Mathematische Zeitschrift

Band 124 · Heft 1 · 1972

Springer-Verlag · Berlin · Heidelberg · New York

Math. Z. – MAZEAX 124(1) 1–88 (1972) 21. XII. 1971

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. Kamke* 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 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.—.

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.

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.

"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 will not be published.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:

H. Wielandt, 74 Tübingen, Math. Inst. d. Universität
T. tom Dieck, 66 Saarbrücken, Math. Inst. d. Universität
E. Heinz, 34 Göttingen, Math. Inst. d. Universität
A. Pfluger, Zürich (Schweiz), Eidgen. Technische Hochschule
K. Zeller, 74 Tübingen, Math. Inst. d. Universität

Springer-Verlag

69 Heidelberg 1
Postfach 1780
Fernsprecher (06221) 49101
Fernschreibnummer 04-61723

1 Berlin 33
Heidelberger Platz 3
Fernsprecher (0311) 822001
Fernschreibnummer 01-83319

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

Math. Z. 124, 1–8 (1972)

Some Characterizations of Finite Miquelian Möbius Planes

NICHOLAS KRIER

In 1965, Hering [3] provided a classification of Möbius planes analogous to the Lenz-Barlotti classification of projective planes. As with the Lenz-Barlotti classification, the Hering classification contains finitely many possibilities but existence is not guaranteed in each case. The purpose of the present paper is to prove that finite Möbius planes of even order and of Hering type at least II.2, III.2, IV.1, or IV.2 are miquelian and that finite Möbius planes of odd order and Hering type at least III.2 must also be miquelian. Since miquelian planes are at least of these types, we have some new characterizations of finite miquelian planes. An alternate way of stating the results is that finite Möbius planes of precisely the above types do not exist.

The author here gives especial thanks to Dr. Jill C.D.S. Yaqub for stimulating conversations.

§ 1. Basic Definitions and Theorems

Further results may be found in Dembowski [2], where a Möbius plane is called an inversive plane.

Definition 1. A *Möbius plane* $\mathfrak{M}=(P,\mathscr{C})$ is a set P of points and a collection $\mathscr{C}$ of subsets of P, called circles, with the usual convention that the point p is on the circle C if and only if $p\in C$, subject to the following conditions: (M.1) each triple of points is on exactly one circle C; (M.2) given $p\in C$, $q\notin C$, there is a unique circle D such that $p,q\in D$ and $D\cap C=\{p\}$; (M.3) there are at least 4 points not all on one circle and each circle contains at least one point.

Theorem A [2, p. 253]. *If $\mathfrak{M}$ is a Möbius plane and p is a point of $\mathfrak{M}$, the "internal structure" $\mathfrak{M}_p=(\{P\in p\},\{C\in\mathscr{C}\colon p\in C\})$ is an affine plane.*

Definition 2. In a Möbius plane $\mathfrak{M}$, a *pencil* $\mathfrak{b}$ with carrier p is a maximal set of circles mutually tangent at p. (In $\mathfrak{M}_p$, the set $\{C\setminus\{p\}\colon C\in\mathfrak{b}\}$ is a parallel class of lines.) A *bundle* with carriers x, y for $x\neq y$ is the set of all circles in $\mathfrak{M}$ containing both x and y, and is denoted by $[x, y]$.

Definition 3. A *translation* of a Möbius plane $\mathfrak{M}$ is an automorphism fixing one point p and all the circles in a pencil with carrier p. M is called $\mathfrak{b}$-*transitive* if given any $q, r\neq p$ with the circle containing p, q, r in $\mathfrak{b}$, there is a translation fixing p and mapping q to r.

Definition 4. An (x, y)-*dilatation* of a Möbius plane $\mathfrak{M}$ is an automorphism fixing x and y and each circle in the bundle $[x, y]$. M is called (x, y)-*transitive* if, given $q, r, \neq x, y$ with q, r, x, y concyclic, there is an (x, y)-dilatation mapping q to r.

The Hering classification is an enumeration of the possible configurations of pencils $\mathfrak{b}$ and bundles $[x, y]$ for which a Möbius plane may be $\mathfrak{b}$- and (x, y)-transitive.

Definition 5. If M is a Möbius plane with more than 5 points, let $\mathscr{F}$ be the set of pencils $\mathfrak{b}$ and bundles $[x, y]$ for which $\mathfrak{M}$ is $\mathfrak{b}$- and (x, y)-transitive. Then $\mathfrak{M}$ is of type

II.2 if $\mathscr{F} = \{\mathfrak{b}\} \cup \{[p, x]: p$ is the carrier of $\mathfrak{b}$; $x \neq p$, x is on a distinguished circle $C \in \mathfrak{b}\}$;

III.2 if $\mathscr{F} = \{\mathfrak{a}$: $\mathfrak{a}$ has carrier $p\} \cup \{[p, x]: x \neq p\}$;

IV.1 if $\mathscr{F}$ consists of all pencils containing the same circle C;

IV.2 if $\mathscr{F} = \{\mathfrak{b}: C \in \mathfrak{b}\} \cup \{[x, y]: x \neq y, x, y \in C\}$;

VII.2 if $\mathscr{F}$ consists of all pencils and all bundles.

Definition 6. An *ovoid* in a projective 3-space R is a set of points $\Omega \neq \emptyset$ such that for each $p \in \Omega$, the set of lines tangent to Ω at p form a plane in R.

Theorem B [2, p. 254]. *If Ω is an ovoid, then $\mathfrak{M}(\Omega) = (\Omega, \{\Omega \cap \pi$: π is a plane and $|\Omega \cap \pi| \geqq 2\})$ is a Möbius plane. A Möbius plane isomorphic to some $M(\Omega)$ will be called ovoidal.*

Theorem C [2, p. 264]. *If $\mathfrak{M}$ contains a finite number of points, then there is a natural number $m > 1$, called the order of $\mathfrak{M}$, such that* (i) $|\mathscr{C}| = m(m^2+1)$; (ii) $|P| = m^2+1$; (iii) *each circle contains $m+1$ points;* (iv) *each point is on $m(m+1)$ circles;* (v) *each bundle contains $m+1$ circles.*

Theorem D (Dembowski) [2, p. 268 and p. 270]. *If $\mathfrak{M}$ is a Möbius plane of even order, then $\mathfrak{M}$ is ovoidal. Each automorphism of $\mathfrak{M}(\Omega)$ is a collineation of R that fixes Ω.*

Theorem E [2, p. 257]. *A Möbius plane is miquelian if and only if $\mathfrak{M}$ is isomorphic to some $\mathfrak{M}(\Omega)$ with Ω a non-ruled quadric surface in R.*

Theorem F [2, p. 279]. *A finite Möbius plane is miquelian if and only if $\mathfrak{M}$ is of type* VII.2.

Definition. An *oval* in an affine or projective plane is a set of points $\mathcal{O} \neq \emptyset$ such that (i) no three points of $\mathcal{O}$ are collinear, and (ii) given $p \in \mathcal{O}$, there is a unique line L such that $\mathcal{O} \cap L = \{p\}$.

Theorem G. *In a Möbius plane $\mathfrak{M}$, each circle of $\mathfrak{M}$ not containing p is an* oval in $\mathfrak{M}_p$.

Theorem H (Qvist) [2, p. 148]. *If $\mathcal{O}$ is an oval in a projective or affine plane of order n, then $\mathcal{O}$ contains $n+1$ points, and* (i) *if n is odd, then no three tangent lines meet at a point and* (ii) *if n is even, then there is a unique point, called the knot of the oval, which is on every tangent line of $\mathcal{O}$.*

§ 2. Möbius Planes of Even Order

Let $\mathfrak{M}$ be a finite Möbius plane of even order larger than 2. Then by a theorem of Dembowski (Theorem D), $\mathfrak{M}$ is isomorphic to some $\mathfrak{M}(\Omega)$ where Ω is an ovoid in the three-dimensional projective space R over the field $F=GF(q)$. Furthermore, each automorphism of $\mathfrak{M}(\Omega)$ extends uniquely to a collineation of R which fixes Ω. Identify $\mathfrak{M}$ with $\mathfrak{M}(\Omega)$ and automorphisms of $\mathfrak{M}$ with the corresponding collineations of R. Choose homogeneous coordinates (x, y, z, w) for R and let u: $(0, 0, 1, 0) \in \Omega$ with tangent plane $w=0$. Let c: $(0, 0, 0, 1) \in \Omega$ with tangent plane $z=0$. Then the points of Ω are u and $\{(x, y, \varphi(x, y), 1)\colon x, y \in F\}$ where φ is a function of x and y into F such that $\varphi(x, y)=0$ if and only if $x=y=0$. By a theorem of Lüneburg [2, p. 270] every involution of $\mathfrak{M}$ is given by a linear collineation of R. Each translation of $\mathfrak{M}$ is an involution since $\mathfrak{M}_p$ is desarguesian for each point p of $\mathfrak{M}$ [2, p. 254]. With these preliminaries, the following lemmas can be easily proved.

Lemma 1. *If T is a translation of $\mathfrak{M}$ fixing u and exchanging c and k: $(0, k, \varphi(0, k), 1)$ for $k \neq 0$, then T may be represented by the matrix*

$$T^* := \begin{pmatrix} 1 & 0 & d(k) & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & k & \varphi(0, k) & 1 \end{pmatrix}$$

with $d(k)$ an element of F depending on k.

Proof. u is a fixed point. Hence, if the third row of T^* is given by (u, v, w, x) then since u is a fixed point, u, v, x are 0. Similarly, since $w=0$ is a fixed plane, the first three entries in the fourth column of T^* are zero. The fourth row is determined up to a scalar factor, since $cT=k$ and we let this factor be such that the $(4, 4)$ entry is 1. Since $T^{*2}=I(T^2=sI$, but the $(4, 4)$ entry is 1, whence $s=1)$ the $(3, 3)$ entry must have square 1, hence it is 1, since F is of characteristic 2. If the entries of the second row of T^* are $(g, h, j, 0)$, from $kT=c$ we obtain the equations $gk=0$, $hk+k=0$, $jk+\varphi(0, k)+\varphi(0, k)=0$, whence $g=0$, $h=1$, and $j=0$. Again since $T^{*2}=I$, the $(1, 1)$ entry must be 1. Since T^* is a translation of $\mathfrak{M}$, it must fix all tangent bundles of circles through u. This corresponds to mapping planes through u not $w=0$ into parallel planes in the affine space $R/\langle w=0\rangle$. If the first row of T^* is given by $(1, b, d, 0)$, then the point $(1, 0, 0, 1)$ is mapped into $(1, b+k, d+\varphi(0, k), 1)$, but the plane $y=0$ must be mapped into the plane $y=k$ since $cT=k$, whence $b+k=k$, so $b=0$. This proves Lemma 1.

Lemma 2. *If T is a translation of $\mathfrak{M}$ fixing k: $(0, k, \varphi(0, k), 1)$ for $k \neq 0$ and exchanging c and u, then T may be represented by the matrix*

$$T^*: \begin{pmatrix} 1 & b(k) & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (\varphi(0, k))^{-1} \\ 0 & 0 & \varphi(0, k) & 0 \end{pmatrix}$$

with $b(k)$ an element of F depending on k.

Proof. The zero entries of the third and fourth rows are obtained from the equations $uT=c$ and $cT=u$. The zero entries of the third and fourth columns are obtained from the fact that the planes $z=0$ and $w=0$ are switched by T^*. Since c, k, u are in the plane $x=0$, this plane is fixed by T^*, whence we obtain the zero entries in the first column. Since $T^{*2}=sI$ and q is even, we scale s to 1 so that the (1, 1) entry is 1. Then the (2, 2) entry must also be 1, and the other terms besides $b(k)$ are derived from the equation $kT^*=k$.

Lemma 3. *If D is a nonidentity dilatation of $\mathfrak{M}$ fixing c and u and mapping $(0, 1, \varphi(0, 1), 1)$ to $(0, k, \varphi(0, k), k)$ for $k \neq 0, 1$, then D may be represented by the matrix*

$$D^*: \quad \begin{pmatrix} k & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & \psi(k) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with $0 \neq \psi(k) \in F$, and with $\psi(k)\varphi(0, 1)=\varphi(0, k)$.

Proof. We first show that D is a projective (linear) collineation of R. Since D fixes all circles of $\mathfrak{M}$ through c and u, it fixes all points of R through c and u as well as $w=0$ and $z=0$, whence it fixes all points on the line $(w=0)\cap(z=0)$, and so D fixes each line in the plane $z=0$ containing c. Let D be represented by $(av)D=a^\sigma(vD)$ with $a\in F$, σ an automorphism of F, and v a vector in R. Let x, y be two linearly independent vectors in the plane $z=0$, with neither x nor y on the plane $w=0$. Then the line joining c and x is fixed by D, so $xD=c_1x$. Similarly $yD=c_2y$, and $(x+y)D=c_3(x+y)$. Thus, $c_3x+c_3y=(x+y)D=xD+yD=c_1x+c_2y$. Since x and y are linearly independent, $c_3=c_1=c_2$. Hence, for each vector v in the plane $z=0$, $vD=c_3v$. Thus, if $0\neq a\in F$, $a^\sigma c_3v=a^\sigma(vD)=(av)D=c_3av$. Hence, $a^\sigma=a$ for each $a\in F$, so that σ is the identity automorphism of F, and so D is linear.

