_j \in S_1\} \subseteqq A$$

existiert eine natürliche Zahl n mit $A_2^n \subseteqq M$. Daraus folgt ersichtlich $X^{n+1} \subseteqq M$.

Ist A M-potent, so ergibt $A^n \subseteqq M$ entsprechend $(SA)^{n+1} \subseteqq M$.

3.4. Aus diesen Lemmata folgt sofort, daß die Vereinigungsmenge $L_M(S)$ aller lokal M-potenten Rechtsideale eines Halbrings S ein Ideal von S ist, welches ebenso die Vereinigungsmenge aller lokal M-potenten Linksideale bzw. Ideale von S ist. Entsprechendes gilt für die Vereinigungsmenge $N_M(S)$ aller M-potenten Rechtsideale von S. Dabei gilt $N_M(S) \subseteqq L_M(S)$, beide Ideale sind lokal M-potent, doch braucht $N_M(S)$ nicht M-potent zu sein.

Folgerung. *Für jedes Ideal M eines Halbrings S gibt es*

a) *ein größtes lokal M-potentes Ideal $L_M(S)$; dabei enthält $L_M(S)$ auch jedes lokal M-potente Rechts- bzw. Linksideal von S;*

b) *ein kleinstes Ideal $N_M(S)$, welches jedes M-potente Rechts- bzw. Linksideal von S enthält.*

Ist S ein Halbring mit annullierendem Nullelement 0 bzw. allgemeiner mit nicht leerem Kern K, so ist $L_0(S)$ bzw. $L_K(S)$ das Levitzkische Radikal von S im Sinne von Barbut [1] bzw. Costa [6, 7]. Entsprechend korrespondiert $N_0(S)$ zum sogenannten kleinen oder klassischen Radikal für Ringe, wofür Costa $N_K(S)$ als Verallgemeinerung vorschlägt. Dagegen werden wir in § 6 (in dieser Arbeit nur für Halbringe mit kommutativer Addition) $L_{\bar{K}}(S)$ als Levitzkisches Radikal von S definieren und für $L_{\bar{K}}(S)$ radikaltypische Eigenschaften analog zur Ringtheorie nachweisen. Dabei stützen wir uns auf die folgenden Aussagen, die allgemein für die Ideale $L_M(S)$ ohne eine bestimmte Festlegung für das zugrunde gelegte Ideal M gelten. Auf die entsprechenden Überlegungen für das klassische Radikal werden wir mit eingehen, ohne jedoch alle Beweise auszuführen.

§ 4. Radikalähnliche Eigenschaften von $L_M(S)$ in beliebigen Halbringen

4.1. Satz. *Es seien M und A Ideale eines Halbrings S. Dann gilt:*

a) $L_M(S) \cap A = L_{M \cap A}(A)$.

b) *Aus* $M \subseteqq A \subseteqq L_M(S)$ *folgt* $L_M(S) = L_A(S)$.

c) $L_{L_M(S)}(S) = L_M(S)$.

Beweis. a) Gilt $L_M(S) \cap A \neq \emptyset$, so ist dieser Durchschnitt ein Ideal von A, welches lokal M-potent, also auch lokal $(M \cap A)$-potent ist. Also gilt $L_M(S) \cap A \subseteqq L_{M \cap A}(A)$[7]. Für die Umkehrung sei $x \in L_{M \cap A}(A)$. Das von x in S erzeugte Rechtsideal $[x)$ besteht aus endlichen Summen $\sum c_i$ mit $c_i = x s_i$, $s_i \in S$ oder $c_i = x$. Sei X eine endliche Teilmenge von $[x)$ und $C = \{c_i\}$ die Menge aller Summanden c_i, die in den Elementen aus X auftreten. Mit C ist $C^2 = \{c_i c_j\}$ endlich, und aus $x \in L_{M \cap A}(A)$ und $x \in A$ folgt $C^2 \subseteqq x A \subseteqq L_{M \cap A}(A)$. Es gibt also eine natürliche Zahl n mit $(C^2)^n \subseteqq M \cap A$, also $X^{2n} \subseteqq M \cap A \subseteqq M$. Daher ist $[x)$ lokal M-potentes Rechtsideal von S, also gilt $x \in L_M(S)$ und $L_{M \cap A}(A) \subseteqq L_M(S)$.

b) Ist X eine endliche Teilmenge von $L_A(S)$, folgt $X^n \subseteqq A \subseteqq L_M(S)$ und damit $(X^n)^m \subseteqq M$ für geeignete natürliche Zahlen n und m. Dies zeigt $L_A(S) \subseteqq L_M(S)$; die Umkehrung folgt sofort aus $M \subseteqq A$.

c) Ist ein Spezialfall von b).

4.2. Satz. *Es sei S ein Halbring, φ ein Epimorphismus von S auf einen Halbring $S' = \varphi(S)$, M' ein Ideal von S' und $\varphi^{-1}(M') = M$. Dann gilt*

$$L_M(S) = \varphi^{-1}(L_{M'}(S')).$$

Beweis. Man zeigt durch direkten Ansatz, daß $\varphi(L_M(S))$ ein lokal M'-potentes Ideal von S' ist, also $\varphi(L_M(S)) \subseteqq L_{M'}(S')$ gilt. Entsprechend folgt, daß $\varphi^{-1}(L_{M'}(S'))$ ein lokal M-potentes Ideal von S ist, also $L_M(S) \supseteqq \varphi^{-1}(L_{M'}(S'))$ gilt.

4.3. Folgerung. *In Satz 4.2 kann das Ideal $M = \varphi^{-1}(M')$ durch ein Ideal M von S ersetzt werden, welches nur $M \subseteqq \varphi^{-1}(M') \subseteqq L_M(S)$ erfüllt.*

Nach Satz 4.1 b) gilt dann nämlich $L_M(S) = L_{\varphi^{-1}(M')}(S)$, und $L_{\varphi^{-1}(M')}(S) = \varphi^{-1}(L_{M'}(S'))$ ist die Behauptung von Satz 4.2.

4.4. Ohne auf die ähnlichen Beweise näher einzugehen, sei bemerkt: Für $N_M(S)$ gelten die Satz 4.1 a) und Satz 4.2 entsprechenden Aussagen ebenfalls, im allgemeinen jedoch nicht die übrigen. Dagegen ist die Satz 4.1 b) entsprechende Aussage (und damit auch c) und Folgerung 4.3) richtig, wenn $N_M(S)$ selbst M-potent ist.

§ 5. *k*-Abschluß (lokal) *M*-potenter Ideale in Halbringen mit kommutativer Addition

5.1. Satz. *Es seien A und M Ideale eines additiv kommutativen Halbrings S. Ist A lokal M-potent (M-potent), so ist $\bar{A}$ lokal $\bar{M}$-potent ($\bar{M}$-potent).*

Beweis. Es sei $X = \{x_1, \ldots, x_k\}$ eine endliche Teilmenge von $\bar{A}$. Zu jedem $x_i \in X$ gibt es ein $a_i \in A$ mit $x_i + a_i = a_i' \in A$. Für die endliche Menge

[7] Daraus folgt $L_M(S) \cap A \neq \emptyset \Rightarrow L_{M \cap A}(A) \neq \emptyset \Rightarrow M \cap A \neq \emptyset \Rightarrow L_M(S) \cap A \neq \emptyset$. Damit bleibt a) formal richtig, wenn einer der auftretenden Durchschnitte leer wird.

$A_1 = \{a_i'\} \subseteqq A$ existiert eine natürliche Zahl n mit $A_1^n \subseteqq M$, also

$$(1) \qquad (x_{i_1} + a_{i_1}) \dots (x_{i_n} + a_{i_n}) \in M.$$

Die k^n Elemente $x_{i_1} \dots x_{i_n}$ der Menge X^n bezeichnen wir in irgendeiner Reihenfolge mit $X_j, j = 1, \dots, k^n$. Das Ausmultiplizieren von (1) zeigt, daß es zu jedem X_j ein $Y_j \in A$ mit

$$(2) \qquad X_j + Y_j \in M$$

gibt. Für die endliche Menge $A_2 = \{Y_j\} \subseteqq A$ existiert eine natürliche Zahl r mit $A_2^r \subseteqq M$. Wir werden zeigen, daß $(X^n)^r = X^{n \cdot r} \subseteqq \overline{M}$ gilt. Dazu betrachten wir ein beliebiges Produkt

$$X_{j_1} \dots X_{j_r} \in (X^n)^r.$$

Für einen Induktionsschluß setzen wir

$$(3) \qquad X_{j_1} \dots X_{j_{t-1}} Y_{j_t} \dots Y_{j_r} + m_{t-1} = m_t \quad \text{mit } m_{t-1}, m_t \in M$$

voraus, wobei der Induktionsanfang $t = 1$ aus $A_2^r \subseteqq M$ folgt. Aus $X_{j_t} + Y_{j_t} \in M$ gemäß (2) ergibt sich

$$(X_{j_1} \dots X_{j_{t-1}}) X_{j_t} (Y_{j_{t+1}} Y_{j_r}) + (X_{j_1} \dots X_{j_{t-1}}) Y_{j_t} (Y_{j_{t+1}} \dots Y_r) \in M.$$

Addiert man dazu m_{t-1} und beachtet (3), so erhält man

$$X_{j_1} \dots X_{j_t} Y_{j_{t+1}} \dots Y_{j_r} + m_t = m_{t+1} \in M.$$

Der letzte Schritt mit $t = r$ liefert

$$X_{j_1} \dots X_{j_r} + m_r \in M, \quad \text{also } X_{j_1} \dots X_{j_r} \in \overline{M}$$

und damit $X^{nr} \subseteqq \overline{M}$; also ist $\bar{A}$ lokal $\overline{M}$-potent.

Ist A M-potent, so folgt aus $A^n \subseteqq M$ mit dem gleichen Beweisgedanken $\bar{A}^{nn} \subseteqq \overline{M}$.

5.2. Folgerung. *Für jedes Ideal M von S gilt*

$$L_M(S) \subseteqq \overline{L_M(S)} \subseteqq L_{\overline{M}}(S) = \overline{L_{\overline{M}}(S)}.$$

Beweis. Der k-Abschluß des lokal $\overline{M}$-potenten Ideals $L_{\overline{M}}(S)$ ist nach Satz 5.1 lokal $(\overline{\overline{M}} = \overline{M})$-potent, also folgt $\overline{L_{\overline{M}}(S)} \subseteqq L_{\overline{M}}(S)$ und damit die Gleichheit dieser Ideale. Die beiden Inklussionen sind damit klar; sie können beide echt sein, wie unser Beispiel in § 7 zeigt.

Insbesondere ist $L_0(S) = \overline{L_0(S)}$ die in der Einleitung behauptete k-Abgeschlossenheit des Levitzkischen Radikals in Barbut [1].

5.3. Entsprechend gilt $N_M(S) \subseteqq \overline{N_M(S)} \subseteqq N_{\overline{M}}(S) = \overline{N_{\overline{M}}(S)}$, doch muß man beim Beweis von $\overline{N_{\overline{M}}(S)} \subseteqq N_{\overline{M}}(S)$ auf die Kennzeichnung von $N_{\overline{M}}(S)$ als Vereinigung aller $\overline{M}$-potenten Ideale von S zurückgehen und neben Satz 5.1 auch Lemma 3.1 heranziehen.

§ 6. Das Levitzkische Radikal in additiv kommutativen Halbringen mit Kern

6.1. Wie bereits angekündigt, definieren wir für additiv kommutative Halbringe S mit Kern K das *Levitzkische Radikal* $L(S)$ *von* S *als das größte lokal* $\bar{K}$*-potente Ideal* $L_{\bar{K}}(S)$, und analog das *klassische Radikal* $N(S)$ *von* S *als das von allen* $\bar{K}$*-potenten Rechtsidealen von* S *erzeugte Ideal* $N_{\bar{K}}(S)$ (vgl. Folgerung 3.4). Diese Begriffsbildungen korrespondieren zu der in der Ringtheorie üblichen, wenn man für Halbringe S mit Kern K „nilpotent" und „lokal nilpotent" durch „$\bar{K}$-potent" bzw. „lokal $\bar{K}$-potent" definiert. Wir bleiben hier aber bei der Sprechweise „(lokal) $\bar{K}$-potent", da wir die von uns auf $\bar{K}$ bezogenen Begriffsbildung mit den von Costa [6, 7] auf K bezogenen vergleichen wollen.

Besitzt S ein annullierendes Nullelement 0, so handelt es sich wegen $\{0\}=K=\bar{K}$ stets um (lokale) Nilpotenz im üblichen Sinne (vgl. 3.1), und $L(S)$ geht in das von Barbut [1] betrachtete Levitzkische Radikal über. Wir bemerken jedoch ohne Beweis, daß für Halbringe S mit (nicht annullierendem) Nullelement $0\notin K$ eintreten kann, dagegen stets $0\in\bar{K}$ gilt.

6.2. Hauptsatz. *Für das oben definierte Levitzkische Radikal* $L(S)$ *gilt:*

a) $L(A)=L(S)\cap A$ *für jedes Ideal* A *von* S, *also insbesondere* $L(L(S))=L(S)$.

b) $L(S)$ *ist* k*-abgeschlossen, also* $\overline{L(S)}=L(S)$.

c) *Der Restklassenhalbring von* S *nach* $L(S)$ *ist radikalfrei, also* $L(S/L(S))=\{0'\}$ *mit dem Nullelement* $0'$ *von* $S/L(S)$.

d) *Im folgenden sei* φ *ein Epimorphismus von* S *auf einen Halbring* S' *mit annullierendem Nullelement* $0'$.

1) *Ist* $\varphi^{-1}(0')\subseteqq L(S)$, *so bildet* φ *genau die Elemente von* $L(S)$ *auf* $L(S')$ *ab, d.h. es gilt*

$$\varphi^{-1}(L(S'))=L(S).$$

2) *Ist* S' *lokal nilpotent und das Ideal* $\varphi^{-1}(0')$ *von* S *lokal* $\bar{K}$*-potent, so ist der Halbring* S *lokal* $\bar{K}$*-potent*[8].

3) *Ist* $\varphi^{-1}(0')=L(S)$, *so hat* S' *keine lokal nilpotenten Ideale außer* $\{0'\}$.

Beweis. a) Gemäß 2.3 ist $K\cap A=K$ der Kern des Halbrings A, und der k-Abschluß von K in A ist ersichtlich der Durchschnitt des k-Abschlusses $\bar{K}$ von K in S mit A. Damit folgt aus Satz 4.1 a)

$$L(A)=L_{\bar{K}\cap A}(A)=L_{\bar{K}}(S)\cap A=L(S)\cap A.$$

b) Es gilt $\overline{L_{\bar{K}}(S)}=L_{\bar{K}}(S)$ nach Folgerung 5.2.

[8] Man beachte, daß „lokal $\bar{K}$-potent" für unsere Verallgemeinerung von „lokal nilpotent" steht.

d 1) Das Ideal $\varphi^{-1}(0')$ von S ist nach 2.5 k-abgeschlossen, also gilt $\bar{K} \subseteqq \varphi^{-1}(0') \subseteqq L_{\bar{K}}(S)$. Daraus folgt nach Folgerung 4.3 die Behauptung

$$L(S) = L_{\bar{K}}(S) = \varphi^{-1}(L_{0'}(S')) = \varphi^{-1}(L(S')).$$

d 2) Wegen $\varphi^{-1}(0') \subseteqq L(S)$ können wir d 1) mit $L(S') = S'$ anwenden, was $S = L(S)$ ergibt.

d 3) Aus d 1) folgt $\varphi^{-1}(L(S')) = L(S) = \varphi^{-1}(0')$, also $L(S') = \{0'\}$.

c) Für den natürlichen Epimorphismus φ von S auf $S' = S/L(S)$ gilt $\varphi^{-1}(0') = L(S)$ wegen b). Damit ist c) ein Spezialfall von d 3).

6.3. Wir wollen nun die Aussagen von 6.2 betrachten, wenn man sie anstelle von $L(S) = L_{\bar{K}}(S)$ für das in Costa [6, 7] vorgeschagene Radikal $L_K(S)$ formuliert, wobei natürlich alles auf K statt auf $\bar{K}$ zu beziehen ist. Dazu stellen wir zunächst fest:

Satz. *Folgende Aussagen sind für einen Halbring S mit Kern K gleichwertig:*

a) $L_K(S) = L_{\bar{K}}(S)$,

b) $\bar{K}$ *ist lokal K-potent, also* $\bar{K} \subseteqq L_K(S)$,

c) $L_K(S)$ *ist k-abgeschlossen.*

Bemerkung. Die triviale Implikation $K = \bar{K} \Rightarrow L_K(S) = L_{\bar{K}}(S)$ läßt sich nicht umkehren, wie man durch Gegenbeispiele zeigen kann.

Beweis. Ist b) erfüllt, so folgt aus $K \subseteqq \bar{K} \subseteqq L_K(S)$ gemäß Satz 4.1 b) bereits $L_K(S) = L_{\bar{K}}(S)$, also a). Aus a) folgt c), da $L_{\bar{K}}(S)$ k-abgeschlossen ist. Schließlich gilt mit c) wegen $K \subseteqq L_K(S)$ auch $\bar{K} \subseteqq \overline{L_K(S)} = L_K(S)$, also b).

6.4. Satz. *Für $L_K(S)$ als Radikal gilt die* 6.2 a) *entsprechende Aussage, nicht aber* 6.2 b) *und* 6.2 c). *Die* 6.2 d) *entsprechenden Aussagen für $L_K(S)$ sind zwar formal richtig, doch impliziert ihre Voraussetzung $\varphi^{-1}(0') \subseteqq L_K(S)$ bereits $L_K(S) = L_{\bar{K}}(S)$.*

Beweis. Für jedes Ideal A von S ist $K = K \cap A$ der Kern, und $L_{K \cap A}(A) = L_K(S) \cap A$ gilt nach Satz 4.1 a). Damit ist 6.2 a) für $L_K(S)$ richtig. Gegenbeispiele für b) und c) geben wir in § 7. Da $\varphi^{-1}(0')$ nach 2.5 ein k-abgeschlossenes Ideal von S ist, also $K \subseteqq \bar{K} \subseteqq \varphi^{-1}(0')$ gilt, impliziert $\varphi^{-1}(0') \subseteqq L_K(S)$ nach Satz 6.4 bereits $L_K(S) = L_{\bar{K}}(S)$.

6.5. Im Anschluß an 6.2 geben wir noch folgenden Zusammenhang mit dem in der Ringtheorie betrachteten Levitzkischen Radikal $L(R)$:

Satz. *Ist ein Halbring S mit Nullelement in einen Ring einbettbar, so gilt für den kleinsten Oberring R von S: $L(R) \cap S = L(S)$.*

Beweis. Gemäß [18] ist $R = D(S) = \{a - b \mid a, b \in S\}$ der Differenzenring von S und $D(L(S)) = \{a - b \mid a, b \in L(S)\}$ ein Ideal von R. Ersichtlich ist mit

$L(S)$ auch $D(L(S))$ lokal nilpotent, also gilt $L(S) \subseteqq D(L(S)) \subseteqq L(R)$. Umgekehrt ist $L(R) \cap S$ ein lokal nilpotentes Ideal von S, woraus $L(R) \cap S \subseteqq L(S)$ folgt.

6.6. Für das klassische Radikal $N(S) = N_{\bar K}(S)$ gelten die den Sätzen 6.2 und 6.5 entsprechenden Aussagen, falls man bei Satz 6.2 c) und d) noch voraussetzt, daß $N_{\bar K}(S)$ selbst $\bar K$-potent ist. Es scheint bemerkenswert, daß diese (ebenso einschneidende wie vom ringtheoretischen Standpunkt verständliche) Voraussetzung der $\bar K$-Potenz von $N_{\bar K}(S)$ nur benötigt wird, um $N_{\bar K}(S) = N_A(S)$ für jedes k-Ideal A von S mit $\bar K \subseteqq A \subseteqq N_{\bar K}(S)$ zu gewährleisten (vgl. § 4); nur dies geht über die Folgerung 4.3 in den Beweis von 6.2 d 1) ein.

Das Satz 6.3 entsprechende Kriterium für $N_K(S) = N_{\bar K}(S)$ gilt unter der zusätzlichen Voraussetzung, daß $N_K(S)$ selbst K-potent ist. Im allgemeinen ist es schwieriger, auf $N_K(S) = N_{\bar K}(S)$ zu schließen. Wie zu erwarten, ist $N_K(S)$ als Radikal weniger geeignet als $N_{\bar K}(S)$. Nach dem folgenden Beispiel braucht nämlich $N_K(S)$ nicht k-abgeschlossen zu sein, und der Restklassenhalbring $S/N_K(S)$ kann auch dann nichttriviale nilpotente Ideale enthalten, wenn $N_K(S)$ selbst K-potent ist.

§ 7. Ein Gegenbeispiel

7.1. In diesem Paragraphen konstruieren wir einen Halbring S mit folgenden Eigenschaften (vgl. 5.2, 6.4, 6.6):

a) S hat einen Kern K mit $K \subset \bar K$.

b) Es gilt $K = N_K(S) = L_K(S)$. Diese Ideale sind also K-potent, aber nicht k-abgeschlossen.

c) Es gilt $\bar K = \overline{N_K(S)} = \overline{L_K(S)} \subset N_{\bar K}(S) = L_{\bar K}(S)$; das klassische und das Levitzkische Radikal (in unserem Sinne) stimmen also überein und sind nicht mit K identisch. Sie sind jedoch $\bar K$-potent.

d) Der Restklassenhalbring $S' = S/K = S/N_K(S) = S/L_K(S) = S/\bar K$ enthält ein von seinem Nullelement $0'$ verschiedenes nilpotentes Ideal, und es gilt

$$N_{0'}(S/N_K(S)) = L_{0'}(S/L_K(S)) \neq (0').$$

e) Dagegen gilt im Einklang mit unseren Behauptungen

$$N_{0'}(S/N_{\bar K}(S)) = L_{0'}(S/L_{\bar K}(S)) = (0').$$

Es sei $P = \mathbb{Z}[x_1, x_2, \ldots]$ der Polynomring in den unendlich vielen nichtkommutativen Unbestimmten $x_1, x_2, \ldots$ über dem Ring $\mathbb{Z}$ der ganzen Zahlen und Q der Quotientenkörper des Polynomrings $\mathbb{Z}[k_1, k_2, \ldots]$ in den unendlich vielen kommutativen Unbestimmten $k_1, k_2, \ldots$. Mit $c(f) \in \mathbb{Z}$ bezeichnen wir die Summe der Koeffizienten des Polynoms

$f=f(x_1, x_2, \ldots)\in P$ und definieren von beiden Seiten eine Operatoranwendung von P auf Q gemäß

$$(1) \qquad f\kappa=\kappa f=c(f)\kappa \quad \text{für } \kappa\in Q.$$

Durch rein formale Rechnungen prüft man nach, daß durch (1) und

$$(2) \qquad \begin{aligned} (f,\kappa)+(g,\lambda)&=(f+g, \kappa+\lambda)\\ (f,\kappa)\cdot(g,\lambda)&=(f\cdot g, f\lambda+\kappa g+\kappa\cdot\lambda) \end{aligned}$$

auf der Menge $P\times Q$ ein Ring $R_1=P+Q$ definiert wird (es handelt sich um eine spezielle Everettsche Ringerweiterung, vgl. etwa [15], § 52). Unterscheiden wir die in P enthaltenen ganzen Zahlen als Vielfache des Einselementes e von P von den in Q enthaltenen rationalen Zahlen als Vielfache des Einselementes ε von Q, so können wir P bzw. Q mit den zu ihnen isomorphen Unterringen $\{(f,0)\mid f\in P\}$ bzw. $\{(0,\kappa)\mid\kappa\in Q\}$ von R_1 identifizieren und die Elemente von R_1 in der Form $(f,\kappa)=f+\kappa$ schreiben. Für die Multiplikation gilt dann gemäß (1) und (2)

$$(3) \qquad (f+\kappa)(g+\lambda)=fg+c(f)\lambda+c(g)\kappa+\kappa\lambda;$$

insbesondere ist $e=e+0$ Einselement von R_1.

Weiter sei F_0 der Unterhalbring von P, der aus allen $f\in P$ mit nichtnegativen Koeffizienten besteht und $F=F_0\setminus\{0\}$. Mit K_0 bezeichnen wir den Unterhalbkörper von Q, der aus allen Quotienten $\kappa\in Q$ besteht, deren Zähler und Nenner mit nichtnegativen Koeffizienten schreibbar sind; wir setzen $K=K_0\setminus\{0\}$. Man sieht leicht, daß

$$F_0+K_0=\{f+\kappa\mid f\in F_0,\ \kappa\in K_0\}$$

und

$$S_1=(F_0+K_0)\setminus\{0\}$$

Unterhalbringe von $R_1=P+Q$ sind. Der zu konstruierende Halbring S ist eine Kongruenzklassenstruktur von S_1, und zwar nach der kleinsten Halbringkongruenz $\mathscr{C}$ auf S_1, welche die Relationen

$$(4,1) \qquad x_1^2+\kappa_1 k_1\equiv\varepsilon+\kappa_1 k_1 \qquad \text{für alle } \kappa_1\in K$$

$$(4,\mathrm{i}) \qquad x_i x_1+\kappa_i k_i\equiv x_1 x_i+\kappa_i k_i \qquad \text{für alle } \kappa_i\in K;\ i=2,3,\ldots$$

enthält. Dabei verwenden wir $\equiv$ in diesem Paragraphen nur für die so definierte Kongruenz $\mathscr{C}$ auf S_1. Ersichtlich folgen aus diesen Relationen (4) keine (nichttrivialen) Relationen zwischen Elementen aus F, womit F als Unterhalbring von S aufgefaßt werden kann. Entsprechendes gilt für K; der durch $x_i\to\varepsilon$ für alle $i=1,2,\ldots$ und $\kappa\to\kappa$ für alle $\kappa\in K$ bestimmte Endomorphismus von S_1 bildet nämlich K identisch ab und führt die Relationen (4) in triviale Identitäten über.

7.3. In der Tat ist K als Unterhalbkörper und wegen (3) das kleinste Ideal von S. Weiter gilt a) wegen $x_1^2 \in \bar{K}$ nach (4.1), aber $x_1^2 \notin K$.

7.4. Als nächstes bestimmen wir die Struktur von $S' = S/K = S/\bar{K}$. Dazu betrachten wir den Restklassenring von $P = \mathbb{Z}[x_1, x_2, \ldots]$ modulo dem von x_1^2 und $x_i x_1 - x_1 x_i$, $i = 2, 3, \ldots$ erzeugten Ringideal I. Der von den Restklassen $[e]_I, [x_1]_I, [x_2]_I, \ldots$ erzeugte Unterhalbring U des Ringes P/I ist gemäß

$$[f+\kappa]_K \to [f]_I \tag{5}$$

homomorphes und damit isomorphes Bild von $S/K = S/\bar{K}$, wobei natürlich $[f+\kappa]_K$ die von $f+\kappa \in S$ repräsentierte Restklasse modulo K gemäß 2.5 bezeichnet. Zum Beweis dieser Behauptung geht man am besten von dem durch $f+\kappa \to [f]_I$, $\kappa \to [0]_I$ definierten Homomorphismus ψ von S_1 in P/I aus. Ersichtlich gilt $\psi(S_1) = U$ und ψ läßt sich zerlegen gemäß

$$S_1 \xrightarrow{\psi_1} S_1/\mathscr{C} = S \xrightarrow{\psi_2} S/K \xrightarrow{\psi_3} U,$$

wobei ψ_3 gerade (5) entspricht und bijektiv ist.

7.5. Für den Restklassenhalbring $S' = S/K = S/\bar{K}$ mit dem Nullelement $0' = \bar{K}$ gilt damit (vgl. auch Satz 6.5):

$$N_{0'}(S') = L_{0'}(S') = ([x_1]_K) + (0').$$

Weiterhin besteht $\bar{K} = \psi_2^{-1}(0')$ aus allen Elementen von S, die in der Form $f_2 + \kappa \in F_0 + K_0$ geschrieben werden können, wobei jeder Summand von f_2 (falls f_2 auftritt) wenigstens zweimal den Faktor x_1 enthält.

7.6. Es gilt $N_{\bar{K}}(S) = L_{\bar{K}}(S)$ und dieses Ideal besteht aus allen Elementen der Form $f_1 + \kappa \in S$, wobei jeder Summand von f_1 (falls f_1 auftritt) wenigstens einmal den Faktor x_1 enthält. Jedenfalls ist diese Menge A ein Ideal von S mit $A^2 \subseteqq \bar{K}$ gemäß 7.5, also gilt $A \subseteqq N_{\bar{K}}(S) \subseteqq L_{\bar{K}}(S)$. Andererseits ist kein Element $f + \kappa \in S \setminus A$ $\bar{K}$-potent. In $f + \kappa \equiv f_0 + f_1 + \kappa$ treten nämlich Summanden f_0 auf, die keinen Faktor x_1 enthalten, und für jede natürliche Zahl n gilt

$$(f_0 + f_1 + \kappa)^n \equiv f_0^n + f_1' + \kappa' \notin \bar{K}.$$

Daraus folgt $L_{\bar{K}}(S) \subseteqq A$. Natürlich ergibt sich 7.6 auch aus 7.5 unter Verwendung von Satz 6.2 d 1).

7.7. Wir zeigen schließlich, daß jedes Rechtsideal B von S, welches K-nil ist, mit K zusammenfällt. Daraus folgt $K = N_K(S) = L_K(S)$, also b), damit c) nach 7.6 sowie d) und e) nach 7.5.