The zero elements in the third and fourth rows and columns are derived from the fact that $u, c, z=0$ are fixed by D. In the plane $z=0$ all lines through c are fixed (being the intersection of $z=0$ with a fixed plane through c and u). Hence, the collineation induced on $z=0$ is a dilatation, so the submatrix of the first two rows and columns must be a scalar matrix. Since $(0, 1, \varphi(0, 1), 1)D=(0, k, \varphi(0, k), 1)$, by scaling the (4, 4) entry to 1 we obtain the desired matrix.

Remark. $\Omega\cap(x=0)$ is an oval in the plane $x=0$. The tangent lines at u and c are $(x=0)\cap(w=0)$ and $(x=0)\cap(z=0)$ respectively, meeting at the point $(0, 1, 0, 0)$. Hence this point is the knot of the oval (Theorem H), and every tangent line to $\Omega\cap(x=0)$ contains this point. Thus the function $\varphi(0, m)$ is a bijection of F. Similarly, $\varphi(a, 0)$ is also a bijection of F.

Theorem 1. *If $\mathfrak{M}$ is a finite Möbius plane of even order and of Hering class at least* II.2, *then $\mathfrak{M}$ is miquelian* (hence of class VII.2).

Proof. $\mathfrak{M}$ contains a distinguished point p and circle C such that $\mathfrak{M}$ is $\mathfrak{b}$-transitive with $C\in\mathfrak{b}$ and (p, k)-transitive for each $p\neq k$ on C. Let u be the point p and C the circle $(x=0)\cap\Omega$. Denote by $T_{c,k}$ the translation fixing u and exchanging c and k: $(0, k, \varphi(0,k), 1)$, and by D_k the dilatation fixing c, u and mapping $(0, 1, \varphi(0,1), 1)$ to k. Thus the $T_{c,k}$'s and D_k's form subgroups of the automorphism group of $\mathfrak{M}$, and by multiplying the matrices of Lemmas 1 and 3 we find that $T_{c,k}\cdot T_{c,m}=T_{c,k+m}$ and $d(k)+d(m)=d(k+m)$; $D_r\cdot D_s=D_{rs}$ and $\psi(k)\,\psi(m)=\psi(km)$.

We now show that $d(k)=0$ implies that $k=0$. If $d(k)=0$ for $k\neq 0$, choose $0\neq m$ such that $\varphi(m,0)=\varphi(0,k)$ which can be done by the remark above. Then $(m, 0, \varphi(m,0), 1)\,T_{c,k}=(m, k, \varphi(m,0)+\varphi(0,k), 1)$ so $\varphi(m,k)=\varphi(m,0)+\varphi(0,k)=0$, a contradiction.

The conjugate of a translation is again a translation, so

$$D_r^{*-1}\,T_{c,k}^{*}\,D_r^{*}=\begin{pmatrix} r^{-1} & 0 & 0 & 0\\ 0 & r^{-1} & 0 & 0\\ 0 & 0 & \psi(r^{-1}) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & 0 & d(k) & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & k & \varphi(0,k) & 1\end{pmatrix}\begin{pmatrix} r & 0 & 0 & 0\\ 0 & r & 0 & 0\\ 0 & 0 & \psi(r) & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

$$=\begin{pmatrix} 1 & 0 & \psi(r)\,r^{-1}d(k) & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & rk & \psi(r)\,\varphi(0,k) & 1\end{pmatrix}=T_{c,rk}.$$

Since $(a, b, \varphi(a,b), 1)\,D_r=(ra, rb, \psi(r)\,\varphi(a,b), 1)$,

$$\varphi(ar, br)=\psi(r)\,\varphi(a,b). \tag{1}$$

Hence $\psi(r)\,r^{-1}d(k)=d(rk)$. Similarly, $d(kr)=\psi(k)\,k^{-1}d(r)$. Thus for all $r, k\neq 0$, $rd(r)/\psi(r)=k\,d(k)/\psi(k)$, and it follows that $rd(r)/\psi(r)=t$, where t is fixed, for all $r\neq 0$. Hence $d(r)=t\,\psi(r)\,r^{-1}$ for all $0\neq r\in F$.

Note that $\varphi(0, k+m)=\varphi(0,k)+\varphi(0,m)$ from the equation

$$(0, m, \varphi(0,m), 1)\;T_{c,k}=(0, k+m, \varphi(0,k)+\varphi(0,m), 1).$$

Hence, by (1),

$$\psi(a+b)\,\varphi(0,1)=\varphi(0, a+b)=\varphi(0,a)+\varphi(0,b)=\big(\psi(a)+\psi(b)\big)\,\varphi(0,1),$$

whence ψ is an additive function of F (defining $\psi(0)=0$), and $a\neq 0$ implies $\psi(a)\neq 0$ since the matrix D_a^* is nonsingular. Thus ψ is an automorphism of F, so $\psi(r)=r^{2^f}$ $(0\leqq f<e)$. But then $d(r)=t\,\psi(r)\,r^{-1}=t\,r^{2^f-1}$. Since d is additive it is a polynomial of 2-powers; i.e., $d(k)=a_0k^{2^0}+a_1k^{2^1}+a_2k^{2^2}+\cdots+a_mk^{2^m}$ with $1\leqq m<e$. But we have already represented d as a polynomial and this representation is unique [Artin, 1, p. 37], so 2^f-1 must be a 2-power. But this can only hold if $2^f-1=1$, or $f=1$. Thus $\psi(r)=r^2$ and $d(k)=tk$. Since $(a, 0, \varphi(a,0), 1)\;T_{c,k}=(a, k, \varphi(a,0)+\varphi(0,k)+tak, 1)$, we see that

$$\varphi(a,k)=\varphi(a,0)+\varphi(0,k)+t\,a\,k=a^2\,\varphi(1,0)+k^2\,\varphi(0,1)+t\,a\,k.$$

Thus Ω is given by a quadric surface, so $\mathfrak{M}(\Omega)$ is miquelian (Theorem E and F) and hence of class VII.2.

Corollary. If $\mathfrak{M}$ is a finite Möbius plane of even order and of class III.2 or IV.2, then $\mathfrak{M}$ is miquelian.

Proof. If $\mathfrak{M}$ is of class III.2 or IV.2, then $\mathfrak{M}$ is at least of class II.2, and by Theorem 1 $\mathfrak{M}$ is miquelian.

Theorem 2. *If $\mathfrak{M}$ is a finite Möbius plane of even order and of class at least* IV.1, *then $\mathfrak{M}$ is miquelian.*

Proof. Let $(x=0)\cap\Omega=C$ be the distinguished circle such that $\mathfrak{M}$ is b-transitive for each pencil containing C. Denote by ${}_kT_{c,u}$ the translation exchanging c and u, and fixing k: $(0, k, \varphi(0,k), 1)$. Then

$$(0, a, \varphi(0,a), 1)\,{}_kT_{c,u}=(0, a, \varphi(0,k),\ \varphi(0,a)/\varphi(0,k))$$

from Lemma 2 and may be rewritten $(0, a\varphi(0,k)/\varphi(0,a), (\varphi(0,k))^2/\varphi(0,a), 1)$. We may require that $\varphi(0,1)=1$. Thus, letting $a=1$, $\varphi(0,\varphi(0,k))=(\varphi(0,k))^2$. But by the remark following Lemma 3, $\varphi(0,k)$ is a bijection of F so that $\varphi(0,b)=b^2$ for each $b\in F$. Now

$${}_kT^*_{c,u}\,{}_mT^*_{c,u}=\begin{pmatrix}1 & b(k) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1/k^2\\ 0 & 0 & k^2 & 0\end{pmatrix}\cdot\begin{pmatrix}1 & b(m) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1/m^2\\ 0 & 0 & m^2 & 0\end{pmatrix}$$
$$\cdot\begin{pmatrix}1 & b(k)+b(m) & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & (m/k)^2 & 0\\ 0 & 0 & 0 & (k/m)^2\end{pmatrix}$$

This map sends $(0, a, a^2, 1)$ to $(0, a(m/k)^2, (a(m/k)^2)^2, 1)$ and fixes c and u. Let $k=1$, $m^2=n/a$. Then the above product will conjugate ${}_aT^*_{c,u}$ to ${}_nT^*_{c,u}$ since then the product maps $(0, a, a^2, 1)$ on $(0, n, n^2, 1)$ and hence, geometrically, conjugates ${}_aT^*_{c,u}$ to ${}_nT^*_{c,u}$. But

$$\begin{pmatrix}1 & b(1)+b(m) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & a/n & 0\\ 0 & 0 & 0 & n/a\end{pmatrix}\begin{pmatrix}1 & b(a) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1/a^2\\ 0 & 0 & a^2 & 0\end{pmatrix}\begin{pmatrix}1 & b(1)+b(m) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & n/a & 0\\ 0 & 0 & 0 & a/n\end{pmatrix}$$
$$=\begin{pmatrix}1 & b(a) & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1/n^2\\ 0 & 0 & n^2 & 1\end{pmatrix}.$$

Hence $b(a)=b(n)$, so b is a constant function on the nonzero elements of F, and $b(k)+b(m)=0$ for any $k, m\neq 0$. Thus the maps

$$ {}_1T^*_{c,u}\cdot {}_mT^*_{c,u}=\begin{pmatrix} m^2 & 0 & 0 & 0\\ 0 & m^2 & 0 & 0\\ 0 & 0 & m^4 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} $$

are dilatations with centers c, u. Since $x\to x^2$ is an automorphism of F, $\mathfrak{M}$ is (c, u)-transitive, and hence of class at least IV.2. By the corollary, $\mathfrak{M}$ is miquelian.

§3. Möbius Planes of Odd Order

Theorem 3. *Any finite Möbius plane of odd order and of class at least* III.2 *is miquelian.*

Proof. Let $\mathfrak{M}$ be a Möbius plane of odd order q and of class at least III.2, and let u be the point such that $\mathfrak{M}$ is (u, p)-transitive for all $u\neq p$ and b-transitive for each pencil with carrier u. Then $\mathfrak{M}_u$ is (x, L_∞)-transitive for every point x in $\mathfrak{M}_u$. Hence $\mathfrak{M}_u$ is desarguesian by Pickert [4, p. 83], and each circle in $\mathfrak{M}$ not containing u is an oval, and by a a theorem of Segre [2, p. 49] is an irreducible conic in $\mathfrak{M}_u$. Let G be the group generated by the translations and dilatations fixing p. Then $|G|=q^2(q-1)$. By a proper choice of coordinates in $\mathfrak{M}_u$ we can choose a representative C of a G-orbit of circles by the equation C: $x^2+cy^2+f=0$, with $-c$ a nonsquare in $GF(q)$. The G-orbit contains the $q^2(q-1)/2$ circles $\{(x, y)\colon (x-a)^2+c(y-b)^2+fr^2=0\}$, but the number of circles not containing u in $\mathfrak{M}$ is $q^2(q-1)$ (Theorem C), whence there are precisely two G-orbits of circles. Since G preserves parallels, by using M.2 in the definition of a Möbius plane, it is easily seen that the circles of distinct G-orbits must have complementary sets of tangent slopes.

In general, if K is the conic: $x^2+2bxy+dy^2+h=0$, with b^2-d a nonsquare, then $y=mx+k$ is a tangent line to K if and only if

$$ k^2(b^2-d)=h(1+2bm+dm^2) \tag{1} $$

and $x=k$ is a tangent line to K if and only if

$$ k^2(b^2-d)=dh. \tag{2} $$

Denote the set of tangent slopes (possibly including ∞) by M_K. Let Q': $x^2+2bxy+dy^2+e'=0$ be a representative of the second G-orbit. Then Q: $x^2+2bxy+dy^2+e=0$, with ee' a nonsquare must have the same set of tangent slopes as C, that is $M_C=M_{Q'}$. We may assume without loss of generality that $\infty\notin M_C$, so that by (2) cf is a square. (In case $q\equiv 1 \pmod 4$) f must be a nonsquare, and if $q\equiv 3 \pmod 4$, f must be a square.) Note that by (1) $m\in M_C$ if and only if $-m\in M_C$ since $b=0$ for C. $m\in M_Q$ if and only if $f(1+2bm+dm^2)$ is a nonsquare. If $q\equiv 3 \pmod 4$, then -1 is a nonsquare. Since $-c$ is a nonsquare and cf is a square, f must be a square. If $q\equiv 1 \pmod 4$, then -1 is a square, so in this case f is a nonsquare. In either case $1+2bm+dm^2=-x^2$ for some x, or

$$x^2+d(m+b/d)^2+(1-b^2/d)=0. \tag{3}$$

Now, $m\in M_C$ if and only if $-m\in M_C$, so that $-m\in M_Q$, or, for some x

$$x^2d(m-b/d)^2+(1-b^2/d)=0. \tag{4}$$

If $m\in M_C$, then m satisfies Eq. (4) so that $m+2b/d$ satisfies (3). Hence if $m\in M_C$, then $m+2b/d\in M_C$. Let $q=p^a$ with p an odd prime. Since $\infty\notin M_C$, the mapping $m\to m+2b/d$ permutes the elements of M_C in cycles of length p if $b\neq 0$. But this implies that $p|(q+1)/2$ a contradiction. Hence $b=0$.