Zum Beweis dieser Behauptung sei $f + \kappa \in F_0 + K_0$ ein Element aus B, welches nicht in K liegt. Wegen $K \subseteqq B \subseteqq L_{\bar{K}}(S)$ enthält jeder Summand von f wenigstens einmal den Faktor x_1, und aus $f + \kappa \notin K$ und der Form der

Relationen (4) folgt, daß $g_1+\kappa\notin K$ für wenigstens ein in f auftretendes Monom g_1 gelten muß. Wir schreiben $f=g_1+\cdots$ zur Andeutung eventuell vorhandener weiterer Summanden aus F und wählen x_t so, daß kein k_t in $\kappa\in K$ vorkommt. Wegen $b=(f+\kappa)\,x_t=g_1\,x_t+\cdots+\kappa\in B$ gibt es eine natürliche Zahl n mit

(6) $$b^n=(g_1\,x_t)^n+\cdots+\kappa^*\in K,$$

wobei κ^* alle auftretenden Summanden aus K bezeichnet. Aus $g_1+\kappa\notin K$ ergeben sich folgende Fälle:

(7) g_1 enthält genau einmal x_1.

(8) Zwischen je zwei der $r+1$ Faktoren x_1 von g_1 steht ein Faktor x_{i_ρ} mit $i_\rho\neq 1$ gemäß

$$g_1=\ldots x_1\ldots x_{i_1}\ldots x_1\ldots x_{i_2}\ldots\ldots x_{i_r}\ldots x_1\ldots,$$

ohne daß Summanden $\kappa_\rho\,k_{i_\rho}$ mit $\kappa_\rho\in K$ in κ auftreten;

(9) in g_1 *tritt* $\ldots x_1^2\ldots$ auf, aber kein Summand $\kappa_1\,k_1$ mit $\kappa_1\in K$ in κ.

Nun ist (7) unverträglich mit (6), da in $(g_1\,x_t)^n$ die Faktoren x_1 und x_t nur vertauschbar wären, wenn k_t in κ^* und damit in κ auftreten würde, was nach Wahl von t nicht gilt[9]. Für (8) zeigt das gleiche Argument, daß in $(g_1\,x_t)^n$ in (6) kein x_1 mit einem x_t vertauschbar ist. Zwei Faktoren x_1 müssen also innerhalb $g_1\,x_t$ zusammenkommen; da aber κ keinen der dafür benötigten Summanden der Form $\kappa_\rho\,k_{i_\rho}$ enthält, gilt das gleiche für κ^*, da κ^* aus Summen von Produkten von Summanden aus κ besteht. Entsprechend fehlt bei (9) ein Summand der Form $\kappa_1\,k_1$ in κ und damit in κ^*.

§ 8. Zusammenhang mit dem Jacobsonschen Radikal

8.1. Die folgenden Begriffsbildungen und Ergebnisse für additiv kommutative Halbringe S mit annullierendem Nullelement stammen von Bourne [2], Bourne-Zassenhaus [3] und Iizuki-Nakahara [11]. In der hier gegebenen Formulierung gelten sie auch für Halbringe ohne Nullelement, wie eine entsprechende Überprüfung der in [2], [3] und [11] gegebenen Beweise zeigt.

[9] Damit haben wir bereits $B\subseteqq\bar K$, also $K\subseteqq N_K(S)\subseteqq L_K(S)\subseteqq\bar K$ gezeigt. Daraus folgen schon alle unsere Behauptungen in 7.1 mit Ausnahme von b), und übrigens (vgl. Satz 6.3) auch $L_K(S)\subset\bar K$. Dabei haben wir bei allen Überlegungen bisher von den jeweils unendlich vielen Relationen (4, 1), (4, 2), (4, 3), ... nur je eine benötigt. Die damit gekennzeichneten Variationsmöglichkeiten unseres Beispiels beeinflussen nur, welche Elemente von K jeweils in $N_K(S)$ bzw. $L_K(S)$ liegen, sowie die Frage, ob $N_K(S)$ selbst K-potent ist. Für das Folgende machen wir in der Tat von allen Relationen (4, 1), (4, 2), ... Gebrauch.

Ein Rechtsideal A von S heißt *rechts semiregulär* [2] bzw. *rechts quasiregulär* [10, 3], wenn für beliebige Elemente $a_1, a_2 \in A$ Elemente $t_1, t_2 \in A$ bzw. $t_1, t_2 \in S$ existieren, so daß

$$a_1 + t_1 + a_1 t_1 + a_2 t_2 = a_2 + t_2 + a_1 t_2 + a_2 t_1$$

gilt. Letzteres ist nach [3] gleichwertig mit der Existenz von Elementen $t_1, t_2 \in A$ und $t \in S$ mit

$$a_1 + t_1 + a_1 t_1 + a_2 t_2 + t = a_2 + t_2 + a_1 t_2 + a_2 t_1 + t.$$

Die Summe (oder schon die Vereinigung) aller rechts semiregulären bzw. aller rechts quasiregulären Rechtsideale von S ist entweder leer oder ein rechts semireguläres Ideal $J(S)$ bzw. ein rechts quasireguläres Ideal $\sigma(S)$ von S. Beide stimmen mit der entsprechenden rechts-links-dualen Begriffsbildung überein; $J(S)$ ist das in [2] eingeführte *Jacobsonsche Radikal* von S, $\sigma(S)$ das in [3] zunächst auf andere Weise definierte Semiradikal von S. Dabei gilt nach Definition $J(S) \subseteqq \sigma(S)$. Die Übereinstimmung beider wurde erst in [11] gezeigt, und zwar gilt sogar: Ein rechts quasireguläres (zweiseitiges) Ideal von S ist auch rechts semiregulär. Gilt $J(S) \neq \varnothing$, so ist $J(S)$ nach [3] ein k-Ideal (und sogar ein h-Ideal, vgl. [12]), und der Restklassenhalbring $S' = S/J(S)$ mit $J(S) = 0'$ als annullierendem Nullelement erfüllt $J(S') = (0')$ [2].

Hat S ein annullierendes Nullelement 0, so ist $J(S)$ wegen $0 \in J(S)$ nicht leer. Hat S einen Kern K, gilt $K \subseteqq \bar{K} \subseteqq J(S) \Leftrightarrow J(S) \neq \varnothing$; es gibt aber Halbringe mit Kern K und $J(S) = \varnothing$ (etwa der Halbring S in § 7, doch findet man leicht einfachere Beispiele).

8.2. Satz. *Es sei S ein Halbring mit Kern K und Jacobsonschen Radikal $J(S) \neq \varnothing$. Dann gilt für die in § 6 definierten Radikale $N(S) \subseteqq L(S) \subseteqq J(S)$. Für $L(S) = J(S)$ hat $S/J(S)$ keine nichttrivialen lokal nilpotenten Ideale.*

Beweis. Ähnlich wie im Beweis von Theorem 7 in [2] zeigen wir zunächst, daß für beliebige Elemente $a_1, a_0 \in L(S)$ Elemente $t_0, t_1 \in L(S)$ und $k_0, k_1 \in \bar{K}$ existieren, so daß

$$(1) \qquad a_1 + t_0 + a_1 t_0 + a_0 t_1 + k_0 = a_0 + t_1 + a_1 t_1 + a_0 t_0 + k_1$$

gilt. Zu $A = \{a_0, a_1\}$ gibt es eine natürliche Zahl n mit $A^n \subseteqq \bar{K}$. Wir bilden dann in $L(S)$ die Elemente

$$t_\nu = \sum_{j=1}^{n-1} \Big(\sum_{\substack{i_1, \ldots, i_j = 0, 1 \\ \text{mit } 2 \mid i_1 + \cdots + i_j + \nu}} a_{i_1} a_{i_2} \cdots a_{i_j} \Big),$$

wobei also t_0 bzw. t_1 jeweils die Summe aller Produkte aus höchstens $n-1$ Faktoren a_0, a_1 ist, in denen a_1 in gerader bzw. ungerader Anzahl auftritt.

Ersichtlich gilt dann

$$a_1 + t_0 + a_1 t_0 + a_0 t_1 = t_1 + t_0 + k_1, \tag{2}$$

$$a_0 + t_1 + a_1 t_1 + a_0 t_0 = t_0 + t_1 + k_0, \tag{3}$$

wobei k_0 bzw. k_1 jeweils die Summe aller Produkte aus n Faktoren a_0, a_1 ist, in denen a_1 in gerader bzw. ungerader Anzahl auftritt. Damit gilt $k_0, k_1 \in A^n \subseteqq \bar{K}$, und die Addition von k_0 bzw. k_1 zu (2) bzw. (3) zeigt (1).

Es sei nun φ der natürliche Epimorphismus von S auf $S' = S/J(S)$. Wie oben bemerkt, gilt $J(S') = (0')$. Für beliebige Elemente $\varphi(a_1) = a'_1$, $\varphi(a_0) = a'_0$ des Ideals $\varphi(L(S))$ von S' folgt aus (1) wegen $\bar{K} \subseteqq J(S)$

$$a'_1 + t'_0 + a'_1 t'_0 + a'_0 t'_1 = a'_0 + t'_1 + a'_1 t'_1 + a'_0 t'_0$$

mit Elementen $t'_\nu = \varphi(t_\nu) \in \varphi(L(S))$. Damit ist das Ideal $\varphi(L(S))$ rechts semiregulär, also gilt $\varphi(L(S)) \subseteqq J(S') = (0')$, woraus

$$L(S) \subseteqq \varphi^{-1}(0') = \overline{J(S)} = J(S)$$

folgt. Die letzte Behauptung entspricht Satz 6.2 d 3).

Literatur

1. Barbut, E.: On nil semirings with ascending chain conditions. Fundamenta Math. **58**, 261 – 264 (1970).
2. Bourne, S.: The Jacobson radical of a semiring. Proc. Nat. Acad. Sci. USA **37**, 163 – 170 (1951).
3. Bourne, S., Zassenhaus, H.: On the semiradical of a semiring. Proc. Nat. Acad. Sci. USA **44**, 907 – 914 (1958).
4. Clifford, A.H., Preston, B.G.: The algebraic theory of semigroups I, II, Amer. Math. Soc., Providence, R.I., 1961 | 1967.
5. Cohn, P.M.: Universal algebra. New York-London: Harper & Row 1965.
6. Costa, A.: Sur la théorie générale des demi-anneaux, Publ. Math. Debrecen **10**, 14 – 29 (1963).
7. Costa, A.: Sur la théorie générale des demi-anneaux I. Seminaire Dubreil-Pisot, Paris, 1961, exposés n° 24.
8. Henriksen, M.: Ideals in semirings with commutative addition. Notices Amer. Math. Soc. **6**, 321 (1958).
9. Herstein, I.N.: Topics in ring theory. Chicago Lectures in Math., 1969.
10. Iizuki, K.: On the Jacobson radical of a semiring. Tôhoku Math. J. (2), **11**, 409 – 421 (1959).
11. Iizuki, K., Nakahara, I.: A note on the semiradical of a semiring. Kumatoto J. Sci., Ser. A, **4**, 1 – 3 (1959).
12. La Torre, D.R.: On h-ideals and k-ideals in hemirings. Publ. Math. Debrecen **12**, 219 – 226 (1965).
13. Levitzki, J.: On the radical of a general ring. Bull. Amer. Math. Soc. **49**, 462 – 466 (1943).
14. Levitzki, J.: Semi-nilpotent ideals. Duke Math. J. **10**, 553 – 556 (1943).
15. Rédei, L.: Algebra I. Leipzig: Geest & Portig 1959.

16. Schwarz, Š.: Zur Theorie der Halbgruppen. Sb. Praci Prirod., Univ. Bratislawa No. 6 (1953).
17. Weinert, H.J.: Über Halbringe und Halbkörper, I. Acta Math. Acad. Sci. Hungar. **13**, 365–378 (1962).
18. Weinert, H.J.: Über Halbringe und Halbkörper, II. Acta Math. Acad. Sci. Hungar. **14**, 209–227 (1963).
19. Weinert, H.J.: Referat zu [1], Math. Reviews **42**, 3126 (1971).

Prof. H.J. Weinert
Institut für Mathematik
Technische Universität Clausthal
D-3392 Clausthal-Zellerfeld
Adolf-Römer-Str. 2A
Bundesrepublik Deutschland

(Eingegangen am 10. Juli 1972)

Math. Z. 128, 343–361 (1972)

Gaußsche Wahrscheinlichkeitsmaße auf Corwinschen Gruppen

Herbert Heyer und Christian Rall

1. Einleitung

Gaußsche Wahrscheinlichkeitsmaße (Gauß-Maße) auf abelschen lokalkompakten Gruppen wurden im Zusammenhang mit dem zentralen Grenzwertproblem erstmalig systematisch untersucht von Parthasarathy, Rao und Varadhan, deren Studien inzwischen in das Buch [4] von Parthasarathy Eingang gefunden haben.

Mehr oder weniger befriedigende Versuche zur Definition von Gauß-Maßen auf nicht notwendigerweise abelschen kompakten Gruppen sind in [3] dargelegt. Die erste Definition des Gauß-Maßes auf einer beliebigen lokalkompakten Gruppe stammt von Siebert, in dessen Arbeit [7] auch der Zusammenhang zwischen der neuen Definition und den bisher bekannten vollständig geklärt wird.

Dabei ist in [7] das Gauß-Maß als einbettbares Wahrscheinlichkeitsmaß (W-Maß) definiert, welches erst im abelschen Fall eine wahrscheinlichkeitstheoretische Deutung zuläßt. Diese wiederum verdankt man den erstgenannten Autoren.

In der vorliegenden Abhandlung wird der Reihe der Definitionen von Gauß-Maßen auf abelschen lokalkompakten Gruppen mit abzählbarer Basis der Topologie (LKAB-Gruppen) eine weitere hinzugefügt, welche im Spezialfall der Gruppe der reellen Zahlen mit der durch den Satz von Bernstein [6] nahegelegten übereinstimmt.

Für Maße beliebigen Vorzeichens wurde die neue Definition von Corwin in seinen Arbeiten [1] eingeführt.

Schränkt man diese Definition auf W-Maße ein, so ergibt sich für gewisse Klassen von abelschen LKAB-Gruppen ein vollständiger Vergleich mit der Definition in [4].

In genauerer Formulierung erhält man (Satz 6.6), daß auf gewissen Corwinschen Gruppen (vgl. 2), zu denen die stark Corwinschen Gruppen, aber auch der nicht stark Corwinsche Torus $\mathbb{T}$ gehören, jedes Gauß-Maß im Sinne von Corwin dargestellt werden kann als Faltung eines idempotenten Gauß-Maßes mit einem Gauß-Maß im Sinne von Parthasarathy [4].

Insbesondere erhält man (Korollar 6.7) für Corwinsche Gruppen mit zusammenhängender Charaktergruppe, also speziell für die Vektorgruppen $\mathbb{R}^p$ $(p \geqq 1)$, daß Gauß-Maße im Sinne von Corwin und Gauß-Maße im Sinne von Parthasarathy zusammenfallen.

Als vorbereitendes Resultat mit eigener Bedeutung ergibt sich eine Charakterisierung der idempotenten Gauß-Maße im Sinne von Corwin (Satz 5.1).

In einem abschließenden Abschnitt wird das Problem der Darstellung von Gauß-Maßen im Sinne von Corwin für separable Hilberträume behandelt. Satz 8.3 ist das Analogon zu Satz 6.6.

In eine andere Richtung zielende die Fälle der LKAB-Gruppen und des separablen Hilbertraumes umfassende Studien zu Gauß-Maßen im Sinne von Corwin nehmen ihren Ausgang mit der Arbeit [5] von Rao und wurden neuerdings von Schmidt in [8] weitergeführt.

2. Vorbereitungen

Es sei G eine (additiv geschriebene) abelsche LKAB-Gruppe mit neutralem Element 0. Mit $\mathcal{M}^b(G)$ bzw. $\mathcal{M}^1(G)$ bezeichnen wir die vag-topologischen Faltungshalbgruppen der beschränkten Borel-Maße bzw. der (Borelschen) W-Maße auf (der Borelschen σ-Algebra $\mathfrak{B} := \mathfrak{B}(G)$ von) G, mit $\mathscr{E}(G)$ die Menge der Dirac-Maße $\varepsilon_x (x \in G)$ in $\mathcal{M}^1(G)$. Von besonderem Interesse ist ferner die Unterhalbgruppe $\mathscr{I}(G)$ der (bez. Faltung) idempotenten Elemente von $\mathcal{M}^1(G)$. Bekanntlich ist $\mathscr{I}(G)$ gerade die Menge der von kompakten Untergruppen H von G getragenen H-invarianten Maße in $\mathcal{M}^1(G)$, welche auch die normierten Haar-Maße von G genannt werden.

Darüber hinaus ist $\mathscr{I}(G)$ eine Unterhalbgruppe der Halbgruppe $\mathscr{U}_0(G)$ der unendlich teilbaren Maße in $\mathcal{M}^1(G)$. Dabei heißt $\mu \in \mathcal{M}^1(G)$ unendlich teilbar, wenn es zu jedem $n \geqq 1$ ein $\mu_n \in \mathcal{M}^1(G)$ gibt mit $\mu_n^n = \mu$.

Da $\mathscr{U}_0(G)$ im allgemeinen keine (vag-)abgeschlossene Halbgruppe ist, wird den Betrachtungen in [4] der Begriff des schwach unendlich teilbaren Maßes in $\mathcal{M}^1(G)$ zugrundegelegt. $\mu \in \mathcal{M}^1(G)$ heißt dabei schwach unendlich teilbar, wenn es zu jedem $n \geqq 1$ ein $\mu_n \in \mathcal{M}^1(G)$ und ein $x_n \in G$ gibt, so daß $\mu = \mu_n^n * \varepsilon_{x_n}$ erfüllt ist.

Die Menge $\mathscr{U}(G)$ der schwach unendlich teilbaren Maße in $\mathcal{M}^1(G)$ ist in der Tat eine abgeschlossene Unterhalbgruppe von $\mathcal{M}^1(G)$.

Details über die Halbgruppen $\mathcal{M}^1(G)$ und $\mathscr{U}(G)$ sind in [3] und [4] dargestellt.

Für jedes $\mu \in \mathcal{M}^1(G)$ bezeichnet $\tilde{\mu}$ das zu μ adjungierte Maß. $\mu \in \mathcal{M}^1(G)$ heißt symmetrisch, wenn $\mu = \tilde{\mu}$ gilt. Für jedes $\mu \in \mathcal{M}^1(G)$ werde der Träger von μ mit T_μ bezeichnet.

Es sei nunmehr G^* die Charaktergruppe von G mit neutralem Element (Einheitscharakter) χ_0.

Zu jedem $\mu \in \mathscr{M}^1(G)$ ist die Fouriertransformierte $\hat{\mu}$ als komplexwertige Funktion auf G^* definiert durch

$$\hat{\mu}(\chi) := \int \chi(x)\, \mu(dx)$$

für alle $\chi \in G^*$. Die Abbildung $\mu \to \hat{\mu}$ (Fouriertransformation) ist bekanntlich eine multiplikative bistetige Injektion von $\mathscr{M}^1(G)$ in die Menge $\mathscr{C}^b(G^*)$ der stetigen beschränkten komplexwertigen Funktionen auf G^*.

Die im Zusammenhang mit der Fouriertransformation benötigten grundlegenden Tatsachen der harmonischen Analyse entnehme der Leser der Monographie [2].

Es sei nun ξ die durch

$$\xi(x, y) := (x + y, x - y)$$

für alle $x, y \in G$ definierte Abbildung von $G \times G$ in sich.

Definition. Ein Maß $\mu \in \mathscr{M}^1(G)$ heißt *(C-)Gaußsch* (Gaußsch im Sinne von Corwin), wenn die Gleichung

$$\xi(\mu \otimes \mu) = (\mu * \mu) \otimes (\mu * \mu^{\sim})$$

erfüllt ist.

Die Gesamtheit der C-Gauß-Maße auf G werde durch $\mathscr{G}_C(G)$, die Teilmenge der symmetrischen Maße in $\mathscr{G}_C(G)$ durch $\mathscr{G}_C^s(G)$ abgekürzt.

Unmittelbar aus der Definition ergibt sich das folgende

Lemma 2.1. *Für jedes $\mu \in \mathscr{M}^1(G)$ mit $\varphi := \hat{\mu}$ sind nachstehende Aussagen äquivalent:*

(i) $\mu \in \mathscr{G}_C(G)$,

(ii) $\varphi(\chi + \psi)\, \varphi(\chi - \psi) = \varphi(\chi)^2\, \varphi(\psi)\, \varphi(-\psi)$ *für alle* $\chi, \psi \in G^*$.

Beweis. Durch Anwendung des Transformationssatzes für Integrale und des Satzes von Fubini erhält man die folgenden für alle $\chi, \psi \in G^*$ gültigen Gleichungsketten:

$$\begin{aligned}
\widehat{\xi(\mu \otimes \mu)}(\chi, \psi) &= \iint (\chi, \psi)(x, y)\, \xi(\mu \otimes \mu)(d(x, y)) \\
&= \iint (\chi, \psi)(\xi(x, y))(\mu \otimes \mu)(d(x, y)) \\
&= \iint (\chi, \psi)(x + y, x - y)(\mu \otimes \mu)(d(x, y)) \\
&= \iint \chi(x + y)\, \psi(x - y)(\mu \otimes \mu)(d(x, y)) \\
&= \iint (\chi + \psi)(x)(\chi - \psi)(y)(\mu \otimes \mu)(d(x, y)) \\
&= \int (\chi + \psi)(x)\, \mu(dx) \int (\chi - \psi)(y)\, \mu(dy) \\
&= \hat{\mu}\, (\chi + \psi)\, \hat{\mu}(\chi - \psi) = \varphi(\chi + \psi)\, \varphi(\chi - \psi)
\end{aligned}$$

sowie

$$\widehat{(\mu * \mu) \otimes (\mu * \mu^{\sim})}(\chi, \psi) = \widehat{\mu * \mu}(\chi)\, \widehat{\mu * \mu^{\sim}}(\psi)$$
$$= \hat{\mu}(\chi)^2 \hat{\mu}(\psi)\, \hat{\mu}(-\psi) = \varphi(\chi)^2 \varphi(\psi)\, \varphi(-\psi).$$

Korollar 2.2. *Die Klassen $\mathscr{G}_C(G)$ und $\mathscr{G}_C^s(G)$ sind abgeschlossene Unterhalbgruppen von $\mathscr{M}^1(G)$, und es gilt $\mathscr{E}(G) \subset \mathscr{G}_C(G)$.*

Die im folgenden zu definierende Klasse von abelschen LKAB-Gruppen ist für die Darstellung der C-Gaußschen Maße von Wichtigkeit.

Mit ζ bzw. ζ^* werden die durch $\zeta(x) := 2x$ für alle $x \in G$ bzw. $\zeta^*(\chi) := 2\chi$ für alle $\chi \in G^*$ definierten Homomorphismen ζ bzw. ζ^* von G bzw. G^* in sich bezeichnet.

Definition. Eine abelsche LKAB-Gruppe G heißt *Corwinsche Gruppe*, wenn die Abbildung ζ von G in sich ein Epimorphismus ist.

Eine Corwinsche Gruppe G heißt *stark Corwinsch*, wenn ζ zudem ein Automorphismus ist.

Eigenschaften. 1. ζ ist genau dann ein Monomorphismus, wenn $\overline{\zeta^*(G^*)} = G^*$ erfüllt ist.

[Ist nämlich ζ ein Monomorphismus und gilt $H := \overline{\zeta^*(G^*)} \neq G^*$, so hat man $H^{\perp} = (G^*/H)^* \neq \{0\}$. Für jedes $y \in H^{\perp} \setminus \{0\}$ gilt $\psi(y) = 1$, sofern $\psi \in H$. Insbesondere ist $2\chi(y) = \chi(2y) = 1$ für alle $\chi \in G^*$, also $2y = 0$ im Widerspruch zur Injektivität von ζ.

Ist umgekehrt $\overline{\zeta^*(G^*)} = G^*$ erfüllt, so ergibt sich für $x \in G$ mit $2x = 0$ offenbar $2\chi(x) = \chi(2x) = 0$, sofern $\chi \in G^*$. Da die Elemente von G^* die Punkte von G trennen, folgt somit $x = 0$.]

2. Sind die Homomorphismen ζ und ζ^* beide surjektiv, so sind sie auch bijektiv.

[Dies ergibt sich aus 1. mittels des Dualitätssatzes von Pontrjagin.]

3. Mit G ist G^* stark Corwinsch.

4. Das Beispiel des Torus $\mathbb{T}$ zeigt, daß die Surjektivität von ζ und von ζ^* im allgemeinen unabhängig voneinander sind.

5. In der Klasse der abelschen lokalkompakten Gruppen gelten folgende Abgeschlossenheitseigenschaften:

(i) Mit $(G_\alpha)_{\alpha \in A}$ sind auch die Produkte $\prod\limits_{\alpha \in A} G_\alpha$ bzw. $\prod\limits_{\alpha \in A}^* G_\alpha$ Corwinsch bzw. stark Corwinsch.

(ii) Für jede Familie $(G_\alpha)_{\alpha \in A}$ stark Corwinscher Gruppen G_α ist auch der projektive Limes $\varprojlim\limits_{\alpha \in A} G_\alpha$ stark Corwinsch.

6. Beispiele für nicht Corwinsche Gruppen sind die endlichen Gruppen, aber auch die Gruppe $\mathbb{Z}$ der ganzen Zahlen.

Der folgende in [1], I, S. 416 bewiesene Satz liefert einen Einblick in die Struktur der stark Corwinschen Gruppen.

Satz 2.3. *Zu jeder stark Corwinschen Gruppe G existieren natürliche Zahlen $p, q \geqq 0$ sowie eine abelsche LKAB-Gruppe K, welche eine stark Corwinsche kompakte, offene Untergruppe enthält, so daß gilt:*

$$G \cong \mathbb{R}^p \times \mathbb{Q}_2^q \times K.$$

3. Gauß-Maße im Sinne von Bernstein

Die in 2. angegebene Definition des Gauß-Maßes auf einer LKAB-Gruppe G läßt sich wahrscheinlichkeitstheoretisch deuten.

Ausgangspunkt für eine Definition in dieser Richtung ist der bekannte Satz von Bernstein [6], S. 355.

Definition. Ein Maß $\mu \in \mathcal{M}^1(G)$ heißt *B-Gaußsches Maß* (Gauß-Maß im Sinne von Bernstein), wenn es einen Wahrscheinlichkeitsraum (W-Raum) $(\Omega, \mathfrak{A}, P)$ sowie unabhängige G-Zufallsvariable X und Y auf $(\Omega, \mathfrak{A}, P)$ gibt mit den Eigenschaften

(i) $X(P) = Y(P) = \mu$.

(ii) $X + Y$ und $X - Y$ sind unabhängig.

Mit $\mathcal{G}_B(G)$ werde die Gesamtheit der B-Gauß-Maße auf G bezeichnet.

Satz 3.1. *Für jede abelsche LKAB-Gruppe G gilt*

$$\mathcal{G}_B(G) = \mathcal{G}_C(G).$$

Beweis. 1) Wir zeigen: $\mathcal{G}_C(G) \subset \mathcal{G}_B(G)$. Es seien $\mu \in \mathcal{M}^1(G)$ sowie $\nu_1, \nu_2 \in \mathcal{M}^1(G)$ mit $\xi(\mu \otimes \mu) = \nu_1 \otimes \nu_2$. Definiert man einen W-Raum $(\Omega, \mathfrak{A}, P)$ durch die Festsetzungen $\Omega := G \times G$, $\mathfrak{A} := \mathfrak{B} \otimes \mathfrak{B}$, $P := \mu \otimes \mu$ und G-Zufallsvariable X und Y auf $(\Omega, \mathfrak{A}, P)$ durch $X := \mathrm{pr}_{G \times .}$ und $Y := \mathrm{pr}_{. \times G}$ bzw., so sind X und Y unabhängig, und es gilt $X(P) = Y(P) = \mu$.

Nach Voraussetzung hat man offenbar

$$\big((X+Y) \otimes (X-Y)\big)(P) = \xi(X \otimes Y)(P) = \xi(\mu \otimes \mu) = \nu_1 \otimes \nu_2,$$

nach Konstruktion des Produktmaßes also

$$(X+Y)(P) = \nu_1 \quad \text{sowie} \quad (X-Y)(P) = \nu_2.$$

Wegen der Unabhängigkeit von X und Y gilt somit

$$\nu_1 = \mu * \mu \quad \text{und} \quad \nu_2 = \mu * \mu^{\sim},$$

also

$$\big((X+Y) \otimes (X-Y)\big)(P) = (X+Y)(P) \otimes (X-Y)(P),$$

womit die Unabhängigkeit von $X + Y$ und $X - Y$ nachgewiesen ist.

2) Zu zeigen: $\mathscr{G}_B(G)\subset\mathscr{G}_C(G)$: Nach Voraussetzung für die in der Definition des B-Gauß-Maßes auftretenden G-Zufallsvariablen X und Y auf $(\Omega, \mathfrak{A}, P)$ gilt

$$\big((X+Y)\otimes(X-Y)\big)(P)=(X+Y)(P)\otimes(X-Y)(P)$$

sowie

$$(X\otimes Y)(P)=X(P)\otimes Y(P).$$

Da nach Definition von ξ die Relation

$$\xi\circ(X\otimes Y)=(X+Y)\otimes(X-Y)$$

erfüllt ist, erhält man

$$\xi(\mu\otimes\mu)=\xi\big(X(P)\otimes Y(P)\big)=\big(\xi\circ(X\otimes Y)\big)(P)=(X+Y)(P)\otimes(X-Y)(P)$$

und wegen der Unabhängigkeit von X und $-Y$ schließlich

$$\xi(\mu\otimes\mu)=(\mu*\mu)\otimes(\mu*\mu^{\sim}),$$

also die Behauptung.

4. Gauß-Maße im Sinne von Parthasarathy

In [4], S. 97 wird die folgende auf Parthasarathy, Rao und Varadhan zurückgehende Definition angegeben.

Definition. Ein Maß $\mu\in\mathscr{M}^1(G)$ heißt *P-Gaußsch* (Gaußsch im Sinne von Parthasarathy), wenn gilt:

(i) $\mu\in\mathscr{U}(G)$.