Case 1. $q\equiv 3\ (\mathrm{mod}\ 4)$. Q: $x^2+dy^2+e=0$, with $c=dt^2$, and without loss of generality $e=f=1$. Then T, the linear transformation which fixes x and maps y to ty, maps lines with slope m to lines with slope tm, maps the conic C to Q and hence fixes M_C. If $t^k=1$, then $k|(q-1)$, and since $f=1$, and c is a square, 0, $\infty\notin M_C$ so that $k|(q+1)/2$, which implies that $k|2$, or $c=d$.

Case 2. $q\equiv 1\ (\mathrm{mod}\ 4)$. By (1), $M_C=M_Q$ implies that $1+cm^2$ is a square if and only if $1+dm^2$ is a square. Since $-d$ is a nonsquare, $1+dm^2$ never vanishes, so that $(1+cm^2)/(1+dm^2)$ is always a nonzero square and $(1+cm^2)/(1+dm^2)=s$ has at most two solutions. Hence there are at least $(q-1)/2$ different square values for nonzero m, but there are only $(q-1)/2$ nonzero squares in $GF(q)$. Hence there is a $z\neq 0$, such that $(1+cz^2)/(1+dz^2)=1$, so that $c=d$.

In either case there is only one possibility for the second G-orbit. This shows that there is a unique Möbius plane of odd order q and of class at least III.2, but the miquelian plane of order q is of class at least III.2, so that $\mathfrak{M}$ is miquelian.

References

1. Artin, E.: Galois theory. Notre Dame: University of Notre Dame Press 1959.
2. Dembrowski, P.: Finite geometries. Berlin-Heidelberg-New York: Springer 1968.
3. Hering, C.: Eine Klassifikation der Möbius-Ebenen. Math. Z. **87**, 252–262 (1965).
4. Pickert, G.: Projektive Ebenen. Berlin-Göttingen-Heidelberg: Springer 1955.

Dr. Nicholas Krier
Department of Mathematics
Colorado State University
Ft. Collins, Colorado 80521
U.S.A.

(Received February 16, 1970)

Math. Z. 124, 9–19 (1972)

A General Injectivity for Modules

SHEILA R. O'DONNELL

Throughout this paper, R will denote an associative ring with unity, all the R-modules ${}_RA=A$ will be left unital R-modules, and all mappings $f: X \to Y$ between R-modules will be R-homomorphisms, unless otherwise specified.

We shall be concerned with an arbitrary, but fixed class Γ of R-monomorphisms $\alpha: A \to B$. If $C = \operatorname{Coker} \alpha$, then we shall say that the arising exact sequence $E: 0 \to A \xrightarrow{\alpha} B \to C \to 0$ is a *Γ-sequence*. $\mathscr{A}$ will denote the class of the domains of the maps of Γ, $\mathscr{B}$ the class of codomains, and $\mathscr{C}$ the class of the cokernels.

An R-module M will be called *injective with respect to* α or *E-injective* if for each homomorphism $f: A \to M$, there exists an $\bar{f}: B \to M$ such that $\bar{f} \circ \alpha = f$. We define an R-module M to be *Γ-injective* if M is injective with respect to α, for each $\alpha \in \Gamma$.

This paper will examine Γ-injectivity in general, establishing its properties and pointing out how these properties generalize known results for previously studied types of injectivity and divisibility. Examples including two new ones are discussed in II. Sections III and IV are devoted to the question as to when the class of Γ-injective modules is closed with respect to epimorphic images. Section V gives a necessary and sufficient condition that all R-modules contain maximum Γ-injective submodules. In VI and VII, respectively, new concepts of purity and torsionfreeness are defined and studied in relation to Γ-injectivity.

This paper is part of the doctoral dissertation submitted to Tulane University. The writer would like to express her deep appreciation for the assistance and inspiration given to her by Professor Laszlo Fuchs.

I. Elementary Properties

We begin by stating several basic properties of Γ-injectivity, which do not require proof.

(i) Let $E: 0 \to A \to B \to C \to 0$ and $E': 0 \to A \to B' \to C \to 0$ be two equivalent exact sequences of modules. Then M is E-injective if and only if it is E'-injective.

(ii) Splitting exact sequences can be included in the Γ-sequences without changing the concept of Γ-injectivity.

(iii) Let Γ, Γ' be two classes of R-monomorphisms such that $\Gamma \subset \Gamma'$. If M is Γ'-injective, then M is Γ-injective.

(iv) Let I be an index set. If M is injective with respect to $\alpha_i: A_i \to B_i$, for all $i \in I$, then M is injective with respect to the direct sum

$$\bigoplus_i \alpha_i: \bigoplus_i A_i \to \bigoplus_i B_i, \quad \text{where naturally} \quad \bigoplus_i \alpha_i | A_i = \alpha_i, \quad \text{for all } i \in I.$$

Conditions for closure of the class of Γ-injectives with respect to direct sums, direct limits or direct products are analogous to the case of injectivity.

Proposition 1.1. *Let I be any index set. $\prod_{i\in I} X_i$ is Γ-injective if and only if X_i is Γ-injective, for all $i\in I$.*

Proof. Standard argument.

Let $\mathbf{2}^I$ denote the Boolean algebra of all subsets of the set I, and $\mathscr{I}$ a dual filter of $\mathbf{2}^I$. Then define

$$\bigoplus_{\mathscr{I}} X_i = \{x=(x_i)_i \in \prod_{i\in I} X_i\colon s(x)\in\mathscr{I}\},$$

where $s(x)$ is the support of x, i.e. $s(x)=\{i\in I\colon x_i \neq 0\}$.

Proposition 1.2. *Let $\mathscr{A}$ be contained in the class of finitely generated modules. If X_i is Γ-injective, for all $i\in I$, then $\bigoplus_{\mathscr{I}} X_i$ is also Γ-injective.*

Proof. Standard argument.

In particular, we can conclude that if $\mathscr{A}$ is contained in the class of finitely generated R-modules, then the class of Γ-injective modules is closed under direct sums.

In order to obtain a similar result for direct limits, either the class $\mathscr{A}$ or the homomorphisms in the direct system must be further restricted.

Proposition 1.3. (i) *If $\mathscr{A}$ is contained in the class of finitely presented modules, then the class of Γ-injective modules is closed under direct limits.*

(ii) *If $\mathscr{A}$ is contained in the class of finitely generated modules, then the class of Γ-injective modules is closed with respect to direct limits of direct systems in which the homomorphisms are monic.*

The proof of part (ii) is routine, while part (i) will be a direct consequence of the following lemma.

Lemma 1.4. *Let $\{M_i, \pi_i^j\}_{i,j\in I}$ be a direct system of modules with direct limit M and natural homomorphisms $\pi_i\colon M_i \to M$. If X is a finitely presented module and if $\phi\colon X\to M$ is a homomorphism, then there exists a $k\in I$ and a homomorphism $\theta\colon X\to M_k$ such that $\pi_k\circ\theta=\phi$.*

Proof. Let X be generated by $x_1,\dots,x_n$ with defining relations

$$\sum_{t=1}^{n} r_{tj}x_t=0, \qquad \text{where } 1\leqq j\leqq s \text{ and } r_{tj}\in R.$$

Now for each t, $\phi x_t=\pi_{i_t} m_{i_t}$, for some $m_{i_t}\in M_{i_t}$; thus

$$\sum_{t=1}^{n} r_{tj}\pi_{i_t} m_{i_t}=0.$$

Hence, there exists a k with $k\geqq i_t$ for all t, such that

$$\sum_{t=1}^{n} r_{tj}\pi_{i_t}^{k} m_{i_t}=0, \qquad \text{for all } j.$$

So define $\theta\colon X\to M_k$ by $\theta x_t=\pi_{i_t}^k m_{i_t}$, for $1\leqq t\leqq n$. This θ is as desired.

II. Examples

There are several examples of Γ-injectivity.

1) If Γ is the class of all monomorphisms, then Γ-injectivity is exactly injectivity.

2) If the Γ-sequences are precisely the splitting exact sequences, then all modules are Γ-injective.

3) If the Γ-sequences are precisely the pure-exact sequences in the sense of Cohn [2], then the Γ-injective modules are the so-called algebraically compact modules.

4) Fix the module M and let Γ_M be the set of all natural inclusion maps $i: N \to M$ where N is some submodule of M. Then M is Γ_M-injective if and only if M is quasi-injective. In view of Baer's criterion, for any module X, X is Γ_R-injective if and only if X is injective.

5) Let S be a subring of R which contains the unity of R. We can choose, for instance, Γ to be the class of all R-monomorphisms α: ${}_RA \to {}_RB$ such that $\operatorname{Im} \alpha$ is a direct summand of B as an S-module. Such sequences are frequently used in relative homological algebra.

6) Let R be a commutative domain. Under the usual definition of torsion, let an exact sequence E: $0 \to A \to B \to C \to 0$ be called *torsion-splitting* if Ei splits for i: $T(C) \to C$ where $T(C)$ is the torsion part of C and i is the inclusion map (see [4]). If we let the Γ-sequences consist of the torsion-splitting exact sequences, then a module M is Γ-injective if and only if M is cotorsion in the sense of Matlis [12].

Γ-injectivity is a generalization of numerous previously studied types of divisibility and injectivity for modules:

7) The usual definition of divisibility when R is a principal ideal domain has $\Gamma = \{\alpha: R\lambda \to R: \lambda \in R\}$. In [6], Hattori uses the same class for Γ, but he allows R to be any ring. In [9], Levy restricts Γ to the natural inclusion maps of principal left ideals generated by non-divisors of zero, into R.

8) The type of injectivity introduced by Maddox [10], but actually examined by Megibben [13], also generalizes classical divisibility: now R is replaced by projectives and principal ideals $R\lambda$ by finitely generated modules.

9) In [7], Helzer defines a divisibility in terms of the class of all natural inclusion maps from finitely generated, projective left ideals of R into R.

10) In [5], Goldman starts with a kernel functor σ, that is a functor from the category of left R-modules into itself such that (i) σM is a submodule of M; (ii) σf is the restriction of f to σM; (iii) if M' is a submodule of M, then $\sigma M' = M' \cap \sigma M$. He then defines his σ-injectivity in terms of the class of all natural inclusion maps $N \to M$, where $\sigma(M/N) = M/N$.

The number of examples of Γ-injectivity can obviously be increased. We mention two which have not as yet occurred in the literature.

11) The one which we call *cyclic injectivity* has Γ equal to the class of all monomorphisms α: $A \to B$ where both A and B are cyclic. This generalizes

the notion of divisibility discussed by Hattori where $B=R$. (Actually, cyclic injectivity has an equivalent form in terms of left ideals of R and R: A module M is cyclically injective if and only if each homomorphism ϕ: $L\to M$, where L is a left ideal of R and $L/\mathrm{Ker}\,\phi$ is cyclic, extends to R.)

12) Let Q be Utumi's maximal ring of left quotients of R [8], and let Γ contain only the natural inclusion map i: $R\to Q$. The corresponding Γ-injectivity, which we may call *Q-divisibility*, reduces to classical divisibility for Dedekind domains. This definition was suggested by that of Yun-dee Wei [16]: She called an R-module M divisible if $\mathrm{Hom}_R(Q, M/N)\neq 0$ for every proper submodule N of M. To see that Q-divisibility generalizes that of Yun-dee Wei, let $m\in M\backslash N$ and define α: $R\to M$ by $\alpha\,1=m$. If M is Q-divisible, there exists a map $\bar{\alpha}$: $Q\to M$ such that $\bar{\alpha}\,1=\alpha\,1$. Then $0\neq v\circ\bar{\alpha}\in\mathrm{Hom}_R(Q, M/N)$, where v: $M\to M/N$ is the canonical epimorphism.

13) A more explicit example of Γ-injectivity is the following. Let R be a local uniserial ring, i.e. a ring with a unique composition series

$$0\subset R\,a_1\subset R\,a_2\subset\cdots\subset R\,a_n=R. \tag{1}$$

It is well known that two cyclic R-modules have the same composition series if and only if they are isomorphic.

Proposition 2.1. *Let R be a local uniserial ring with composition series* (1) *and η: $R\,a_k\to R\,a_{k+t}$ the natural inclusion map for some k and t, where $1\leqq t\leqq n-k$ and $1\leqq k\leqq n$. Then $R\,a_j$ is injective with respect to η if and only if $j\geqq k+t$.*

Proof. Let $j<k+t$. Then there exists an f: $R\,a_k\to R\,a_s\subseteq R\,a_j$, where $s=\min(j,k)$ and $\mathrm{Ker}f=R\,a_{k-s}$ (put $a_0=0$). If f extends to $\bar{f}$: $R\,a_{k+t}\to R\,a_j$, then $\mathrm{Ker}\bar{f}=\mathrm{Ker}f$ implies $\mathrm{Im}\bar{f}$ is isomorphic to $R\,a_{t+s}$ contradicting $t+s>j$.

Let $j\geqq k+t$. As a local uniserial ring, R is quasi-Frobenius and hence self-injective [15]. Thus, given f: $R\,a_k\to R\,a_j\subset R$, there exists an $\bar{f}$: $R\,a_{k+t}\to R$ such that $\bar{f}\circ\alpha=f$. Obviously, $\mathrm{Im}\bar{f}\subseteq R\,a_{k+t}\subseteq R\,a_j$, proving that $R\,a_j$ is injective with respect to η. This completes the proof of Proposition 2.1.