(ii) Wenn immer $\mu=\exp(\nu)*\lambda$ mit $\nu\in\mathscr{M}^b(G)$ und $\lambda\in\mathscr{U}(G)$, so folgt $\nu=\rho\varepsilon_0$ für ein $\rho\in\mathbb{R}_+$.

Die Klasse der P-Gaußschen Maße in $\mathscr{M}^1(G)$ werde mit $\mathscr{G}_P(G)$ bezeichnet, die Teilklasse der symmetrischen Elemente von $\mathscr{G}_P(G)$ mit $\mathscr{G}_P^s(G)$.

Offenbar besitzt ein $\mu\in\mathscr{G}_P(G)$ keinen idempotenten Faktor (s. [3], S. 193).

Es sei $\mathscr{C}(G, \mathbb{R})$ die Menge der stetigen reellen Funktionen auf G.

Wir setzen nunmehr

$$Q(G):=\{\varphi\in\mathscr{C}(G,\mathbb{R}): \varphi(x+y)+\varphi(x-y)=2[\varphi(x)+\varphi(y)] \text{ für alle } x, y\in G\}$$

sowie

$$Q_+(G):=\{\varphi\in Q(G): \varphi\geqq 0\}.$$

Dann ergibt sich der folgende in [4], S. 97 formulierte

Satz 4.1. *Für jede abelsche LKAB-Gruppe G und jede komplexwertige Funktion φ auf G^* sind folgende Aussagen äquivalent:*

(i) *Es existiert ein* $\mu \in \mathscr{G}_P(G)$ *mit* $\varphi = \hat{\mu}$

(ii) *Es existieren* $x_0 \in G$ *und* $\Phi \in Q_+(G^*)$ *mit* $\varphi(\chi) = \chi(x_0)\exp(-\Phi(\chi))$ *für alle* $\chi \in G^*$.

Korollar 4.2. $\mathscr{G}_P(G)$ *ist eine abgeschlossene Unterhalbgruppe von* $\mathscr{U}(G)$.

Beweis. Zu zeigen bleibt die Abgeschlossenheit von $\mathscr{G}_P(G)$ in $\mathscr{M}^1(G)$. Es sei also $(\mu_n)_{n \geq 1}$ eine Folge in $\mathscr{G}_P(G)$ mit $\lim_{n \to \infty} \mu_n = \mu \in \mathscr{M}^1(G)$. Nach dem Satz existieren (eindeutig bestimmte) Folgen $(x_n)_{n \geq 1}$ in G und $(\Phi_n)_{n \geq 1}$ in $Q_+(G^*)$ mit $\hat{\mu}_n = \hat{\varepsilon}_{x_n} \exp(-\Phi_n)$ für alle $n \geq 1$, so daß sich wegen der Bistetigkeit der Fouriertransformation für alle $\chi \in G^*$ unmittelbar

$$\log|\hat{\mu}(\chi)| = \lim_{n \to \infty} \log|\hat{\mu}_n(\chi)| = \lim_{n \to \infty} (-\Phi_n(\chi))$$

ergibt.

Die Funktion $\Phi := -\log|\hat{\mu}|$ ist stetig und definitionsgemäß ein Element von $Q_+(G^*)$.

Setzt man noch für alle $\chi \in G^*$:

$$\Psi(\chi) := \frac{\hat{\mu}(\chi)}{|\hat{\mu}(\chi)|} = \lim_{n \to \infty} \frac{\hat{\mu}_n(\chi)}{|\hat{\mu}_n(\chi)|} = \lim_{n \to \infty} \chi(x_n),$$

so ergibt sich $\Psi \in (G^*)^*$. Nach dem Satz von Pontrjagin existiert also ein $x_0 \in G$ mit

$$\Psi(\chi) = \chi(x_0) \qquad \text{für alle } \chi \in G^*$$

und damit die Darstellung

$$\hat{\mu} = \hat{\varepsilon}_x \exp(-\Phi),$$

so daß erneute Anwendung des Satzes die Behauptung liefert.

In [4], S. 101 wird ferner gezeigt, daß für jedes $\mu \in \mathscr{G}_P(G)$ mit $\hat{\mu} = \exp(-\Phi)$ für ein $\Phi \in Q_+(G^*)$ die Inklusion $T_\mu \subset G_0$ gilt.

Hieraus ergibt sich zunächst

$$\mathscr{G}_P(G) = \mathscr{E}(G) * \mathscr{G}_P(G_0)$$

und schließlich als Antwort auf die Frage nach der Existenz P-Gaußscher Maße in $\mathscr{M}^1(G)$:

$$\mathscr{G}_P(G) \setminus \mathscr{E}(G) \neq \varnothing$$

genau dann, wenn $G_0 \neq \{0\}$. Hierbei bezeichnet G_0 die Zusammenhangskomponente von 0 in G.

Lemma 4.3. *Es seien* G *eine abelsche LKAB-Gruppe,* H *eine offene Untergruppe von* G *und* $\Phi_0 \in Q_+(H)$.

Dann existiert ein $\Phi \in Q_+(G)$ *mit* $\Phi_0 = \mathrm{Res}_H \Phi$.

Beweis (unter Heranziehung einer Idee in [4], S. 106).

Da H offene Untergruppe von G ist und G eine abzählbare Basis der Topologie besitzt, ist G/H abzählbar. Es existiert also eine Folge $(x_j)_{j \geqq 1}$ in G mit

$$(x_j+H) \cap (x_k+H) = \varnothing \qquad \text{für alle } j \neq k \ (j, k \geqq 1)$$

und

$$G = \bigcup_{j \geqq 1} (x_j+H).$$

(a) Es sei $H_1 := x_1 \mathbb{Z} + H = \bigcup_{m \in \mathbb{Z}} (x_1 m + H)$ die von x_1 und H erzeugte offene Untergruppe von G.

1. Fall. Es gelte $x_1 \mathbb{Z} \cap H = \varnothing$.

In diesem Fall definiert man eine reelle Funktion Φ_1 auf H_1 durch

$$\Phi_1(x) := \Phi_1(x_0 + m x_1) := \Phi_0(x_0)$$

für alle $x = x_0 + m x_1 \in H_1$ mit $x_0 \in H$ und $m \in \mathbb{Z}$. Φ_1 ist eine nichtnegative stetige Fortsetzung von Φ_0, und es gilt

$$\Phi_1(x+y) + \Phi_1(x-y) = 2[\Phi_1(x) + \Phi_1(y)]$$

für alle $x, y \in G$.

2. Fall. Es gelte $x_1 \mathbb{Z} \cap H \neq \varnothing$.

Dann existiert eine kleinste ganze Zahl $r_0 > 0$ mit $r_0 x_1 \in H$. Für beliebiges $x \in H_1$ ist $r_0 x \in H_1$.

Hat nämlich $x \in H_1$ die Darstellung $x = x_0 + m x_1$ mit $x_0 \in H$ und $m \in \mathbb{Z}$, so ist wegen $r_0 x = r_0 x_0 + m r_0 x_1$ mit $r_0 x_1$ auch $r_0 x \in H$.

Nun definiert man eine reelle Funktion Φ_1 auf H_1 durch

$$\Phi_1(x) := \frac{1}{r_0^2} \Phi_0(r_0 x)$$

für alle $x \in H_1$.

Wegen $\Phi_0(m y) = m^2 \Phi_0(y)$ für alle $m \in \mathbb{Z}$ und $y \in H$ gilt für alle $x \in H$:

$$\Phi_1(x) = \frac{1}{r_0^2} \Phi_0(r_0 x) = \frac{r_0^2}{r_0^2} \Phi_0(x) = \Phi_0(x),$$

und es ist Φ_1 eine nichtnegative stetige Fortsetzung von Φ_0 auf H_1.

Für $x, y \in H_1$ erhält man

$$\begin{aligned}
\Phi_1(x+y) + \Phi_1(x-y) &= \frac{1}{r_0^2} [\Phi_0(r_0(x+y)) + \Phi_0(r_0(x-y))] \\
&= \frac{1}{r_0^2} [\Phi_0(r_0 x + r_0 y) + \Phi_0(r_0 x - r_0 y)] \\
&= \frac{2}{r_0^2} [\Phi_0(r_0 x) + \Phi_0(r_0 y)] = 2[\Phi_1(x) + \Phi_1(y)],
\end{aligned}$$

also ist $\Phi_1 \in Q_+(H_1)$.

(b) Es bleibt die Aufgabe, die gesuchte Funktion $\Phi \in Q_+(G)$ zu konstruieren, indem man sie durch Funktionen $\Phi_n \in Q_+(H_n)$ für offene Untergruppen H_n von G $(n \geqq 1)$ zusammensetzt: Der Beweis erfolgt durch vollständige Induktion: Für $n=1$ ist der Beweis in (a) erbracht.

Es seien für jedes $n \geqq 1$ eine offene Untergruppe H_n von G mit $H_n \supset H_{n-1}$ sowie eine Funktion $\Phi_n \in Q_+(H_n)$ mit $\mathrm{Res}_{H_{n-1}} \Phi_n = \Phi_{n-1}$ bereits konstruiert.

Mit $(x_j)_{j \geqq 1}$ werde die zu Beginn des Beweises gewählte Folge in G bezeichnet:

Ist $x_j \in H_n$ für alle $j \geqq 1$, so folgt wegen $H_n \supset H$ sofort

$$G = \bigcup_{j \geqq 1} (x_j + H_n)$$

und daher $G = H_n$.

In diesem Fall setzt man $H_{n+1} := H_n = G$ und $\Phi_{n+1} := \Phi_n$. Im anderen Fall kann o. B. d. A. $x_{n+1} \notin H_n$ vorausgesetzt und daher $H_{n+1} := x_{n+1} \mathbb{Z} + H_n$ sowie analog zu (a) Φ_n zu einem $\Phi_{n+1} \in Q_+(H_{n+1})$ fortgesetzt werden.

Wegen $G = \bigcup_{j \geqq 1} (x_j + H)$ erhält man aus $x_{n+1} + H \subset H_{n+1}$ für alle $n \geqq 1$ sofort

$$G = \bigcup_{n \geqq 1} H_n,$$

und man definiert die gewünschte reelle Funktion Φ auf G durch

$$\Phi(x) := \Phi_n(x)$$

für alle $x \in G$ mit $x \in H_n$.

Nach Konstruktion ist Φ wohldefiniert. Da H_n eine offene Untergruppe von G ist für jedes $n \geqq 1$, ist Φ zunächst stetig. Da Φ_n die $Q_+(G)$ definierende Funktionalgleichung erfüllt für jedes $n \geqq 1$, ist $\Phi \in Q_+(G)$. Schließlich folgt nach Konstruktion auch $\Phi_0 = \mathrm{Res}_H \Phi$.

Lemma 4.4. *Es seien G eine abelsche LKAB-Gruppe, H eine kompakte Untergruppe von G und p die kanonische Abbildung von G auf G/H. Dann existiert zu jedem Maß $\nu \in \mathscr{G}_P(G/H)$ mit $\hat{\nu} = \exp(-\Phi_0)$ für ein $\Phi_0 \in Q_+(H^\perp)$ ein Maß $\mu \in \mathscr{G}_P^s(G)$ mit $p(\mu) = \nu$.*

Beweis. Bekanntlich gelten die Relationen $(G/H)^* \cong H^\perp$ sowie $(G^*/H^\perp)^* \cong H$.

Da H nach Voraussetzung kompakt ist, ist $G^*/H^\perp$ diskret, und $H^\perp$ ist eine offene Untergruppe von G^*.

Φ_0 kann damit nach Lemma 4.3 zu einer Funktion $\Phi \in Q_+(G^*)$ fortgesetzt werden. Dann existiert aber ein Maß $\mu \in \mathscr{G}_P^s(G)$ mit $\hat{\mu} = \exp(-\Phi)$.

Aus $\widehat{p(\mu)} = \mathrm{Res}_{H^\perp} \hat{\mu} = \exp(-\Phi_0) = \hat{\nu}$ folgt die Behauptung.

5. Idempotente Gauß-Maße

Eine einfache Rechnung zeigt, daß für jede abelsche LKAB-Gruppe G die Inklusion

$$\mathscr{G}_P(G) \subset \mathscr{G}_C(G)$$

erfüllt ist.

Es sei nämlich $\mu \in \mathscr{G}_P(G)$. Dann existieren nach Satz 4.1 ein $x_0 \in G$ und ein $\Phi \in Q_+(G^*)$ mit der Eigenschaft:

$$\hat{\mu}(\chi) = \chi(x_0) \exp(-\Phi(\chi))$$

für alle $\chi \in G^*$.

Für alle $\chi, \psi \in G^*$ ergibt sich hieraus mit $\varphi := \hat{\mu}$ die Gleichungskette

$$\begin{aligned}
\varphi(\chi+\psi)\,\varphi(\chi-\psi) &= (\chi+\psi)(x_0)\exp(-\Phi(\chi+\psi))\,(\chi-\psi)(x_0)\exp(-\Phi(\chi-\psi)) \\
&= \chi(x_0)^2 \exp\{-2[\Phi(\chi)+\Phi(\psi)]\} \\
&= \chi(x_0)^2 \{\exp(-\Phi(\chi))\}^2 \{\exp(-\Phi(\psi))\}^2 \\
&= \varphi(\chi)^2 \exp(-\Phi(\psi)) \exp(-\Phi(-\psi)) \\
&= \varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi),
\end{aligned}$$

also $\mu \in \mathscr{G}_C(G)$.

Satz 5.1 (Charakterisierung der idempotenten Gauß-Maße). *Es seien G eine abelsche LKAB-Gruppe, H eine kompakte Untergruppe von G sowie w_H das normierte Haar-Maß auf H. Dann sind folgende Aussagen äquivalent:*

(i) $w_H \in \mathscr{G}_C(G)$,

(ii) *Für alle* $\chi \in \complement H^\perp$ *ist* $2\chi \in \complement H^\perp$.

Beweis. 1) (i)$\Rightarrow$(ii) Es sei w_H ein Maß in $\mathscr{G}_C(G)$ mit Fouriertransformierter $\varphi := \hat{w}_H$.

Offenbar ist $\varphi = 1_{H^\perp}$, und es gilt nach Lemma 2.1 für alle $\chi, \psi \in G^*$:

$$\varphi(\chi+\psi)\,\varphi(\chi-\psi) = \varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi).$$

Ist nun $\chi \in \complement H^\perp$, so erhalten wir aus dieser Gleichung mit $\psi := \chi$ sofort:

$$\varphi(2\chi) = \varphi(\chi+\chi)\,\varphi(\chi-\chi) = \varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi) = 0$$

und damit $2\chi \in \complement H^\perp$.