Since now modules are direct sums of cyclic modules, we conclude:

Corollary 2.2. *Let R be a local uniserial ring with composition series* (1) *and let Γ be a set of inclusion maps η: $R\,a_k\to R\,a_{k+t}$ where $1\leqq t\leqq n-k$ and $1\leqq k\leqq n$. If m_0 is the maximum $k+t$ in Γ, then an R-module M is Γ-injective if and only if each of its cyclic direct summands has composition series of length $\geqq m_0$.*

III. Epimorphic Images of Injectives

Epimorphic images of injectives are not necessarily injective. In fact, the class of injectives is closed under epimorphic images exactly when R is left hereditary. For Γ-injectivity in general, it is rather difficult to give a satisfactory necessary and sufficient condition for closure with respect to epimorphic images. However, we do have one for closure under epimorphic images of injectives. In IV, we will explain why we believe this result to be the best possible with such a general definition as that of Γ-injectivity.

We prove two lemmas which immediately give the necessity statement. First recall that if given the exact sequence E in the top row and a map ξ in the diagram

$$\begin{tikzcd} E\colon 0 \arrow[r] & F \arrow[r, "\phi"] \arrow[d, "\xi"] & T \arrow[r, "\psi"] \arrow[d, "\sigma"] & C \arrow[r] \arrow[d, equal] & 0 \\ \xi E\colon 0 \arrow[r] & A \arrow[r, "\alpha"] & B \arrow[r] & C \arrow[r] & 0, \end{tikzcd} \tag{2}$$

then there exists an exact sequence ξE, as in the bottom row, making diagram (2) commutative. Notice that B is the pushout of (ϕ, ξ), see [14].

Lemma 3.1. *If F is projective in the commutative diagram* (2), *then every epimorphic image of an E-injective module is ξE-injective.*

Proof. We are given the commutative diagram (2) and an epimorphism $\gamma\colon N \to M$ with an E-injective module N. To show M is ξE-injective, let $f\colon A \to M$ be any homomorphism. Since F is projective and γ is epic, there exists a $\rho\colon F \to N$ such that $\gamma \circ \rho = f \circ \xi$. Also, there exists a $\tau\colon T \to N$ with $\tau \circ \phi = \rho$, because N is E-injective. Now B is the pushout of (ϕ, ξ) and $(\gamma \circ \tau) \circ \phi = \gamma \circ \rho = f \circ \xi$, so that there exists a homomorphism $\bar{f}\colon B \to M$ such that $\bar{f} \circ \alpha = f$. Consequently, M is ξE-injective. This completes the proof.

Let $0 \to A \xrightarrow{\alpha} B \to C \to 0$ be a Γ-sequence. Given an epimorphism $\xi\colon F \to A$, let

$$\xi_*\colon \operatorname{Ext}^1_R(C, F) \to \operatorname{Ext}^1_R(C, A)$$

be the induced group-homomorphism. We now prove:

Lemma 3.2. *Let $E\colon 0 \to A \to B \to C \to 0$ be an exact sequence and let $\xi\colon F \to A$ be an epimorphism with F projective. If $E \in \operatorname{Im} \xi_*$, then epimorphic images of injective modules are E-injective.*

Proof. $E \in \operatorname{Im} \xi_*$ means that there exists an exact sequence $E'\colon 0 \to F \to T \to C \to 0$ such that $\xi E' = E$. By Lemma 3.1, epimorphic images of E'-injectives and hence of injectives are E-injective.

From Lemmas 3.1 and 3.2 we obviously have:

Theorem 3.3. *If for each Γ-sequence $E\colon 0 \to A \to B \to C \to 0$ there exists an epimorphism $\xi_E\colon F_E \to A$ where F_E is projective and $E \in \operatorname{Im}(\xi_E)_*$, then epimorphic images of injectives are Γ-injective.*

An immediate consequence of this result is that the largest class Γ for which epimorphic images of injectives are Γ-injective can be obtained in the following way: Take all epimorphisms $\xi\colon F \to A$ with F projective and all exact sequences $E\colon 0 \to F \to T \to C \to 0$, and let all ξE form the class of Γ-sequences.

The proof of the converse of Theorem 3.3 is more complicated.

Theorem 3.4. *If epimorphic images of injective modules are Γ-injective, then for each Γ-sequence $E\colon 0 \to A \to B \to C \to 0$ and for each epimorphism $\xi\colon F \to A$, we must have $E \in \operatorname{Im} \xi_*$.*

Proof. Let E: $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ be any Γ-sequence and let $0 \to K \xrightarrow{\mu} F \xrightarrow{\xi} A \to 0$ be exact. Let I be an injective module such that there is a monomorphism η: $K \to I$, and choose a homomorphism $\bar{\eta}$: $F \to I$ such that $\eta = \bar{\eta} \circ \mu$.

Define N as the pushout of $(\xi, \bar{\eta})$ with maps θ: $A \to N$, γ: $I \to N$; here γ is epic since ξ is epic. By assumption, N must be Γ-injective; therefore, there exists a homomorphism $\bar{\theta}$: $B \to N$ with $\bar{\theta} \circ \alpha = \theta$. So we have constructed part of the following commutative diagram:

$$
\begin{tikzcd}
K \arrow[dr, "\mu"] & & & & & \\
& F \arrow[rr, dashed, "\phi"] \arrow[drr, "\xi"] \arrow[dd, "\bar{\eta}"'] & & T \arrow[dr, "\rho"] \arrow[ddll, dashed, "\tau"] & & \\
& & & A \arrow[r, "\alpha"] \arrow[dd, "\theta"] & B \arrow[r, two heads, "\beta"] \arrow[ddl, "\bar{\theta}"] & C \\
& I \arrow[drr, "\gamma"'] & & & & \\
& & & N & &
\end{tikzcd}
\tag{3}
$$

Now let T be the pullback of $(\gamma, \bar{\theta})$ with maps ρ: $T \to B$, τ: $T \to I$. Since γ is epic, ρ is also epic. Because $\bar{\theta} \circ \alpha \circ \xi = \theta \circ \xi = \gamma \circ \bar{\eta}$ and T is the pullback of $(\gamma, \bar{\theta})$, there exists a unique map ϕ: $F \to T$ with $\rho \circ \phi = \alpha \circ \xi$ and $\tau \circ \phi = \bar{\eta}$. We claim $0 \to F \xrightarrow{\phi} T \xrightarrow{\beta \circ \rho} C \to 0$ is exact and is mapped by ξ_* onto E. The proof of this claim is straightforward except for showing $\operatorname{Ker} \beta \circ \rho \subset \operatorname{Im} \phi$. Setting $L = \operatorname{Ker} \beta \circ \rho$, we have $\rho L \subset \operatorname{Ker} \beta = \operatorname{Im} \alpha$. Define $\bar{\alpha} = \alpha^{-1} \circ \rho$: $L \to A$ and $\bar{\tau} = \tau | L$; then $\gamma \circ \bar{\tau} = \theta \circ \bar{\alpha}$. Because η is monic and N is the pushout of $(\xi, \bar{\eta})$, F with maps ξ and $\bar{\eta}$ is a pullback of (θ, γ) (see [3], Prop. 2.53). Consequently, there exists a unique map σ: $L \to F$ with $\xi \circ \sigma = \bar{\alpha}$ and $\bar{\eta} \circ \sigma = \bar{\tau}$. It follows very easily that $\phi \circ \sigma = 1_L$; thus $x \in \operatorname{Ker} \beta \circ \rho$ implies $x = (\phi \circ \sigma) x \in \operatorname{Im} \phi$. Hence, $\operatorname{Ker} \beta \circ \rho \subset \operatorname{Im} \phi$, completing the proof of Theorem 3.4.

Summarizing Theorems 3.3 and 3.4, we have:

Corollary 3.5. *Epimorphic images of injective modules are Γ-injective if and only if for each Γ-sequence E: $0 \to A \to B \to C \to 0$, there exists an epimorphism ξ: $F \to A$ with F projective such that $E \in \operatorname{Im} \xi_*$.*

From the above proofs we see: In order that Corollary 3.5 be valid we need only consider epimorphisms of projectives onto $A \in \mathscr{A}$ rather than all epimorphisms. In fact, this can be easily proven without the aid of Theorems 3.3 and 3.4: Supposing that for each Γ-sequence E: $0 \to A \to B \to C \to 0$ and for each epimorphism ξ_F: $F \to A$ with F projective, $E \in \operatorname{Im}(\xi_F)_*$ holds, then we have $E \in \operatorname{Im}(\xi_G)_*$, for each Γ-sequence E and for each epimorphism ξ_G: $G \to A$.

IV. Closure Under Epimorphic Images

Corollary 3.5 does not answer the question: When is the class of Γ-injective modules closed under epimorphic images? Obviously, we cannot expect an informative answer in general. However, assuming certain, not too restrictive, properties of the Γ-sequences, we can obtain a satisfactory result.

In certain special cases (e.g. for injectivity, Megibben injectivity, Hattori divisibility), we know that epimorphic images of Γ-injectives are again Γ-injective if and only if the modules in $\mathscr{A}$ are projective (i.e. when R is left hereditary, left semi-hereditary, and a so-called left P.P. ring, respectively). One notices that in all of these cases Γ-injectivity can be defined in terms of Γ-sequences with $\mathscr{B}$ contained in the class $\mathscr{P}$ of projective modules. We generalize these results by proving the following theorem:

Theorem 4.1. *Let $\mathscr{B}$ be contained in the class $\mathscr{P}$. The class of Γ-injective modules is closed under epimorphic images if and only if $\mathscr{A}$ is also contained in $\mathscr{P}$.*

Proof. Assume the closure property and let $A \in \mathscr{A}$. To check the projectivity of A, it suffices to consider maps $\phi\colon A \to M$ and $\sigma\colon N \to M$, where N is injective and σ is epic (see Prop. 5.1 in Chapter I of [1]). M is Γ-injective by assumption; thus, taking an $\alpha\colon A \to B \in \Gamma$, there exists a map $\bar{\phi}\colon B \to M$ with $\bar{\phi} \circ \alpha = \phi$. Because σ is epic and B is projective, there exists a map $\psi\colon B \to N$ with $\sigma \circ \psi = \bar{\phi}$. So $\psi \circ \alpha\colon A \to N$ satisfies $\sigma \circ (\psi \circ \alpha) = \phi$ and A is thus projective. The converse is very easy, and it should be mentioned that it holds whether or not $\mathscr{B} \subset \mathscr{P}$.

For general Γ-injectivity, that is, without assuming $\mathscr{B} \subset \mathscr{P}$, we cannot force $\mathscr{A}$ to be contained in $\mathscr{P}$. [Take Γ-injectivity, where all R-modules are Γ-injective.]

From Corollary 1.5 in Chapter VII of [1], we see that if R is a principal ideal domain, injectivity is closed with respect to taking tensor products. For Γ-injectivity, we have the following result:

Theorem 4.2. *Let the class of Γ-injective R-modules be closed under R-epimorphic images and assume that $\mathscr{A}$ is contained in the class of finitely generated R-modules. Let S be an associative ring with unity. If the bimodule ${}_RM_S$ is Γ-injective as an R-module, then $M \otimes_S N$ is a Γ-injective, left R-module, for all left S-modules N.*

Proof. Recall that $M \otimes_S N$ becomes in a natural way a left R-module by defining $r(\sum m_i \otimes n_i) = \sum (r\, m_i) \otimes n_i$, for all $m_i \in M$, $n_i \in N$ and $r \in R$. There exists an S-epimorphism $g\colon F \to N$ for some free S-module F. Then $1_M \otimes g\colon M \otimes_S F \to M \otimes_S N$ is an R-epimorphism and $M \otimes_S F$ is R-isomorphic to a direct sum of copies of $M \otimes_S S$ which in turn is R-isomorphic to M. Since M is Γ-injective, $M \otimes_S F$ is a Γ-injective, left R-module by Proposition 1.3. Therefore, $M \otimes_S N$ as an R-epimorphic image of $M \otimes_S F$ is a Γ-injective, left R-module.

Notice that the types of divisibility and injectivity in Examples 7), 8), 9), and 12) are covered by Theorem 4.2, provided they are closed under epimorphic images.

V. Γ-Injective Submodules

Generally, modules do not contain maximal injective submodules or if they do, these submodules need not be unique. In fact, Matlis [11] showed that every R-module has a unique maximal injective submodule if and only if R is left Noetherian and left hereditary. We now consider the analogous questions for Γ-injectivity.

Proposition 5.1. *If $\mathscr{A}$ is contained in the class of finitely generated modules, then every module has a maximal Γ-injective submodule.*

Proof. By Proposition 1.3, Γ-injectivity is an inductive property, so Zorn's lemma can be applied.

Call a submodule M_Γ of M the *maximum Γ-injective submodule* of M if M_Γ is Γ-injective and contains every Γ-injective submodule of M.

Theorem 5.2. *Every R-module has a maximum Γ-injective submodule exactly if the class of Γ-injective R-modules is closed under* (i) *direct sums and* (ii) *epimorphic images.*

Proof. Assume every R-module has a maximum Γ-injective submodule. To show (i), we prove a little more. Let $\{X_i, \pi_i^j\}_{i,j\in I}$ be a direct system with limit X, where X_i is Γ-injective, for all $i\in I$, and π_i^j is monic, for all i, j with $i\leq j$. Then the canonical R-homomorphisms π_i: $X_i\to X$ are monic and so $\pi_i X_i$ is Γ-injective, for all $i\in I$. Since the $\pi_i X_i$ generate X and X has a maximum Γ-injective submodule, X must be Γ-injective.

To prove (ii), suppose ξ: $M\to N$ is an R-epimorphism where M is Γ-injective. By assumption, N has a maximum Γ-injective submodule N_Γ. Since M and N_Γ are Γ-injective, $M\oplus N_\Gamma$ is also Γ-injective. Now ϕ: $M\oplus N_\Gamma\to M\oplus N$ defined by $\phi(m,n)=(m,\xi(m)+n)$ with $m\in M$, $n\in N_\Gamma$, is monic, so that $\phi(M\oplus N_\Gamma)$ is Γ-injective in $M\oplus N$. It is readily checked that M, N_Γ, and $\operatorname{Im}\phi$ generate $M\oplus N$. So $M\oplus N$ is Γ-injective and thus N is Γ-injective.

The sufficiency of conditions (i) and (ii) is clear, since the union of all Γ-injective submodules of a module is an epimorphic image of their (outer) direct sum.

VI. Γ-Purity

In order to state our next theorem we need a definition of purity which generalizes that introduced by Cohn [2].

Definition 6.1. A submodule M of N is *Γ-pure* in N if for every commutative square of the form

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "\theta"'] & B \arrow[d, "\gamma"] \arrow[dl, dashed, "h"'] \\ M \arrow[r, "i"] & N \end{tikzcd} \qquad (\alpha\in\Gamma), \tag{4}$$

(with i the inclusion map) there exists an h: $B\to M$ with $h\circ\alpha=\theta$.

In the special case when Γ consists of all R-monomorphisms α: $A\to B$ with A finitely generated and B finitely generated free, Γ-purity is equivalent to Cohn's purity. In this case, pure submodules of Γ-injectives are Γ-injective and the Γ-injectives are precisely the pure submodules of injectives [10]. The corresponding conclusion is true in the general setting.

Theorem 6.2. (i) *Γ-pure submodules of Γ-injectives are Γ-injective.*

(ii) *An R-module M is Γ-injective if and only if it is Γ-pure in its injective hull $\hat{M}$.*

Proof. (i) Let M be Γ-pure in a Γ-injective N. Suppose $\alpha\colon A\to B$ is in Γ and $\theta\colon A\to M$ is any homomorphism. From the Γ-injectivity of N, we obtain the existence of a $\gamma\colon B\to M$ such that $\gamma\circ\alpha=\mathrm{i}\circ\theta$. As M is Γ-pure in N, there exists an $h\colon B\to M$ such that $h\circ\alpha=\theta$, proving M is Γ-injective.

(ii) Assume M is Γ-injective and the commutativity of the square in (4) holds with $N=\hat{M}$. Since M is Γ-injective, there exists an $h\colon B\to M$ such that $h\circ\alpha=\theta$. So M is Γ-pure in $\hat{M}$. The converse follows from (i).

VII. Γ-Injectivity and Γ-Torsionfreeness

As with the notion of purity, we will find it useful to have a definition of torsionfreeness in terms of the class Γ.

Definition 7.1. An R-module M is *Γ-torsionfree* if for each diagram

$$A\xrightarrow{\alpha}B\underset{h}{\overset{g}{\rightrightarrows}}M \quad \text{with } \alpha\in\Gamma, \tag{5}$$

$g\circ\alpha=h\circ\alpha$ implies $g=h$. Equivalently, M is Γ-torsionfree if and only if $\operatorname{Hom}_R(C,M)=0$ for each Γ-sequence $0\to A\xrightarrow{\alpha}B\to C\to 0$.

Notice that Γ-torsionfreeness reduces to the usual torsionfreeness when R is a commutative domain and $\Gamma=\{\alpha\colon R\lambda\to R\colon \lambda\in R\}$.

We state several properties of Γ-torsionfreeness for arbitrary rings. Their proofs are straightforward and therefore have been omitted.

(i) Submodules of Γ-torsionfree R-modules are Γ-torsionfree.

(ii) Extensions of Γ-torsionfree modules by Γ-torsionfree modules are again Γ-torsionfree.

(iii) $\prod_{i\in I} X_i$ (or $\bigoplus_{i\in I} X_i$) is Γ-torsionfree if and only if X_i is Γ-torsionfree, for all $i\in I$.

(iv) The class of Γ-torsionfree modules is closed under inverse limits.

(v) If $\mathscr{B}$ is contained in the class of finitely generated modules, then the class of all Γ-torsionfree modules is closed under direct limits of direct systems whose maps are monic.

(vi) Let $\mathscr{B}\subset\mathscr{P}$. If N is Γ-torsionfree and M is Γ-pure in N, then N/M is Γ-torsionfree.

(vii) If N/M is Γ-torsionfree, then M is Γ-pure in N.

(viii) Let $\mathscr{B}\subset\mathscr{P}$. If X_i is Γ-pure in X, for all $i\in I$, and X is Γ-torsionfree, then $X/(\bigcap_{i\in I} X_i)$ is Γ-torsionfree. In particular, $\bigcap_{i\in I} X_i$ is Γ-pure in X.

(ix) If N is Γ-torsionfree with maximum Γ-injective submodule N_Γ, then every Γ-pure submodule M of N has a maximum Γ-injective submodule, namely $M \cap N_\Gamma$.

Recall that a domain R is a right Ore domain if for all $x, x' \in R$, $x' \neq 0$, there exist $u \neq 0$, $v \in R$ with $xu = x'v$. In the next proposition, x^{-1} denotes the inverse of $x \in R$ in the field of right quotients Q of R.

Proposition 7.2. *Let R be a right Ore domain and M a left R-module. If M is injective and torsionfree with respect to all inclusions $R \to R x^{-1}$ with $0 \neq x \in R$, then M is injective with respect to the inclusion map $R \to Q$.*

Proof. It is not hard to see that as a left R-module, Q is the direct limit of the left R-modules $R x^{-1}$ with natural inclusions. The proof of Proposition 7.2 follows from the lemma below.

Lemma 7.3. *Let the exact sequence $0 \to A \xrightarrow{\alpha} B \to C \to 0$ be the direct limit of the Γ-sequences $0 \to A_i \xrightarrow{\alpha_i} B_i \to C_i \to 0$ $(i \in I)$. If M is Γ-injective and Γ-torsionfree, then M is injective with respect to α.*

Proof. Given f: $A \to M$, we have a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & A_i \arrow[r, "\alpha_i"] \arrow[d, "\pi_i^j"'] & B_i \arrow[r] \arrow[d, "\phi_i^j"'] & C_i \arrow[r] \arrow[d, "\tau_i^j"'] & 0 \\
0 \arrow[r] & A_j \arrow[r, "\alpha_j"] \arrow[d, "\pi_j"'] & B_j \arrow[r] \arrow[d, "\phi_j"'] & C_j \arrow[r] \arrow[d, "\tau_j"'] & 0 \quad (i \leq j) \\
0 \arrow[r] & A \arrow[r, "\alpha"] \arrow[d, "f"'] & B \arrow[r] \arrow[dl, dashed, "\bar{f}"] & C \arrow[r] & 0 \\
& M & & &
\end{tikzcd}$$

where π_i^j, ϕ_i^j, τ_i^j are the respective maps of the direct systems and π_j, ϕ_j, τ_j are the respective induced maps. Since M is Γ-injective, there exists a map f_i: $B_i \to M$ such that $f_i \circ \alpha_i = f \circ \pi_i$, for each $i \in I$. So for all i, j with $i \leqq j$, we have $f_j \circ \phi_i^j \circ \alpha_i = f_j \circ \alpha_j \circ \pi_i^j = f \circ \pi_j \circ \pi_i^j = f \circ \pi_i = f_i \circ \alpha_i$ whence $f_j \circ \phi_i^j = f_i$, because M is Γ-torsionfree. Thus, by the universal property of direct limits, there exists a unique map $\bar{f}$: $B \to M$ such that $\bar{f} \circ \phi_i = f_i$, for all $i \in I$. So for all $i \in I$, $\bar{f} \circ \alpha \circ \pi_i = f \circ \pi_i$. Hence, $\bar{f} \circ \alpha = f$ because the $\pi_i A_i$ generate A.

Next, we generalize the well-known result that over a principal ideal domain, inverse limits of torsionfree injectives are again injective.

Proposition 7.4. *The inverse limit of Γ-injective, Γ-torsionfree modules is again Γ-injective and Γ-torsionfree.*

Proof. Let $X = \varprojlim \{X_i, \pi_i^j\}_{i,j \in I}$ where X_i is Γ-injective and Γ-torsionfree, for all $i \in I$, and denote by π_i: $X \to X_i$ the induced maps. By property (iv) above, X is Γ-torsionfree. To show that X is Γ-injective, let $0 \to A \xrightarrow{\alpha} B \to C \to 0$ be a Γ-sequence and f: $A \to X$ be any map. Let $i, j \in I$ with $i \leqq j$. Since X_i is Γ-injective, for each i, there exists an f_i: $B \to X_i$ with $f_i \circ \alpha = \pi_i \circ f$. Then $\pi_i^j \circ f_j = f_i$ because $\pi_i^j \circ f_j \circ \alpha = \pi_i^j \circ \pi_j \circ f = \pi_i \circ f = f_i \circ \alpha$ and because X_i is Γ-torsionfree. By the universal property of inverse limits, there exists a unique map $\bar{f}: B \to X$ with $\pi_i \circ \bar{f} = f_i$, for all $i \in I$. This implies $\pi_i \circ \bar{f} \circ \alpha = \pi_i \circ f$, for all $i \in I$. Since π_i is the projection map, for all $i, \bar{f} \circ \alpha = f$. So $\bar{f}$ is as desired.

By Proposition 1.4 of Chapter VII of [1], if R is a commutative domain, the functor $\mathrm{Hom}_R(M, .)$: $N \to \mathrm{Hom}_R(M, N)$ preserves torsionfreeness, for every ${}_RM$. It also preserves the divisibility of torsionfree modules ${}_RN$. Our final theorem extends this result to Γ-injectivity.

Theorem 7.5. *Let Δ be a class of right R-monomorphisms and S an associative ring with 1. If ${}_SN_R$ is Δ-torsionfree and Δ-injective as a right R-module, then $H = \mathrm{Hom}_S({}_SM, {}_SN_R)$ is again a Δ-torsionfree and Δ-injective right R-module, for all ${}_SM$.*

Proof. Let $0 \to A_R \xrightarrow{\alpha} B_R \to C_R \to 0$ be a Δ-sequence. By hypothesis, $0 = \mathrm{Hom}_R(C, N) \to \mathrm{Hom}_R(B, N) \xrightarrow{\alpha^*} \mathrm{Hom}_R(A, N) \to 0$ is exact, and so α^* is an S-isomorphism. Because $\mathrm{Hom}_S(M, \mathrm{Hom}_R(B, N))$ is in a natural way group-isomorphic to $\mathrm{Hom}_R(B, H)$, $\mathrm{Hom}_R(B, H)$ is group-isomorphic to $\mathrm{Hom}_R(A, H)$ where the isomorphism is induced by α. Thus, H is necessarily Δ-injective. Since N is also Δ-torsionfree, $\mathrm{Hom}_R(C, N) = 0$. Therefore, $\mathrm{Hom}_R(C, H)$ is group-isomorphic to $\mathrm{Hom}_S(M, \mathrm{Hom}_R(C, N)) = 0$, and so H is Δ-torsionfree. (In proving H was Δ-torsionfree, the Δ-injectivity of N was not used.)

Corollary 7.6. *Let R be a commutative ring. The functor* $\mathrm{Hom}_R(M, .)$: $N \to \mathrm{Hom}_R(M, N)$ *takes Γ-injective, Γ-torsionfree modules into Γ-injective, Γ-torsionfree modules, for all ${}_RM$.*

Bibliography

1. Cartan, H., Eilenberg, S.: Homological algebra. Princeton: University Press 1956.
2. Cohn, P.M.: On the free products of associative rings. Math. Z. **71**, 380–389 (1959).
3. Freyd, P.: Abelian categories. New York, Evanston and London: Harper & Row 1964.
4. Fuchs, L.: Infinite Abelian groups. New York: Academic Press 1970.
5. Goldman, O.: Rings and modules of quotients. J. Algebra **13**, 10–47 (1969).
6. Hattori, A.: A foundation of torsion theory for modules over general rings. Nagoya Math. J. **17**, 147–158 (1960).
7. Helzer, G.: On divisibility and injectivity. Canadian J. Math. **18**, 901–919 (1966).
8. Lambek, J.: Lectures on rings and modules. Waltham, Mass.: Blaisdell Publ. Co. 1966.
9. Levy, L.: Torsion-free and divisible modules over non-integral-domains. Canadian J. Math. **15**, 132–151 (1963).
10. Maddox, B.H.: Absolutely pure modules. Proc. Amer. Math. Soc. **18**, 155–158 (1967).
11. Matlis, E.: Injective modules over Noetherian rings. Pacific J. Math. **8**, 511–528 (1958).
12. – Cotorsion modules. Mem. Amer. Math. Soc. Number 49, 1964.
13. Megibben, C.: Absolutely pure modules. Proc. Amer. Math. Soc. **26**, 561–566 (1970).
14. Mitchell, B.: Theory of categories. New York: Academic Press 1965.
15. Nakayama, T.: Note on uniserial and generalized uniserial rings. Imp. Acad. Tokyo **16**, 285–289 (1940).
16. Yun-dee Wei, D.: On the concept of torsion and divisibility for general rings. Illinois J. Math. **13**, 414–431 (1969).