2) (ii)$\Rightarrow$(i) Da $H^\perp$ eine Untergruppe von G^* ist, folgt mit $\chi, \psi \in H^\perp$ auch $\chi+\psi$, $\chi-\psi \in H^\perp$, es gilt also für $\varphi := \hat{w}_H = 1_{H^\perp}$ und alle $\chi, \psi \in H^\perp$:

$$1 = \varphi(\chi+\psi)\,\varphi(\chi-\psi) = \varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi).$$

Ist $\chi \in H^\perp$ und $\psi \in \complement H^\perp$, so ist $\chi+\psi \notin H^\perp$, und man hat

$$\varphi(\chi+\psi)\,\varphi(\chi-\psi)=0=\varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi).$$

Ist aber $\chi \in \complement H^\perp$ und $\psi \in \complement H^\perp$, so folgt $\chi+\psi \in \complement H^\perp$ oder $\chi-\psi \in \complement H^\perp$:

Wären nämlich $\chi+\psi, \chi-\psi \in H^\perp$, so gälte auch $\chi+\psi+\chi-\psi=2\chi \in H^\perp$ im Widerspruch zur Voraussetzung. Also ergibt sich auch in diesem Fall:

$$\varphi(\chi+\psi)\,\varphi(\chi-\psi)=\varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi),$$

und somit ist $w_H \in \mathscr{G}_C(G)$.

Korollar 5.2. *Für abelsche LKAB-Gruppen G ist im allgemeinen*

$$\mathscr{G}_C(G) \not\subset \mathscr{G}_P(G).$$

Beweis. Wir setzen $G:=\mathbb{T}$. Dann gibt es neben G eine Folge $(H_n)_{n\geqq 1}$ von nichttrivialen kompakten Untergruppen H_n von G mit der Eigenschaft (ii) des Satzes: man setze

$$H_n:=\mathbb{T}_n:=\left\{\exp\left(i\frac{2\pi l}{2n+1}\right):\ l=0,1,\ldots,n\right\}$$

für jedes $n\geqq 1$. Die Maße w_G und w_{H_n} für $n\geqq 1$ sind in $\mathscr{G}_C(G)$, aber nicht in $\mathscr{G}_P(G)$, da Elemente aus $\mathscr{G}_P(G)$ keine idempotenten Faktoren besitzen.

6. Darstellung von Gauß-Maßen auf Corwinschen Gruppen

In diesem Hauptteil der Untersuchung soll der volle Vergleich der Klassen $\mathscr{G}_P(G)$ und $\mathscr{G}_C(G)$ durchgeführt werden. Zunächst werde der Zusammenhang zwischen $\mathscr{G}_C(G)$ und $\mathscr{G}_C^s(G)$ geklärt:

Satz 6.1. *Es sei G eine Corwinsche Gruppe. Dann gilt*

$$\mathscr{G}_C(G)=\mathscr{E}(G)*\mathscr{G}_C^s(G).$$

Beweis. Offenbar hat man $\mathscr{E}(G)\subset\mathscr{G}_C(G)$. Ferner ist nach Korollar 2.2 $\mathscr{G}_C(G)$ eine Unterhalbgruppe von $\mathscr{M}^1(G)$ und damit

$$\mathscr{E}(G)*\mathscr{G}_C^s(G)\subset\mathscr{G}_C(G).$$

Zu zeigen bleibt die umgekehrte Inklusion. Für $\mu\in\mathscr{G}_C(G)$ mit $\varphi:=\hat{\mu}$ gilt stets

$$\varphi(\chi+\psi)\,\varphi(\chi-\psi)=\varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi)$$

für alle $\chi, \psi\in G^*$.

Die Menge $A:=A_\varphi:=\{\chi\in G^*:\ \varphi(\chi)\neq 0\}$ ist eine offene Untergruppe von G^*:

Denn einerseits ist A offen, da φ stetig ist. Wegen $\varphi(\chi_0)=1$ ist $\chi_0\in A$, und wegen der für alle $\chi, \psi\in A$ geltenden Relation

$$\varphi(\chi+\psi)\,\varphi(\chi-\psi)=\varphi(\chi)^2\,\varphi(\psi)\,\varphi(-\psi)=\varphi(\chi)^2\,|\varphi(\psi)|^2\neq 0$$

sind auch $\chi+\psi$ und $\chi-\psi\in A$, womit A auch als Untergruppe von G^* nachgewiesen ist.

Wir definieren nun eine Abbildung P_0: $A\to\mathbb{T}$ durch

$$P_0(\chi):=\frac{\varphi(\chi)}{\varphi(-\chi)}=\frac{\varphi(\chi)}{\overline{\varphi(\chi)}}$$

für alle $\chi\in A$.

Eine einfache Rechnung liefert

$$P_0(\chi+\psi)=P_0(\chi)\,P_0(\psi)$$

für alle $\chi,\psi\in G^*$, d.h. P_0 ist ein stetiger Charakter von A, welcher zu einem stetigen Charakter P auf ganz G^* fortgesetzt werden kann.

Nach dem Dualitätssatz von Pontrjagin existiert nun ein $s\in G$ mit $s(\chi)=P(\chi)$ für alle $\chi\in A$. Da G Corwinsch und daher die Abbildung $x\to 2x$ von G in sich eine Surjektion ist, existiert zu s ein $y\in G$ mit $2y=s$.

Wir setzen $x:=-y$.

Dann ergibt sich, daß $\nu:=\varepsilon_x*\mu\in\mathscr{G}_C^s(G)$ ist: Da ε_x und μ offenbar zu $\mathscr{G}_C(G)$ gehören, bleibt zu zeigen, daß ν symmetrisch ist.

Wir setzen $\Phi:=\hat{\nu}$.

Dann gilt $\Phi(\chi)=\chi(x)\,\varphi(\chi)$ für alle $\chi\in G^*$. Mit χ gehört $-\chi$ zu $\complement A$, da A eine Untergruppe von G^* ist, es gilt also für $\chi\in\complement A$ offenbar

$$\Phi(\chi)=\Phi(-\chi)=0.$$

Für $\chi\in A$ hingegen erhält man

$$\begin{aligned}\hat{\nu}(\chi)&=\Phi(\chi)=\chi(x)\,\varphi(\chi)=\chi(-y)\,\varphi(\chi)=P_0(\chi)\,\frac{\varphi(-\chi)}{\chi(y)}\\&=\frac{\chi(2y)\,\varphi(-\chi)}{\chi(y)}=\frac{\chi(y)^2\,\varphi(-\chi)}{\chi(y)}=\chi(y)\,\varphi(-\chi)\\&=\chi(-x)\,\varphi(-\chi)=(-\chi)(x)\,\varphi(-\chi)=\Phi(-\chi)=\widehat{\tilde{\nu}}(\chi)\end{aligned}$$

und damit $\hat{\nu}=\widehat{\tilde{\nu}}$, d.h. ν ist symmetrisch.

Es gilt somit $\mu=\varepsilon_{-x}*\nu$ mit $\nu\in\mathscr{G}_C^s(G)$ und daher

$$\mathscr{G}_C(G)\subset\mathscr{E}(G)*\mathscr{G}_C^s(G).$$

Dem folgenden Satz stellen wir drei Lemmata voran:

Lemma 6.2. *Es seien G eine abelsche LKAB-Gruppe und $\mu\in\mathscr{G}_C^s(G)$ mit $\varphi:=\hat{\mu}\geqq 0$. Wir setzen $A:=A_\varphi:=\{\chi\in G^*\colon \varphi(\chi)\neq 0\}$ sowie $H:=A^\perp$. Dann existiert ein $\nu\in\mathscr{G}_P^s(G)$, so daß $\mu=w_H*\nu$ gilt mit $w_H\in\mathscr{G}_C(G)$.*

Beweis. Bekanntlich hat man $(G/H)^*\cong A$. Nach dem Satz von Bochner existiert zu $\sigma:=\mathrm{Res}_A\,\varphi$ ein Maß $\nu_0\in\mathscr{M}^1(G/H)$ mit $\hat{\nu}_0=\sigma$. Nach Definition von σ gilt nun $0<\sigma(\chi)\leqq 1$ für alle $\chi\in A$. Die Abbildung Φ: $A\to\mathbb{R}_+$ definiert durch

$$\Phi(\chi) := -\log\sigma(\chi)$$

für alle $\chi\in A$ ist stetig und genügt der Funktionalgleichung

$$\Phi(\chi+\psi)+\Phi(\chi-\psi)=2[\Phi(\chi)+\Phi(\psi)]$$

für alle $\chi,\psi\in A$. Es ist also $\Phi\in Q_+(A)$ mit

$$\hat{\nu}_0=\exp(-\Phi).$$

Nach Lemma 4.4 existiert ein Maß $\nu\in\mathscr{G}_P^s(G)$ mit $\mathrm{Res}_A\,\hat{\nu}=\hat{\nu}_0$.

Da A eine offene Untergruppe von G^* ist, ist H eine kompakte Untergruppe von G. Wir weisen nun für H Bedingung (ii) von Satz 5.1 nach. Hierzu sei $\chi\in G^*$ mit $2\chi\in A$. Unter Berücksichtigung der Symmetrie von μ folgt aus der Funktionalgleichung (ii) von Lemma 2.1:

$$0\neq\varphi(2\chi)=(\varphi(\chi))^4.$$

Es ist also $\chi\in A$ und damit $w_H\in\mathscr{G}_C(G)$.

Nach Konstruktion von ν erhält man somit wegen $\hat{w}_H=1_A$ schließlich $\hat{\mu}=\hat{w}_H\cdot\hat{\nu}$ und damit $\mu=w_H*\nu$ wie behauptet.

Lemma 6.3. *Es seien G eine abelsche LKAB-Gruppe mit Charaktergruppe G^*, so daß der Index von $\overline{2G^*}$ höchstens 2 ist, und Φ eine auf G^* definierte stetige reelle Funktion mit den folgenden Eigenschaften:*

(i) $\Phi(G^*)\subset\{-1,1\}$ *und* $\Phi(\chi_0)=1$,

(ii) Φ *ist auf den Nebenklassen nach der Untergruppe* $\overline{2G^*}$ *konstant.*

Dann ist Φ ein Charakter von G^* (und damit nach dem Satz von Pontrjagin ein Element von G).

Beweis. Im Falle $\overline{2G^*}=G^*$ ist Φ offenbar der Einheitscharakter auf G^*.

Im Falle $\overline{2G^*}\neq G^*$ kann o.B.d.A. $\Phi(\complement\,\overline{2G^*})=\{-1\}$ angenommen werden. Da einerseits $\complement\,\overline{2G^*}+\complement\,\overline{2G^*}\subset\overline{2G^*}$ gilt und andererseits $\overline{2G^*}$ eine Untergruppe von G^* ist, hat man auch in diesem Fall $\Phi(\chi+\psi)=\Phi(\chi)\,\Phi(\psi)$ für alle $\chi,\psi\in G^*$ und damit $\Phi\in(G^*)^*\cong G$.

Bemerkung. Die Bedingung im Lemma über den Index von $\overline{2G^*}$ kann nicht ersatzlos fortgelassen werden, wie das Beispiel der Gruppe $G:=\mathbb{Z}(2)^2$ zeigt.

Lemma 6.4. *Es seien G eine abelsche LKAB-Gruppe mit Charaktergruppe G^*, so daß der Index von $\overline{2G^*}$ höchstens 2 ist, $\mu\in\mathscr{G}_C^s(G)$ mit $\varphi:=\hat{\mu}$ und $A:=A_\varphi:=\{\chi\in G^*:\varphi(\chi)\neq 0\}$.*

Dann ist der Index von $\overline{2A}$ in A ebenfalls höchstens 2.

Beweis. In der üblichen Weise erschließt man aus

$$\varphi(2\chi)=\varphi(\chi)^4$$

für alle $\chi\in G^*$, daß $\chi\in A$ genau dann gilt, wenn $2\chi\in A$ vorliegt.

Wir zeigen:

$$\overline{2G^*}\cap A=\overline{2A}.$$

Da A offene und damit abgeschlossene Untergruppe von G^* ist, gilt $\overline{2A}\subset A$ und somit $\overline{2A}\subset\overline{2G^*}\cap A$.

Umgekehrt existiert zu jedem $\chi\in\overline{2G^*}\cap A$ eine Folge $(\chi_n)_{n\geqq 1}$ in $2G^*$ mit $\lim\limits_{n\to\infty}\chi_n=\chi\in A$.

Für jedes $n\geqq 1$ gilt $\chi_n=2\psi_n$ mit $\psi_n\in G^*$. Wegen der Offenheit von A kann o.B.d.A. $\chi_n\in A$ für alle $n\geqq 1$ angenommen werden.

Für jedes $n\geqq 1$ ist $\psi_n\in A$ und daher auch $\chi_n=2\psi_n\in 2A$. Also ist $\lim\limits_{n\to\infty}\chi_n=\chi\in\overline{2A}$ und damit insgesamt

$$\overline{2G^*}\cap A\subset\overline{2A}.$$

Sind nunmehr $\chi,\psi\in A$ mit $\chi,\psi\notin\overline{2A}$ vorgegeben, so gilt nach Voraussetzung $\chi\in\psi+\overline{2G^*}$ und wegen $\chi-\psi\in A$ schließlich $\chi\in\psi+\overline{2G^*}\cap A=\psi+\overline{2A}$, also die Behauptung.

Satz 6.5. *Es seien G eine abelsche LKAB-Gruppe mit Charaktergruppe G^*, so daß der Index von $\overline{2G^*}$ höchstens 2 ist. Setzt man noch*

$$\mathscr{I}_C(G):=\mathscr{I}(G)\cap\mathscr{G}_C(G),$$

so gilt

$$\mathscr{G}_C^s(G)=\mathscr{I}_C(G)*\mathscr{G}_P^s(G).$$

Beweis. Nach der einführenden Bemerkung in 5. ist noch die Inklusion

$$\mathscr{G}_C^s(G)\subset\mathscr{I}_C(G)*\mathscr{G}_P^s(G)$$

zu zeigen.

Es sei also $\mu\in\mathscr{G}_C^s(G)$. Anwendung von Lemma 6.2 auf das Maß $\mu*\mu\in\mathscr{G}_C^s(G)$ mit $\varphi_0:=\widehat{\mu*\mu}\geqq 0$, $A:=A_{\varphi_0}$ sowie $H:=A^\perp$ liefert die Existenz eines Maßes $\nu\in\mathscr{G}_P^s(G)$, so daß

$$\mu*\mu\doteq w_H*\nu$$

gilt mit $w_H\in\mathscr{I}_C(G)$, d.h. die Existenz einer Funktion $\Phi\in Q_+(G^*)$, so daß

$$\varphi_0=1_A\exp(-\Phi)$$

erfüllt ist.

Zu zeigen ist die Existenz eines $x\in G$ mit $2x=0$, so daß für $\varphi:=\hat{\mu}$ die Darstellung

$$\varphi=1_A\exp\left(-\frac{\Phi}{2}\right)\hat{\varepsilon}_x$$

vorliegt.

Es sei Ψ diejenige Abbildung von A in $\mathbb{C}$, für welche

$$\varphi = \exp\left(-\frac{\Phi}{2}\right) \cdot \Psi$$

gilt. Da μ symmetrisch ist, ist Ψ nur der Werte 1 und -1 fähig, und es gilt $\Psi(\chi_0) = 1$.