Dr. Sheila R. O'Donnell
Department of Mathematics
Tulane University
New Orleans, La. 70118
USA

(Received April 19, 1971)

Math. Z. 124, 20–22 (1972)

The Stufe of Number Fields

IAN G. CONNELL

A result of Hilbert states that a totally positive element in an algebraic number field K is a sum of four squares of elements in K; the first published proof was given by Siegel [6]. In particular, if K has no real archimedean primes then -1 is totally positive and so can be written as a sum of four or fewer squares. The smallest number of squares possible in such a representation of -1 is called the stufe of K, and is denoted by $s=s(K)$. If K has a real prime one writes $s(K)=\infty$. Since $-1=x^2+y^2+z^2$ implies

$$-1=\left(\frac{xy+z}{y^2+z^2}\right)^2+\left(\frac{xz-y}{y^2+z^2}\right)^2,$$

$s\neq 3$, and therefore $s=1$, 2, 4 *or* ∞. Of course $s=1$ simply means that $i\in K$. (Pfister [5] has shown that for an arbitrary field, s is a power of 2 (or ∞), thus solving a long outstanding problem; cf. Kneser [4].)

Thus, given a number field K with no real primes and not containing i, there remains the problem of deciding whether $s=2$ or 4.

Theorem. *Let K be a number field with no real primes and not containing i. Then $s(K)=2$ if and only if for each prime P lying over 2 the local degree $[K_P:Q_2]$ is even; otherwise $s(K)=4$.*

Remark. If $(2)=P_1^{e_1}\dots P_g^{e_g}$ is the ideal factorization of 2 in K and P_i has absolute norm $NP_i=2^{f_i}$, then $[K_{P_i}:Q_2]=e_i f_i$, and so the requirement is that $2|e_i f_i$ for each i.

Proof. By Hasse's principle (cf. Ex. 4.8; all references of this form are to the exercises at the end of [1]), $s(K)=2$ if and only if $s(K_P)\leqq 2$ for each local completion K_P of K. The archimedean P are of no concern, so let $K_P\supseteq Q_p$. When $p\geqq 3$, $s(Q_p)\leqq 2$, hence $s(K_P)\leqq 2$: $-1\equiv x^2+y^2 \bmod p$ has a solution which lifts to a solution in Q_p by Hensel's lemma.

Thus let $K_P\supseteq Q_2$, say $[K_P:Q_2]=ef$ in the usual notation. By Ex. 4.5, $s(K_P)\leqq 2$ if and only if the quadratic norm residue symbol $(-1,-1)_P=1$. Dropping the subscript P, by Ex. 2.2, 2.5

$$\begin{aligned}(-1,-1)&=(-1,-1)(-1,2)=(-1,-2)\\&=(-1,-2)(3,-2)=(-3,-2).\end{aligned}$$

Now $-3=1+4(-1)$ and in the notation of Ex. 2.12,

$$\zeta=-1,\quad c=-1,\quad S(\bar c)=-f,\quad v(-2)=e,$$

so

$$(-1,-1)=(-3,-2)=(-1)^{ef}.$$

This completes the proof.

We conclude by discussing some examples which were worked out in an elementary way recently ([2, 3]). Let $K=Q(\zeta_p)$ where ζ_p is a primitive p-th root of unity, and p is an odd prime. Here $e=1$ and f is the exponent of 2 mod p (for each P lying over 2).

If $p=8k+7$ then $(2/p)=1$ whence $2^{(p-1)/2}=2^{4k+3}\equiv 1 \bmod p$, $f|4k+3$, f is odd and therefore by the theorem $s(K)=4$.

If $p=8k+5$, then $(2/p)=-1$ whence $2^{(p-1)/2}=2^{4k+2}\equiv -1 \bmod p$. It follows that $4|f$ and $s(K)=2$. Similarly if $p=8k+3$, $2|f$ and $s(K)=2$.

If $p=8k+1$ (a case left undecided in [3]) $s(K)=4$ or 2 according as the exponent of 2 mod p is odd or even, and both cases occur.

In the last example Tate suggested that the relative frequency of even versus odd should be computable using the Čebotarev density theorem. Here are the details.

For $n\geqq 3$ let $K_n=Q(\zeta_{2^n})$ and $L_n=K_n(2^{1/2^n})$; note that $K_n\cap Q(2^{1/2^n})=Q(\sqrt{2})$ and $L_n\colon Q=2^{2n-2}$. Then $G_n=\mathrm{Gal}(L_n/Q)$ consists of the elements σ_{ij} where $0\leqq i\leqq 2^n-1$, j is odd, $1\leqq j\leqq 2^n-1$, i is even when $j\equiv \pm 1 \bmod 8$, and i is odd when $j\equiv \pm 3 \bmod 8$, where, writing ζ for ζ_{2^n}, $\sigma_{ij}(2^{1/2^n})=\zeta^i\, 2^{1/2^n}$, and $\sigma_{ij}(\zeta)=\zeta^j$. There are exactly two elements in G_{n+1} satisfying $\sigma_{ij}\in\mathrm{Ker}(G_{n+1}\to G_n)$, $\sigma_{ij}\notin\mathrm{Ker}(G_{n+1}\to\mathrm{Gal}(K_{n+1}/Q))$, namely those with $i=0, 2^n$, $j=1+2^n$; they form a single conjugacy class. The rational primes p whose Frobenius automorphisms with respect to L_{n+1}/Q comprise this class are precisely those p which are of the form $2^n h+1$, h odd (because p splits in K_n but not in K_{n+1}) and for which 2 has a 2^n-th root mod p (because p splits in L_n); for these p, f is odd. Letting $n=3, 4, \ldots$, we obtain the totality of p for which f is odd and by Čebotarev's theorem the density of these primes is $\sum_{n=3}^{\infty} 2/2^{2n}=1/24$. (It is easy to justify the addition of these infinitely many densities.) Since the density of all primes $\equiv 1 \bmod 8$ is 1/4, the density of primes $\equiv 1 \bmod 8$ for which 2 has even exponent is $1/4-1/24=5/24$ and therefore the asymptotic ratio of the number of even cases to odd cases is 5 to 1.

Added in Proof. A different proof of the theorem by Fein, Gordon and Smith has since appeared in J. Number Theory **3**, 310–315 (1971).

References

1. Cassels, J.W.S., Fröhlich, A. (Editors): Algebraic number theory. London: Academic Press 1967.

2. Chowla, P.: On the representation of −1 as a sum of squares in a cyclotomic field. J. Number Theory **1**, 208–210 (1969).

3. — Chowla, S.: Determination of the stufe of certain cyclotomic fields. J. Number Theory **2**, 271–272 (1970).

4. Kneser, H.: Verschwindende Quadratsummen in Körpern. J.-ber. Deutsch. Math.-Verein. **44**, 143–146 (1934).
5. Pfister, A.: Zur Darstellung von —1 als Summe von Quadraten in einem Körper. J. London Math. Soc. **40**, 159–165 (1965).
6. Siegel, C.: Darstellung total positiver Zahlen durch Quadrate. Math. Z. **11**, 246–275 (1921).

Dr. I. Connell
Department of Mathematics
McGill University
Montreal 110, Quebec
Canada

(Received May 24, 1971)

Math. Z. 124, 23–36 (1972)

Choice Mappings of Certain Classes of Finite Sets

MORDECHAI LEWIN

1. Introduction

By a *choice mapping* we mean a function defined on some class of sets mapping a set of that class on some subset or possibly on some superset.

Let Ω be the class of all finite sets of distinct positive integers. Let each set be arranged in a monotonically increasing order. Define by Ω' the subclass of Ω of sets possessing more than one element. Let $\Sigma=(a_1, \ldots, a_n)$ be an arbitrary element of Ω'. It is shown subsequently in this paper that there is precisely one positive integer i_0, such that for all i, $1\leqq i<i_0$ we have

$$a_i-a_{i_0}-2(i-i_0)-1\geqq 0,$$

and for all i, $i_0<i\leqq n$, we have

$$a_i-a_{i_0}-2(i-i_0)\geqq 0.$$

Thus a choice function $\psi(\Sigma)=a_{i_0}$ is defined on Ω'. Instead of this choice function we consider the function

$$\varphi(\Sigma)=\Sigma-\{\psi(\Sigma)\}.$$

The function φ clearly maps Ω' into Ω.

Let n and N be integers such that $1<n\leqq N$. For $k\leqq N$ let $S(k, N)$ denote the class of all k-sets of distinct positive integers not exceeding N. If N is considered fixed as is frequently the case, $S(k, N)$ may simply be written as S_k. Let S_k be arranged in lexicographical order which will be denoted by $(S_k, \prec)$. Consider $(S_{n-1}, \prec)=(\sigma_1, \ldots, \sigma_{\binom{N}{n-1}})$, $(S_n, \prec)=(\Sigma_1, \ldots, \Sigma_{\binom{N}{n}})$. Define $f(\sigma_1)=\Sigma_1$. Clearly $\sigma_1\subset\Sigma_1$. Let now $\sigma_t\in S_{n-1}$. Assume that all predecessors to σ_t have been cared for. Then let $f(\sigma_t)$ assign to σ_t the least element of S_n containing σ_t and which has not yet been assigned to any predecessor of σ_t. If no such element exists in S_n, $f(\sigma_t)$ is labelled *undefined*, whereupon σ_{t+1} is considered. f is thus a mapping of a subset of S_{n-1} into S_n. The question naturally arises under what conditions is f defined on the whole of S_{n-1}. A necessary condition, which is quite trivial, is that $|S_{n-1}|\leqq|S_n|$ (where $|X|$ denotes the power of the set X), or equivalently, $n\leqq\frac{1}{2}(N+1)$. It turns out however as will be shown in this paper that this necessary condition is also sufficient. Of a particular interest is the case of equality where f becomes a bijection of S_{n-1} to S_n.

Both functions φ and f are choice mappings. The following rather surprising connection between them is revealed and of course being used in the course of establishing our results: φ turns out to be the inverse as it were of f; more precisely we establish the following

Main Result. *For $1 < n \leqq \frac{1}{2}(N+1)$, $\varphi f = I$, where I is the identity mapping defined on S_{n-1}; for $\frac{1}{2}(N+1) \leqq n \leqq N$, $f\varphi = I$ defined on S_n.*

For $n = \frac{1}{2}(N+1)$, both φ and f are bijections of S_n and S_{n-1} respectively, in which case φ and f are exact inverses of each other.

Considering S_{N-k} as *dual* sets to S_k, the connection between the functions φ, or f respectively, for dual pairs of sets is established.

To conclude our introduction one may use our Main Result in order to deduce quite easily the following theorem of Sperner:

Theorem. *Among the 2^m subsets of a given m-set there are at most $\binom{m}{[\frac{1}{2}m]}$ subsets such that no one contains the other* [2].

Proof. Introduce inclusion as partial order into the powerset P of the m-set S. An ordered subset of P is a *chain*. Put now $N = m$. Our main result implies that P may be partitioned into $\binom{m}{[\frac{1}{2}m]}$ mutually disjoint chains. On the other hand consider the class of all the $[\frac{1}{2}m]$-sets which are subsets of S. Their number is $\binom{m}{[\frac{1}{2}m]}$ and they do not contain each other. Applying now Dilworth's Theorem [1] to the partial order introduced we obtain immediately that the maximum is equal to the minimum and the result follows.

2. Preliminaries

A sequence of the form $x_1, x_2, \ldots, x_n$ will be written as $x_1,, x_n$.

An element or a subset of a sequence appearing immediately after the sequence with a $\wedge$-sign above it designates the deletion of that particular element or subset from the sequence.

If X and Y are sets, then $X - Y$ means set difference.

All sequences of distinct positive integers will be assumed to be monotonically increasing. The symbol $\prec$ between two such sequences denotes the lexicographical order relation.

No distinction will be made between a positive integer x and the one-element-set $\{x\}$; the curly brackets will simply be omitted.

3. The φ-Function

Let $a_1,, a_n$ be a strictly increasing sequence of positive integers, $n > 1$. Consider the following n-square matrix $[a_{ij}]$: For $i < j$, $a_{ij} = 1$ or -1 according as

$$a_j - a_i - 2(j-i)$$

is nonnegative or negative respectively.

For $i > j$, $a_{ij} = 1$ or -1 according as

$$a_j - a_i - 2(j-i) - 1$$

is nonnegative or negative respectively.

For $i=j$, $a_{ij}=0$.

This n-square matrix is clearly skew-symmetric. The structure of this matrix has an interest of its own as the following results reveal:

Lemma 3.1. *Let $A=[a_{ij}]$ be the associated matrix defined above. Let further* :, *j, k be any three positive integers not exceeding n and such that*

$$a_{ik}=1 \quad \text{and} \quad a_{jk}=-1, \tag{3.1}$$

then $a_{ij}=1$.

Proof. The conditions in (3.1) clearly imply that i, j, k be different.

Let $i<j<k$. We have

$$a_k-a_i-2(k-i)\geqq 0, \tag{3.2}$$

$$a_k-a_j-2(k-j)<0. \tag{3.3}$$

Subtracting (3.3) from (3.2) yields $a_j-a_i-2(j-i)>0$, so that $a_{ij}=1$.

Let now $i<k<j$. Then (3.1) implies

$$a_k-a_i-2(k-i)\geqq 0, \tag{3.4}$$

$$a_k-a_j-2(k-j)-1<0. \tag{3.5}$$

Again subtracting (3.5) from (3.4) we get $a_j-a_i-2(j-i)+1>0$. Then $a_j-a_i-2(j-i)\geqq 0$, so that again $a_{ij}=1$.

The same kind of reasoning will apply to the other different orders of i, j, k, so that Lemma 3.1 is proved.

Lemma 3.2. *Let $A=[a_{ij}]$ be the associated matrix defined above, then $a_{ij}=1$ if and only if the i-th row has more positive entries than the j-th row.*

Proof. We may assume $a_{ij}=1$ and $a_{ji}=-1$. Suppose there exists some k, $1\leqq k\leqq n$ such that $a_{ik}=-1$ and $a_{jk}=1$. Then by Lemma 3.1 $a_{ji}=1$, a contradiction. Then for every k for which $a_{jk}=1$ we have $a_{ik}=1$. Since a_{ij} belongs to the i-th row and a_{ji} to the j-th row it follows that the i-th row has more positive entries than the j-th row. The converse is obtained by assuming $a_{ij}=-1$ (and trivially $a_{ij}=0$). This proves Lemma 3.2.

Corollary 3.1. *In the associated matrix no two distinct rows (columns) have an equal number of positive entries.*

The proof is obvious by Lemma 3.2.

Corollary 3.2. *The associated matrix possesses a row (column) with exactly k positive entries for every k, $0\leqq k<n$.*

The proof is obvious by Corollary 3.1 and by the fact that there are n rows and at most $n-1$ positive entries in each row.

Returning to the original sequence we may now state

Theorem 3.1. *If $a_1, , a_n$ is a strictly increasing sequence of positive integers, then there exists precisely one term a_{i_0} for which $a_j-a_{i_0}-2(j-i_0)-1\geqq 0$ whenever $j<i_0$ and $a_j-a_{i_0}-2(j-i_0)\geqq 0$ whenever $j>i_0$.*

Proof. Consider the associated matrix. By Corollary 3.2 there exists precisely one row with $n-1$ positive entries. Let this row be the i_0-th row. The theorem then follows from the definition of the matrix.

The consequence of Theorem 3.1 suggests a choice function defined on the class of all finite sequences of more than one element of distinct positive integers $\psi(a_1,,a_n)=a_{i_0}$. Instead of the choice function we introduce the following choice mapping: $\varphi(a_1,,a_n)=(a_1,,a_n;\,\hat{a}_{i_0})$. The following quite general property of φ which will be of use later on is expressed by

Lemma 3.3. *Let $\varphi(\Sigma)=(\Sigma;\,\hat{a}_{i_0})$, where $i_0>1$, then*

$$a_{i_0}-1=a_{i_0-1}\in\Sigma.$$

Proof. Since $a_{i_0,i_0-1}=1$ by the definition of i_0, we have $a_{i_0-1}-a_{i_0}+2-1\geqq 0$ and hence $a_{i_0}\leqq a_{i_0-1}+1$. On the other hand clearly $a_{i_0}\geqq a_{i_0-1}+1$. Both inequalities imply equality, which is the lemma.

4. A Special Case

Let $N=2n-1>1$. Put $S(n,N)=S_n$, $S(n-1,N)=S_{n-1}$; $(S_n,\prec)=(\Sigma_1,,\Sigma_{\binom{N}{n}})$, $(S_{n-1},\prec)=(\sigma_1,,\sigma_{\binom{N}{n}})$. We now have

Lemma 4.1. *The restriction of φ to S_n is a bijection of S_n to S_{n-1}.*

Proof. Since we are concerned with the mapping of S_n we shall write φ although referring to the restriction of φ to S_n. φ is clearly singlevalued. We only have to show that it is also univalent or in other words that two distinct elements of S_n are mapped by φ on two distinct elements of S_{n-1}. Suppose to the contrary that Σ_u and Σ_v are both mapped onto the same element σ_k of S_{n-1}. Clearly $\Sigma_u\cap\Sigma_v=\sigma_k$. We may assume $\Sigma_u\prec\Sigma_v$. Then $\Sigma_v-\Sigma_u>\Sigma_u-\Sigma_v$.

Let $\Sigma_u=(a_1,,a_n)$, $\sigma_k=\varphi(\Sigma_u)=(a_1,,a_n;\,\hat{a}_{i_0})$. Consider Σ_u. We have

Case 1. $i_0=1$. Since $a_n\leqq 2n-1$, we have $a_n-2(n-1)\leqq 1$ and hence

$$a_n-a_1-2(n-1)\leqq 1-a_1\leqq 0. \tag{4.1}$$

On the other hand, since $a_{1_n}=1$, $i_0=1$,

$$a_n-a_1-2(n-1)\geqq 0. \tag{4.2}$$

(4.1) and (4.2) together imply equality so that

$$a_1=1 \tag{4.3}$$

and

$$a_n-a_1-2(n-1)=0. \tag{4.4}$$

Then $a_n=2n-1=N$.

Case 2. $i_0>1$. Then, by Lemma 3.3 $a_{i_0-1}=a_{i_0}-1$.

In both cases we have

$$a_{i_0}-a_{i_0-1}\equiv 1 \pmod N,$$

i_0-1 being taken modulo n.

Let $\Sigma_v=(b_1,,b_n)$ and $\varphi(\Sigma_v)=(\Sigma_v;\,\hat{b}_{j_0})$. Since $\Sigma_u\prec\Sigma_v$, Σ_v is obtained by deleting a_{i_0} from Σ_u and replacing it by b_{j_0}. We have $a_{i_0}<b_{j_0}$ and hence $i_0\leqq j_0$. We show that strict inequality is necessary.

Suppose $i_0=j_0$.

Case 1. $i_0=1$. Then $j_0=1$ and hence by (4.3) $a_{i_0}=1$, $b_{j_0}=1=a_{i_0}$, a contradiction.

Case 2. $i_0>1$. Then $j_0>1$ and hence by Lemma 3.3 we have $b_{j_0-1}=b_{j_0}-1\in\Sigma_v$. But then $b_{j_0}-1\neq a_{i_0}$, so $b_{j_0}-1\in\Sigma_u$ and $a_{i_0}<b_{j_0}-1$. Then $b_{j_0-1}=a_{j_0}>a_{i_0}$ and so $i_0<j_0$, again a contradiction, so that $i_0<j_0$.

Comparing the elements of Σ_u and Σ_v we have

$$\text{for } i<i_0 \quad \text{or} \quad i>j_0 \qquad b_i=a_i,$$

$$\text{for } i_0\leqq i<j_0 \qquad b_i=a_{i+1}.$$

Let now $i_0>1$. Then $b_{i_0-1}-b_{j_0}-2[(i_0-1)-j_0]-1\geqq 0$ by the choice of j_0.

Then $b_{j_0}-b_{i_0-1}-2[j_0-(i_0-1)]<0$. But $b_{j_0}=a_{j_0}+1$ by Lemma 3.3 and hence $b_{i_0-1}=a_{i_0-1}=a_{i_0}-1$ so that

$$a_{j_0}+1-a_{i_0}+1-2[j_0-(i_0-1)]=a_{j_0}-a_{i_0}-2(j_0-i_0)<0,$$

a contradiction because of the choice of i_0.

If $i_0=1$, then $a_n=2n-1=N$ by (4.3) and (4.4), so that $b_n=a_n=2n-1$. By the choice of $i_0=1$ we have $a_{j_0}-1-2(j_0-1)\geqq 0$, or $a_{j_0}\geqq 2j_0-1$. Then $b_{j_0}=a_{j_0}+1\geqq 2j_0$. By the choice of j_0 we have $b_n-b_{j_0}-2(n-j_0)\geqq 0$. On the other hand $b_n-b_{j_0}-2(n-j_0)\leqq 2n-1-2j_0-2(n-j_0)<0$, a contradiction. This completes the proof of Lemma 4.1.

As a by-product of Lemma 4.1 we may state

Lemma 4.2. *If* $(a_1,,a_n)$ *is such that* $a_n\leqq 2n-1$, $n>1$, *and* $\varphi(a_1,,a_n)=(a_1,,a_n;\,\hat{a}_1)$, *then* $a_1=1$.

This is the result expressed in (4.3).

Recalling the function f which maps some subset of S_{n-1} into S_n, the criterion being lexicographical order as described in Section 1 we may now state

Theorem 4.1. *φf is the identity mapping of* $S_{n-1}=S(n-1,2n-1)$, $n>1$, *onto itself.*

Proof. Since φ is a bijection by Lemma 4.1, φ^{-1} is a bijection of S_{n-1} to S_n. We shall show $\varphi^{-1}\equiv f$. Suppose φf is not the identity. Put $(S_{n-1},\prec)=\{\sigma_i\}$, $(S_n,\prec)=\{\Sigma_j\}$. Clearly $f(\sigma_1)=\Sigma_1$ and $\varphi f(\sigma_1)=\varphi(\Sigma_1)=\sigma_1$.

Let t be the least positive integer such that $\varphi f(\sigma_t)\neq\sigma_t$.

Case 1. $f(\sigma_t)$ is undefined. Then $\varphi^{-1}(\sigma_t)=f(\sigma_i)$ so that $\sigma_t=\varphi f(\sigma_i)$ for some $\sigma_i\prec\sigma_t$ and such that $\sigma_i=\varphi^{-1}(\sigma_t)$. But then, by the minimal property of t, $\varphi f(\sigma_i)=\sigma_i\neq\sigma_t$, a contradiction.

Case 2. $f(\sigma_t)$ is defined but $\varphi f(\sigma_t)\neq\sigma_t$. Then $f(\sigma_t)\neq\varphi^{-1}(\sigma_t)$. This implies

$$f(\sigma_t)\prec\varphi^{-1}(\sigma_t). \tag{4.5}$$

Suppose $\varphi f(\sigma_t)=\sigma_i\prec\sigma_t$ for some $i<t$. Applying φ^{-1} to both sides we get $f(\sigma_t)=\varphi^{-1}(\sigma_i)=f(\sigma_i)$, the second equality being due to the induction hypothesis. This is an obvious contradiction. Then

$$\varphi f(\sigma_t)\succ\sigma_t. \tag{4.6}$$

Put $f(\sigma_t)-\varphi f(\sigma_t)=z$, $f(\sigma_t)=(\sigma_t,x)$, $\varphi^{-1}(\sigma_t)=(\sigma_t,y)$. (4.6) implies $z<x$. Suppose $y<x$. Then $\varphi^{-1}(\sigma_t)\prec f(\sigma_t)$ so that there exists $\sigma_i\prec\sigma_t$ such that $f(\sigma_i)=\varphi^{-1}(\sigma_t)$.

But then $\varphi f(\sigma_i)=\sigma_i \neq \sigma_t$, a contradiction and hence $x<y$. Then

$$z<x<y. \tag{4.7}$$

Put $\varphi^{-1}(\sigma_t)=(a_1,,a_n)$, $f(\sigma_t)=(b_1,,b_n)$. Let further $y=a_{i_0}$, $x=b_{j_1}$, $z=a_{i_1}=b_{j_2}$. (4.7) clearly implies $j_2=i_1<i_0$ and $j_2<j_1$. Then $i_0>1$ so that by Lemma 3.3 $y-1\in\varphi^{-1}(\sigma_t)$ and hence $y-1\in\sigma_t$; but $x\notin\sigma_t$ which together with (4.7) implies

$$z<x<y-1. \tag{4.8}$$

We now have $y-1=a_{i_0-1}=b_{i_0}$. The corresponding elements of the associated matrices of $\varphi^{-1}(\sigma_t)$ and $f(\sigma_t)$ respectively are $a_{i_0 i_1}=1$ and $b_{i_1 i_0}=1$ so that

$$z-y-2(i_1-i_0)-1\geqq 0, \tag{4.9}$$

$$y-1-z-2(i_0-i_1)\geqq 0. \tag{4.10}$$

Adding (4.9) and (4.10) leads us to a contradiction. This completes the proof of Theorem 4.1.

It follows that f is a bijection mapping $S(n-1,2n-1)$ onto $S(n,2n-1)$.

5. The General Case

In what follows we shall deal with the cases $n<\frac{1}{2}(N+1)$ and $n>\frac{1}{2}(N+1)$. We first proceed to the extension of Theorem 4.1 to the case $n<\frac{1}{2}(N+1)$ which we state as

Theorem 5.1. *If* $1<n\leqq\frac{1}{2}(N+1)$, *then* φf *is the identity mapping defined on* S_{n-1}.