Eine einfache Rechnung ergibt

$$\Psi(\chi+\psi)\,\Psi(\chi-\psi) = (\Psi(\chi)\,\Psi(\psi))^2 = 1$$

für alle $\chi, \psi \in A$.

Für $\chi = \chi_1 + 2\psi \in A$ (mit $\chi_1, \psi \in A$) gilt

$$\Psi(\chi) = \Psi(\chi_1 + \psi + \psi) = \Psi(\chi_1 + \psi - \psi) = \Psi(\chi_1).$$

Wegen der Stetigkeit von Ψ ist Ψ auf den Nebenklassen von A nach $\overline{2A}$ konstant. Mittels Lemma 6.4 folgt aus Lemma 6.3: $\Psi \in A^*$.

Ψ läßt sich wie folgt zu einer stetigen reellen Funktion $\tilde{\Psi}$ auf G^* fortsetzen:

$$\tilde{\Psi}(\chi) := 1 \qquad \text{für alle } \chi \in G^*, \text{ falls } \Psi \equiv 1$$

und

$$\tilde{\Psi}(\chi) := \begin{cases} 1 & \text{für alle } \chi \in \overline{2G^*} \\ -1 & \text{für alle } \chi \notin \overline{2G^*} \end{cases}, \quad \text{falls } \Psi \not\equiv 1.$$

Damit ist $\tilde{\Psi}$ ein Charakter auf G^*, es existiert also nach dem Satz von Pontrjagin ein $x \in G$ mit

$$\hat{\varepsilon}_x = \tilde{\Psi}.$$

Ferner gilt

$$\varphi = \hat{\mu} = 1_A \exp\left(-\frac{\Phi}{2}\right) \hat{\varepsilon}_x$$

und

$$\hat{\varepsilon}_x^2 = 1, \quad \text{d.h.} \quad 2x = 0.$$

Die Funktion $\exp\left(-\frac{\Phi}{2}\right) \hat{\varepsilon}_x$ ist also Fouriertransformierte eines Maßes $\lambda \in \mathscr{G}_P^s(G)$, d.h.

$$\mu = w_H * \lambda$$

mit $w_H \in \mathscr{I}_C(G)$ und $\lambda \in \mathscr{G}_P^s(G)$.

Satz 6.6. *Es sei G eine Corwinsche Gruppe mit Charaktergruppe G^*, so daß der Index von $\overline{2G^*}$ höchstens 2 ist. Dann gilt*

$$\mathscr{G}_C(G) = \mathscr{I}_C(G) * \mathscr{G}_P(G).$$

Beweis ergibt sich unmittelbar aus Satz 6.5 zusammen mit Satz 6.1.

Korollar 6.7. *Ist G^* zusammenhängend, so gilt*

$$\mathscr{G}_C(G) = \mathscr{G}_P(G).$$

Beweis. Falls G^* zusammenhängend ist, besitzt G keine nichttriviale kompakte Untergruppe, es ist also $\mathscr{I}_C(G) = \{\varepsilon_0\}$.

Insbesondere hat man:

Korollar 6.8. *Für jedes* $p \geqq 1$ gilt

$$\mathscr{G}_C(\mathbb{R}^p) = \mathscr{G}_P(\mathbb{R}^p).$$

Korollar 6.9. *Für jede stark Corwinsche Gruppe G gilt*

$$\mathscr{G}_C(G) = \mathscr{I}_C(G) * \mathscr{G}_P(G).$$

Beweis folgt aus der für stark Corwinsche Gruppen gültigen Gleichung $2G^* = G^*$.

Das folgende Resultat zeigt, daß die Aussage des Satzes auch für nicht notwendigerweise stark Corwinsche Gruppen wahr ist.

Korollar 6.10. *Für den Torus* $\mathbb{T}$ *gilt*

$$\mathscr{G}_C(\mathbb{T}) = \mathscr{I}_C(\mathbb{T}) * \mathscr{G}_P(\mathbb{T}).$$

Beweis. $\mathbb{T}$ ist Corwinsch, und es gilt

$$\mathbb{T}^*/\overline{2\mathbb{T}^*} \cong \mathbb{Z}/2\mathbb{Z}.$$

Der Index von $\overline{2\mathbb{T}^*}$ in $\mathbb{T}^*$ ist somit 2, so daß die Voraussetzungen des Satzes erfüllt sind.

Bemerkung. Nach Satz 5.1 ist

$$\mathscr{I}_C(\mathbb{T}) = \{\varepsilon_1, w_{\mathbb{T}}\} \cup \{w_{\mathbb{T}_n} : n \geqq 1\}$$

mit

$$w_{\mathbb{T}_n} = \frac{1}{2n+1} \sum_{k=0}^{2n} \varepsilon_{a_k},$$

wobei $a_k := \exp\left(\frac{2\pi i k}{2n+1}\right)$ für $0 \leqq k \leqq 2n$ $(n \geqq 1)$.

7. Darstellung von Gauß-Maßen im Sinne von Corwin auf einem separablen Hilbertraum

Es sei G eine polnische Gruppe, das ist eine topologische Gruppe mit abzählbarer Basis der Topologie, welche vollständig metrisierbar ist. Für G werden analog zum lokalkompakten Fall mit abzählbarer Basis die schwach topologischen (Faltungs-)Halbgruppen $\mathscr{M}^b(G)$ bzw. $\mathscr{M}^1(G)$ der beschränkten Borel-Maße bzw. der (Borelschen) W-Maße auf G erklärt, so daß insbesondere die Teilmengen $\mathscr{E}(G)$, $\mathscr{U}_0(G)$ und $\mathscr{U}(G)$ der Dirac-Maße, der unendlich und der schwach unendlich teilbaren Maße von $\mathscr{M}^1(G)$ sowie die Teilmengen $\mathscr{G}_P(G)$ und $\mathscr{G}_C(G)$ der Gauß-Maße im Sinne von Parthasarathy und im Sinne von Corwin innerhalb von $\mathscr{M}^1(G)$ definiert sind.

Nunmehr sei $G := E$ ein separabler Hilbertraum mit innerem Produkt $\langle\cdot,\cdot\rangle$.

Für jedes $\mu\in\mathscr{M}^b(E)$ ist das charakteristische Funktional $\hat{\mu}$ als komplexwertige Funktion auf E erklärt mit den Eigenschaften der Fouriertransformierten eines Maßes auf einer abelschen LKAB-Gruppe (vgl. [4], S. 152).

Insbesondere erhält man eine zum Fall abelscher LKAB-Gruppen analoge Charakterisierung der Klasse $\mathscr{G}_P(E)$ ([4], S. 179–181):

Für eine komplexwertige Funktion φ auf E gilt genau dann $\varphi=\hat{\mu}$ für ein Maß $\mu\in\mathscr{G}_P(E)$, wenn es ein $x_0\in E$ und einen S-Operator A auf E gibt, so daß

$$\varphi(y)=\exp(i\langle x_0, y\rangle-\tfrac{1}{2}\langle Ay, y\rangle)$$

für alle $y\in E$ erfüllt ist.

Durch Modifikation der Methode der harmonischen Analyse gewinnt man die Hauptresultate über die Darstellung der Maße von $\mathscr{G}_C(G)$ nunmehr in vereinfachter Form.

Zunächst ist offenbar $\mathscr{G}_P(E)\subset\mathscr{G}_C(E)$.

Satz 7.1. *Für einen separablen Hilbertraum E gilt*

$$\mathscr{G}_C^s(E)\subset G_P^s(E).$$

Beweis. Es seien $\mu\in\mathscr{G}_C^s(E)$ und $\varphi:=\hat{\mu}$, so daß eine zum Beweis von Lemma 2.1 analoge Rechnung die für alle $x, y\in E$ gültige Funktionalgleichung

$$\varphi(x+y)\,\varphi(x-y)=\varphi(x)^2\,\varphi(y)^2 \tag{$*$}$$

liefert.

Mittels vollständiger Induktion erhält man sodann

$$\varphi(rx)=\varphi(x)^{r^2} \quad \text{für alle } r\in\mathbb{Q},\ x\in E$$

und wegen der Stetigkeit von φ schließlich

$$\varphi(tx)=\varphi(x)^{t^2} \quad \text{für alle } t\in\mathbb{R},\ x\in E.$$

Ferner ist φ reell, so daß man gemäß ($*$) auf $\varphi\geqq 0$ schließen kann.

Der nichtleere Unterraum $H_\varphi:=\{x\in E\colon \varphi(x)\neq 0\}$ ist offen und abgeschlossen. Da E zusammenhängend ist, gilt also $H_\varphi=E$ und daher $0<\varphi(x)\leqq 1$ für alle $x\in E$.

Setzt man für alle $x \in E$:

$$\Phi(x) := -\log \varphi(x),$$

so erhält man eine stetige Funktion $\Phi \geqq 0$ auf E mit den folgenden Eigenschaften:

(i) $\Phi(x+y)+\Phi(x-y)=2[\Phi(x)+\Phi(y)]$

(ii) $\Phi(t\,x)=t^2\,\Phi(x)$

für alle $x, y \in E$, $t \in \mathbb{R}$.

Φ ist somit eine stetige positiv definite quadratische Form auf E, welche sich darstellen läßt in der Form

$$\Phi(x)=\langle A\,x, y\rangle \quad \text{für alle } x \in E,$$

wobei A ein hermitescher positiv-definiter Operator auf E ist. Damit ist

$$\varphi(x)=\exp(-\langle A\,x, x\rangle) \quad \text{für alle } x \in E$$

und A nach [4], S. 164 ein S-Operator auf E.

Nach oben zitiertem Charakterisierungssatz für $\mathscr{G}_P(E)$ ergibt sich die Behauptung.

Satz 7.2. *Unter den Voraussetzungen von Satz* 7.1 *gilt*

$$\mathscr{G}_C(E)=\mathscr{E}(E) * \mathscr{G}_C^s(E).$$

Beweis. Zu zeigen ist die Inklusion $\mathscr{G}_C(E) \subset \mathscr{E}(E) * \mathscr{G}_C^s(E)$. Es sei also $\mu \in \mathscr{G}_C(E)$ mit $\varphi := \hat{\mu}$. Wie im Beweis von Satz 7.1 erhält man sodann $H_\varphi := \{x \in E : \varphi(x) \neq 0\} = E$.

Definieren eine Abbildung $\Phi: E \to \mathbb{T}$ durch

$$\Phi(x) := \frac{\varphi(x)}{\varphi(-x)} \quad \text{für alle } x \in E.$$

Nach [4], S. 160 ist Φ stetig bez. der S-Topologie. Man zeigt ferner, daß für alle $x, y \in E$ die Gleichung

$$\Phi(x+y)=\Phi(x)\,\Phi(y)$$

gilt.

Die Abbildung $\Psi: E \to \mathbb{T}$ definiert durch

$$\Psi(x)=\Phi\left(-\frac{x}{2}\right) \quad \text{für alle } x \in E$$

erfüllt die für alle $x, y \in E$ gültige Gleichung

$$\Psi(x+y)=\Psi(x)\,\Psi(y).$$

Offenbar ist dann Ψ positiv definit, und es gilt

$$\Psi(0)=1.$$

Nach [4], S. 160 existiert ein Maß $\lambda \in \mathscr{M}^1(E)$ mit $\hat{\lambda}=\Psi$. Weiterhin hat man wegen $\Psi\bar{\Psi}=1$ noch $\lambda * \lambda^{\sim}=\varepsilon_0$, also die Existenz eines $x_0 \in E$ mit $\lambda=\varepsilon_{x_0}$.

Mit μ ist auch $\nu := \mu * \varepsilon_{x_0}$ in $\mathscr{G}_C(E)$.

Eine kurze Rechnung liefert $\nu=\nu^{\sim}$, so daß mit $\mu=\varepsilon_{-x_0} * \nu$ die Behauptung erwiesen ist.

Als unmittelbare Folgerung der Sätze 7.1 und 7.2 ergibt sich

Satz 7.3. $\mathscr{G}_C(E)=\mathscr{G}_P(E)$.

Damit ist das für separable Hilberträume gültige Analogon von Satz 6.6 gewonnen.

Literatur

1. Corwin, L.: Generalized Gaussian measure and a functional equation I. J. Functional Analysis **5**, 412–427 (1970); II. J. Functional Analysis **6**, 481–505 (1970); III. Advances Math. **6**, 239–251 (1971).
2. Hewitt, E., Ross, K.A.: Abstract harmonic analysis I. Berlin-Heidelberg-New York: Springer 1963.
3. Heyer, H.: Infinitely divisible probability measures on compact groups. Springer Lecture Notes **247**, 55–249 (1972).
4. Parthasarathy, K.R.: Probability measures in metric spaces. New York: Academic Press 1967.
5. Rao, C.R.: Characterization of the distribution of a random variable in linear structural relations. Sankya A **28**, 251–260 (1966).
6. Richter, H.: Wahrscheinlichkeitstheorie. Berlin-Heidelberg-New York: Springer 1956.
7. Siebert, E.: Stetige Halbgruppen von Wahrscheinlichkeitsmaßen auf lokalkompakten maximal fastperiodischen Gruppen. Eingereicht bei Z. Wahrscheinlichkeitstheorie verw. Geb.
8. Schmidt, K.: On a characterization of certain infinitely divisible positive definite functions and measures. Erscheint in J. London Math. Soc. (1972).

Prof. H. Heyer und C. Rall
Mathematisches Institut der Universität
D-7400 Tübingen, Hölderlinstraße 19
Bundesrepublik Deutschland

(Eingegangen am 27. Juli 1972)

Math. Z. 128, 363 – 367 (1972)

Eine Bemerkung über quadratische Formen über einem lokalen Ring der Charakteristik 2

Ricardo Baeza

0. Bezeichnungen und Definitionen

Sei A ein lokaler Ring mit maximalem Ideal $\mathfrak{m}$ und (M, q) eine quadratische Form über A. q heißt nicht entartet, wenn die zugehörige bilineare Form B_q nicht entartet ist, d.h. B_q induziert einen Isomorphismus $M \xrightarrow{\sim} M^*$ durch $y \mapsto B_q(-, y)$. Von jetzt an sei q nicht entartet vorausgesetzt. Ist $2 \in A^*$, so besitzt q eine orthogonale Basis. Ist $2 \in \mathfrak{m}$, so hat M gerade Dimension, und q hat eine Darstellung der Form

$$(0,1) \qquad \begin{bmatrix} a_1 & 1 \\ 1 & b_1 \end{bmatrix} \perp \cdots \perp \begin{bmatrix} a_n & 1 \\ 1 & b_n \end{bmatrix} \quad \text{mit } a_i, b_i \in A,$$

wobei $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ den zwei-dimensionalen Raum $A \cdot x \oplus A \cdot y$ mit $q(x) = a$, $q(y) = b$ und $(x, y) = 1$ darstellt. Wir werden dafür auch $[a, b]$ schreiben. All das ist im Körperfall leicht zu sehen, und für einen lokalen Ring erhält man es durch Liftung einer entsprechenden Zerlegung von $M/\mathfrak{m} \cdot M$ über $A/\mathfrak{m}$.

Jeder quadratischen Form kann man bekanntlich zwei Invarianten zuordnen, nämlich die Diskriminante und die Witt-Invariante (s. [1, 3, 4]). Wir erinnern kurz an ihre Definitionen.

(0.2) *Die Diskriminante.* Sei B eine kommutative Algebra über dem lokalen Ring A, die als A-Modul frei von Rang 2 ist, und derart, daß die Reduktion $B/\mathfrak{m}B$ eine separable $A/\mathfrak{m}$-Algebra sei. Dann heißt B eine quadratisch-étale A-Algebra. Eine solche Algebra hat eine Basis $\{1, z\}$ über A mit $z^2 - z = \beta \in A$ und $1 + 4\beta \in A^*$. Wir werden dann für B die Bezeichnung $A\left(\frac{1}{\wp}\beta\right)$ verwenden. Die Isomorphieklassen quadratisch-étaler A-Algebren bilden eine Gruppe von Exponenten 2 $\Delta(A)$ (s. [4]) vermöge der Multiplikation

$$(0.3) \qquad \left[A\left(\frac{1}{\wp}\beta_1\right)\right]\left[A\left(\frac{1}{\wp}\beta_2\right)\right] = \left[A\left(\frac{1}{\wp}(\beta_1 + \beta_2 + 4\beta_1\beta_2)\right)\right].$$

Ist $2 \in A^*$, so hat man einen Isomorphismus

$$\Delta(A) \xrightarrow{\sim} A^*/A^{*2}$$

gegeben durch

$$\left[A\left(\frac{1}{\wp}\beta\right)\right] \mapsto 1+4\beta \bmod A^{*2}.$$

Ist $2=0$, so erhält man einen Isomorphismus

$$\Delta(A) \xrightarrow{\sim} A/\wp A \qquad (\wp A = \{\alpha^2+\alpha \mid \alpha \in A\})$$

durch

$$\left[A\left(\frac{1}{\wp}\beta\right)\right] \mapsto \beta \bmod \wp A.$$

Wir werden in diesen Fällen $\Delta(A)$ mit A^*/A^{*2} bzw. $A/\wp A$ identifizieren.

Wir bilden nun die Clifford-Algebra $C(q)=C^+(q) \oplus C^-(q)$ der quadratischen Form (M,q). Dann ist der Zentralisator $D(q)$ von $C^+(q)$ in $C(q)$ eine quadratisch-étale A-Algebra, und wir definieren die Diskriminante von q durch $\delta(q)=[D(q)] \in \Delta(A)$ (s. [2, 4, 5]). Ist $2 \in A^*$, so ist $\delta(q)$ die signierte Determinante $d(B_q) \in A^*/A^{*2}$. Ist $2=0$, so ist $\delta(q) \in A/\wp A$ gerade die Arf-Invariante $a_1 b_1 + \dots + a_n b_n \bmod \wp A$, wenn man q in der Form (0.1) dargestellt hat.

(0.4) *Die Witt-Invariante.* Für einen quadratischen Raum (M,q) ist bekanntlich $C(q)$ im gerade-dimensionalen Fall eine Azumaya Algebra über A, sonst aber ist $C^+(q)$ eine Azumaya *A-Algebra* (s. [2]). Wir definieren die Witt-Invariante von q durch $\omega(q)=$ Bild von $C(q)$ bzw. $C^+(q)$ in der Brauer-Gruppe $\mathrm{Br}(A)$ von A.

Die Invarianten δ, ω induzieren Abbildungen δ: $W_q(A) \to \Delta(A)$ und ω: $W_q(A) \to \mathrm{Br}(A)$, wobei $W_q(A)$ die Witt-Gruppe der quadratischen Formen (s. [2]) über A ist. Ist B eine symmetrische Bilinearform über A und q eine quadratische Form, so kann man bekanntlich eine quadratische Form $B \otimes q$ bilden, und wir haben

(0.5) $\qquad \delta(B \otimes q) = n(B)\,\delta(q) \quad$ mit $n(B)=\dim(B) \bmod 2$

(0.6) $\qquad \omega(B \otimes q) = n(B)\,\omega(q) + (d(B), \delta(q)] \quad$ in $\mathrm{Br}(A)$

(s. Paragraph 1 für die Definition von $(\lambda, \alpha]$).

1. Quaternionen-Algebren

Für einen lokalen Ring A hat man eine biadditive Abbildung (s. [4])

(1.1) $\qquad A^*/A^{*2} \times \Delta(A) \to \mathrm{Br}(A),$

die folgendermaßen definiert ist: zu $\lambda \in A^*$ und $A\left(\frac{1}{\wp}\beta\right)$ ordnen wir die Quaternionen-Algebra $(\lambda,\beta] = A \oplus A\cdot z \oplus A\cdot e \oplus A\cdot z e$ mit den Relationen $z^2 - z = \beta$, $e^2 = \lambda$ und $z e + e z = e$. Man sieht leicht, daß $(\lambda,\beta]$ eine Azumaya-Algebra über A ist, die nur von den Klassen $\langle\lambda\rangle \in A^*/A^{*2}$ und $\left[A\left(\frac{1}{\wp}\beta\right)\right] \in \Delta(A)$ abhängt. Das Bild von $(\lambda,\beta]$ in $\mathrm{Br}(A)$ definiert (1.1). Im Fall $2=0$ erhalten wir also eine Abbildung $A^*/A^{*2} \times A/\wp A \to \mathrm{Br}(A)$. Sei von jetzt an stets $2=0$ vorausgesetzt. Wir wollen nun zwei Quaternionen-Algebren $(\lambda,\beta]$ und $(\mu,\gamma]$ vergleichen.

Wir setzen

$$(\lambda,\beta] = A + A\cdot z + A\cdot e + A\cdot z e \quad \text{und} \quad (\mu,\gamma] = A + A\cdot z' + A\cdot e' + A\cdot z' e'$$

mit den entsprechenden Relationen (s. oben). Da jeder A-Homomorphismus zwischen Azumaya A-Algebren ein Isomorphismus ist (s. [2]), so garantiert uns die Existenz eines A-Homomorphismus zwischen $(\lambda,\beta]$ und $(\mu,\gamma]$ ihre Isomorphie. Nun ist aber ein A-Homomorphismus $\varphi_2\colon (\lambda,\beta] \to (\mu,\gamma]$ durch $\varphi(z)$, $\varphi(e)$ mit den Relationen $\varphi(z)^2 - \varphi(z) = \beta$, $\varphi(e)^2 = \lambda$ und $\varphi(z)\,\varphi(e) + \varphi(e)\,\varphi(z) = \varphi(e)$ eindeutig bestimmt. Wir setzen

$$\varphi(z) = \alpha_0 + \alpha_1\cdot z' + \alpha_2\cdot e' + \alpha_3\cdot z' e', \qquad \varphi(e) = \beta_0 + \beta_1\cdot z' + \beta_2\cdot e' + \beta_3\cdot z' e'$$

an. Man erhält dann aus den obigen Relationen durch Koeffizientenvergleich

i) $\beta = \alpha_1^2\gamma + (\alpha_2^2 + \alpha_2\alpha_3 + \alpha_3^2\gamma)\,\mu + \alpha_0^2 + \alpha_0$
$\alpha_1^2 = \alpha_1,\ \alpha_2 = \alpha_2\alpha_1,\ \alpha_3 = \alpha_3\alpha_1.$

ii) $\beta_1 = 0,\ \lambda = \beta_0^2 + (\beta_2^2 + \beta_2\beta_3 + \beta_3^2\gamma)\,\mu$

iii) $\beta_0 = (\alpha_2\beta_3 + \alpha_3\beta_2)\,\mu,\ \beta_2 = \beta_2\alpha_1,\ \beta_3 = \beta_3\alpha_1.$

Aus (i) folgt $\alpha_1 = 0$ oder $\alpha_1 = 1$. Ist $\alpha_1 = 0$, so erhält man $\alpha_2 = \alpha_3 = \beta_0 = \beta_2 = \beta_3 = 0$ und daher auch wegen (ii) $\lambda = 0$, also ein Widerspruch. Somit muß $\alpha_1 = 1$ sein. Wir erhalten

(1.2) **Lemma.** *Sei A ein lokaler Ring der Charakteristik 2. Die Quaternionen-Algebren $(\lambda,\beta]$ und $(\mu,\gamma]$ sind genau dann isomorph, wenn es Elemente $\alpha_0, \alpha_2, \alpha_3, \beta_0, \beta_2, \beta_3 \in A$ gibt, mit*

$$\beta = \gamma + (\alpha_2^2 + \alpha_2\alpha_3 + \alpha_3^2\gamma)\,\mu + \alpha_0^2 + \alpha_0$$
$$\lambda = \beta_0^2 + (\beta_2^2 + \beta_2\beta_3 + \beta_3^2\gamma)\,\mu$$
$$\beta_0 = (\alpha_2\beta_3 + \alpha_3\beta_2)\,\mu.$$

Der Isomorphismus $\varphi\colon (\lambda,\beta] \to (\mu,\gamma]$ ist dann gegeben durch $\varphi(z) = \alpha_0 + z' + \alpha_2\cdot e' + \alpha_3\cdot z' e'$ und $\varphi(e) = \beta_0 + \beta_2\cdot e' + \beta_3\cdot z' e'$.

2. Sei A ein lokaler Ring der Charakteristik 2. Dann gilt

(2.1) **Satz.** *Zwei vier-dimensionale nicht entartete quadratische Formen $(M_1, q_1), (M_2, q_2)$ über A sind genau dann isomorph, wenn sie ein gemeinsames Element aus $A^* \cup \{0\}$ darstellen und gleiche Arf- und Witt-Invariante haben.*

Beweis. Stellen q_1, q_2 die Null dar, so spalten beide Räume eine hyperbolische Ebene ab. Man kann dann schreiben $q_i = [0,0] \perp p_i$ $(i=1,2)$ mit zwei-dimensionalen Formen p_1, p_2. Gilt $\delta(q_1)=\delta(q_2)$ und $\omega(q_1)=\omega(q_2)$, so erhält man auch $\delta(p_1)=\delta(p_2)$ und $\omega(p_1)=\omega(p_2)$ und infolgedessen $p_1 \cong p_2$ (s. [1, 4]). Seien also q_1, q_2 vier-dimensionale Formen, die ein Element $\varepsilon \in A^*$ gemeinsam darstellen und gleiche Arf- und Witt-Invariante haben. Wir setzen $q_i = (\varepsilon) \otimes p_i$ mit vier-dimensionalen, 1 darstellenden Formen p_1 und p_2. Aus (0.5) folgt $\delta(p_1)=\delta(q_1)=\delta(q_2)=\delta(p_2)$. Ebenso erhält man aus (0.6)

$$\omega(q_i)=\omega((\varepsilon)\otimes p_i)=\omega(p_i)+(\varepsilon, \delta(p_i)]$$

und daher $\omega(p_1)=\omega(p_2)$.

Wir können uns also auf den Fall $\varepsilon=1$ beschränken. Daher kann man schreiben

$$q_1=[1,\alpha]\perp(\lambda)\otimes[1,\beta], \qquad q_2=[1,\gamma]\perp(\mu)\otimes[1,\delta]$$

mit $\lambda, \mu \in A^*$ und $\alpha, \beta, \gamma, \delta \in A$. Wegen der Voraussetzung gilt dann

$$\alpha+\beta \equiv \gamma+\delta \pmod{\wp A}, \qquad [(\lambda, \beta]] = [(\mu, \delta]] \quad \text{in } \mathrm{Br}(A).$$

Da aber zwei Azumaya A-Algebren gleicher Dimension und gleichen Bildes in $\mathrm{Br}(A)$ isomorph sind (A ist ja lokal!), so folgt $(\lambda, \beta] \cong (\mu, \delta]$. Wir beweisen nun, daß q_2 einen Unterraum der Form $[1, \alpha]$ enthält.

Sei x, y, u, v eine kanonische Basis von M_2 bzgl. der obigen Darstellung von q_2. Also gilt $q_2(x)=1$, $(x,y)_2=1$, $q_2(y)=\gamma$, $q_2(u)=\mu$, $(u,v)_2=\mu$ und $q_2(v)=\mu\delta$.

Wir setzen $t=\theta\cdot x+y+\rho\cdot u+\eta\cdot v$ an. Dann gilt

$$q_2(t)=\theta^2+\theta+\gamma+(\rho^2+\rho\eta+\eta^2\delta)\mu.$$

Wegen der Relation $\alpha+\beta=\gamma+\delta+\sigma^2+\sigma$ $(\sigma\in A)$ erhält man

$$q_2(t)=(\theta+\sigma)^2+(\theta+\sigma)+\alpha+\beta+\delta+(\rho^2+\rho\eta+\eta^2\delta)\mu.$$

Aus dem Lemma (1.2) entnehmen wir auf Grund des Isomorphismus $(\lambda, \beta] \cong (\mu, \delta]$, daß es Elemente $\alpha_0, \alpha_2, \alpha_3 \in A$ gibt mit

$$\beta+\delta+(\alpha_2^2+\alpha_2\alpha_3+\alpha_3^2\delta)\mu+\alpha_0^2+\alpha_0=0.$$

Wir wählen nun $\theta=\sigma+\alpha_0$, $\rho=\alpha_2$, $\eta=\alpha_3$ und erhalten $q_2(t)=\alpha$. Andererseits ist $(x,t)_2=1$, d.h. der Unterraum $\langle x, t\rangle$ von M_2 ist isomorph zu

$[1,\alpha]$. Wir können somit schreiben $q_2=[1,\alpha]\perp p$ mit einer zwei-dimensionalen Form p. Dann folgt $\delta(q_2)=\alpha+\delta(p)$ und daher $\delta(p)=\beta \bmod \not{p}\, A$. Ebenso ist $\omega(q_2)=\omega(p)$, also $\omega(p)=\omega(q_1)=[(\lambda,\beta]]$. Daraus folgt nun (s. [4]) $p\cong(\lambda)\otimes[1,\beta]$ und somit $q_1\cong q_2$.

Literatur

1. Arf, C.: Untersuchungen über quadratische Formen in Körpern der Charakteristik 2. J. reine angew. Math. **183**, 148–167 (1941).
2. Bass, H.: Lectures on algebraic K-theory. Tata Inst. Fund. Res., Bombay 1967.
3. Bourbaki, N.: Formes sesquilineaires et formes quadratiques. Algebre, Chap. 9. Paris: Hermann 1959.
4. Knebusch, M.: Bemerkungen zur Theorie der quadratischen Formen über semi-lokalen Ringen. Schriften des mathematischen Inst. der Univ. des Saarlandes (A 70-04). Saarbrücken 1971.
5. Kneser, M.: Bestimmung des Zentrums der Cliffordschen Algebren einer quadratischen Form über einem Körper der Charakteristik 2. J. reine angew. Math. **193**, 123–125 (1954).
6. Serre, J. P.: Corps locaux. Paris: Hermann 1962.

Dr. R. Baeza
Mathematisches Institut der
Universität des Saarlandes
D-6600 Saarbrücken
Bundesrepublik Deutschland

(Eingegangen am 20. Juni 1972)