For the proof we suppose the contrary and for this purpose we choose n to be the largest integer for which the conditions of Theorem 5.1 hold and the conclusion fails.

We first prove the auxiliary

Lemma 5.1. *Let* $1<n\leqq\frac{1}{2}(N+1)$. *Then* $\varphi(S_n)=S_{n-1}$.

In other words φ is a surjection of S_n onto S_{n-1}. Lemma 5.1 clearly becomes superfluous once Theorem 5.1 is established.

Proof of Lemma 5.1. Let n be the largest integer for which the lemma fails. Suppose there exists an element σ of S_{n-1} which is not an image of any element of S_n under the mapping φ. It is easy to see that if $n<\frac{1}{2}N$, then $n+1\leqq\frac{1}{2}(N+1)$. Therefore if $n<\frac{1}{2}N$, then since n was chosen largest,

$$\varphi[S(n+1,N)]=S(n,N). \tag{5.1}$$

If $n=\frac{1}{2}N$, then $n+1=\frac{1}{2}[(N+1)+1]$ and hence by Theorem 4.1

$$\varphi[S(n+1,N+1)]=S(n,N+1). \tag{5.2}$$

Let $\alpha=(a_1,,a_{n-1})\notin\varphi(S_n)$. For $n<\frac{1}{2}N$ let b_0 be the largest integer not exceeding N which is not an element of α. For $n=\frac{1}{2}N$, put $b_0=N+1$. Consider now $\beta=(b_1,,b_n)=(\alpha,b_0)$. Since β is an n-sequence it is by (5.1) and (5.2) an image of some $(n+1)$-sequence from $S(n+1,N)$ or $S(n+1,N+1)$ according as

$n<\frac{1}{2}N$ or $n=\frac{1}{2}N$ respectively. Denote the $(n+1)$-sequence by $\gamma=(c_1, , c_{n+1})=(\alpha, b_0, c_0)$. Put $c_0=c_{i_0}$, $b_0=c_{i_1}$. The choice of b_0 implies

$$i_0<i_1 \tag{5.3}$$

and $c_{i_0}<c_{i_1}\leqq N+1$. Considering the i_0-th row of the associated matrix for γ we have $c_{i_0 j}=1$ for all $j\neq i_0$.

Case 1. $j>i_0$. Then

$$c_j-c_{i_0}-2(j-i_0)\geqq 0. \tag{5.4}$$

If b_0 is deleted from γ, then $j-i_0$ will decrease by one if $j>i_1$ and will remain unchanged if $i_0<j<i_1$. Inequality (5.4) will therefore remain intact with possibly some suitable change of the index j.

Case 2. $j<i_0$. Since $c_{j i_0}=-1$, we have

$$c_{i_0}-c_j-2(i_0-j)<0. \tag{5.5}$$

The deletion of b_0 will not affect (5.5).

Cases 1 and 2 taken together imply $c_{i_0 i}=1$ for all the elements of the i_0-th row of the associated matrix of $(\gamma; \hat{b}_0)$ with $i\neq i_0$. Then by the definition of φ, $\varphi(\gamma; \hat{b}_0)=(\gamma; \hat{b}_0, \hat{c}_0)=\alpha$. This brings us to a contradiction and thus proves the lemma.

Again let $1<n\leqq\frac{1}{2}N$. If $\sigma_i\in(S_{n-1}, \prec)$, define b_i as the largest integer $\leqq N$ which does not belong to σ_i if $n<\frac{1}{2}N$; if $n=\frac{1}{2}N$, put $b_i=N+1$. We then have

Lemma 5.2. *For* $1<n\leqq\frac{1}{2}N$, $f(\sigma_i)=[f(\sigma_i, b_i); \hat{b}_i]$ *for all i.*

Proof. Choose n such that Theorem 5.1 holds for every n' for which $n<n'\leqq\frac{1}{2}(N+1)$ and choose t such that for all i, $i<t$ we have $\varphi f(\sigma_i)=\sigma_i$ and $f(\sigma_i)=[f(\sigma_i, b_i); \hat{b}_i]$. Assuming these conditions we prove $f(\sigma_t)=[f(\sigma_t, b_t); \hat{b}_t]$. Suppose the contrary. Put $N_0=N$ if $n<\frac{1}{2}N$ and $N_0=N+1$ if $n=\frac{1}{2}N$. Clearly $\sigma_t\subset[f(\sigma_t, b_t); \hat{b}_t]$.

Case 1. There exists $i<t$ such that

$$f(\sigma_i)=[f(\sigma_t, b_t); \hat{b}_t]. \tag{5.6}$$

Then

$$\sigma_i\subset[f(\sigma_t, b_t); \hat{b}_t]. \tag{5.7}$$

By hypothesis we have

$$f(\sigma_i)=[f(\sigma_i, b_i); \hat{b}_i]. \tag{5.8}$$

(5.6) and (5.8) imply

$$[f(\sigma_i, b_i); \hat{b}_i]=[f(\sigma_t, b_t); \hat{b}_t]. \tag{5.9}$$

$\sigma_i\prec\sigma_t$ and (5.7) imply $b_i\geqq b_t$.

Suppose $b_i>b_t$. If $n=\frac{1}{2}N$, then $b_i=b_t=N+1$, and hence $n<\frac{1}{2}N$. By (5.9) $b_i\notin[f(\sigma_t, b_t); \hat{b}_t]$ so that $b_i\notin f(\sigma_t, b_t)$ and hence $b_i\notin\sigma_t$ which is a contradiction since b_t was chosen maximal not in σ_t. Then $b_i=b_t$.

(5.9) may be written as

$$[f(\sigma_i, b_t); \hat{b}_t]=[f(\sigma_t, b_t); \hat{b}_t].$$

It follows that $f(\sigma_i, b_t) = f(\sigma_t, b_t)$ and hence $(\sigma_i, b_t) = (\sigma_t, b_t)$ which is a contradiction. This completes Case 1.

Case 2. $f(\sigma_t) = (\sigma_t, c) \prec (\sigma_t, d) = [f(\sigma_t, b_t); \hat{b}_t]$. Then $f(\sigma_t, b_t) = (\sigma_t, d, b_t), d < b_t$.

If $\varphi(\sigma_t, c, b_t) \succ \varphi(\sigma_t, d, b_t) = (\sigma_t, b_t)$, then there exists $\alpha \in S(n, N_0)$ such that $\alpha \prec (\sigma_t, b_t)$ and such that $f(\alpha) = (\sigma_t, c, b_t)$. Then by hypothesis $\varphi(\sigma_t, c, b_t) = \alpha$, a contradiction, so that

$$\varphi(\sigma_t, c, b_t) \prec \varphi(\sigma_t, d, b_t) = (\sigma_t, b_t).$$

We therefore have

$$\varphi(\sigma_t, c, b_t) = (\sigma_t, c, b_t; \hat{e}), \qquad e > c.$$

By assumption $f(\sigma_t, b_t) = (\sigma_t, d, b_t)$. This implies the existence of $\Sigma \in S(n, N_0)$ such that $\Sigma \prec (\sigma_t, b_t)$ and $f(\Sigma) = (\sigma_t, c, b_t)$. Then $\varphi(\sigma_t, c, b_t) = \Sigma$ and hence $\Sigma = (\sigma_t, c, b_t; \hat{e})$.

Subcase 2.1. $e < b_t$. Put $(\sigma_t, c; \hat{e}) = \sigma_i \prec \sigma_t$. Then

$$\begin{aligned} f(\sigma_i) &= [f(\sigma_i, b_i); \hat{b}_i] = [f(\sigma_i, b_t); \hat{b}_t] \\ &= [f(\sigma_t, c; \hat{e}; b_t); \hat{b}_t] = [f(\Sigma); \hat{b}_t] = (\sigma_t, c, b_t; \hat{b}_t) = (\sigma_t, c), \end{aligned}$$

a contradiction.

Subcase 2.2. $e = b_t$. Then $f(\sigma_t, c) = (\sigma_t, c, b_t)$, and hence $\varphi(\sigma_t, c, b_t) = \varphi f(\sigma_t, c) = (\sigma_t, c)$. Then $b_t = N_0$. We have $a_{n+1,j} = 1$ for all $j \leqq n$ owing to the n inequalities for the respective entries of the associated matrix. If c is now replaced by d, all the n inequalities are preserved and hence $\varphi(\sigma_t, d, b_t) = (\sigma_t, d)$, a contradiction.

Subcase 2.3. $e > b_t$. Then, as in Subcase 2.2, we have $e = N = N_0$. Thus $f(\sigma_t, c, b_t; \hat{N}) = (\sigma_t, c, b_t)$. Then $\varphi(\sigma_t, c, b_t) = (\sigma_t, c, b_t; \hat{N})$. By the same reasoning as in Subcase 2.2 we have $\varphi(\sigma_t, d, b_t) = (\sigma_t, d, b_t; \hat{N})$, a contradiction.

The hypothetical form of Lemma 5.2 is thus proved.

We now claim that under the same hypothetical conditions as Lemma 5.2 we have

$$\varphi[f(\sigma_t, b_t); \hat{b}_t] = [\varphi f(\sigma_t, b_t); \hat{b}_t]. \tag{5.10}$$

Proof. Put $f(\sigma_t, b_t) = (\sigma_t, x, b_t)$. Clearly $x < b_t$. Let $x = a_{i_0} \in (\sigma_t, x, b_t)$ and $x = b_{j_0} \in (\sigma_t, x)$. Clearly $j_0 = i_0$. The defining inequalities for $b_{j_0 j}$ for all j are preserved and hence $\varphi(\sigma_t, x) = \sigma_t$. This proves (5.10). By Lemma 5.2 the left hand side of (5.10) is $\varphi f(\sigma_t)$. Thus we have $\varphi f(\sigma_t) = \sigma_t$ which is Theorem 5.1.

Passing from t to $t+1$ we prove Lemma 5.2 and then (if necessary) Theorem 5.1 in turn. After having exhausted all the t's we pass from n to $n-1$. Having exhausted all the necessary n's in decreasing order we have proved Lemma 5.2 and Theorem 5.1 simultaneously.

Consider the following mapping of the class of n-sets onto the class of $(N-n)$-sets: For $0 < n < N$, let $\sigma \in S_n$. Then $x \in \bar{\sigma}$ if and only if $N+1-x \notin \sigma$. By definition $\bar{\sigma} \in S_{N-n}$ and such a mapping is clearly a bijection of S_n onto S_{N-n}. We now have the following

Lemma 5.3. *Let* $\frac{1}{2}(N+1) \leqq n < N$ *and let* $\Sigma \in S_n$, *then if* $\varphi(\Sigma) = \sigma$, *then* $\varphi(\bar{\sigma}) = \bar{\Sigma}$.

Proof. Let $\Sigma = (\sigma, x)$; then

$$\bar{\sigma} = (1, , N; N + \hat{1} - a_i) \quad \text{for all } i \neq i_0,\ a_{i_0} = x,$$
$$\bar{\Sigma} = (1, , N; N + \hat{1} - a_i) \quad \text{for all } i.$$

It follows that $\bar{x} = \bar{\sigma} - \bar{\Sigma} = N + 1 - x$. Let $\bar{x} = \bar{a}_{j_0} = \psi(\bar{\sigma})$. Put

$$a_{i_0+t} - a_{i_0} = t + \lambda_t, \qquad t \geqq 0, \tag{5.11}$$

where λ_t denotes the number of elements between a_{i_0} and a_{i_0+t} which are not in Σ. (For $t=0$ we have trivially $\lambda_0 = 0$.) The property of a_{i_0} implies

$$a_{i_0+t} - a_{i_0} - 2t \geqq 0. \tag{5.12}$$

Substituting from (5.11) we get $\lambda_t - t \geqq 0$, or

$$\lambda_t \geqq t. \tag{5.13}$$

Let $x + r \notin \Sigma$. Then $N + 1 - x - r \in \bar{\Sigma}$. Let k be the number of elements of Σ greater than x and less than $x + r$. Put $r = k + \mu_k$. Clearly $\mu_k > \lambda_k$, but $\lambda_k \geqq k$ and hence $\mu_k \geqq k + 1$. We now have

$$\bar{a}_{j_0} - \mu_k - \bar{a}_{j_0} + 2\mu_k - 1 = -r + 2\mu_k - 1 = \mu_k - k - 1 \geqq 0. \tag{5.14}$$

Let now $a_{i_0-t} \in \Sigma$, $t > 0$. Put $a_{i_0} - a_{i_0-t} = t + \lambda_t$.

As in (5.12) we now have

$$a_{i_0-t} - a_{i_0} + 2t - 1 \geqq 0. \tag{5.15}$$

Substitution leads to $t - \lambda_t - 1 \geqq 0$, or

$$\lambda_t \leqq t - 1. \tag{5.16}$$

Let now $x - r \notin \Sigma$. Then $N + 1 - x + r = \bar{x} + r \in \bar{\Sigma}$. Again put $r = k + \mu_k$ where k is the number of elements of Σ less than x and greater than $x - r$. We claim

$$\mu_k \leqq k. \tag{5.17}$$

We first show

$$a_{i_0} \leqq 2i_0 - 1. \tag{5.18}$$

Suppose $a_{i_0} > 2i_0 - 1$. Then for t such that $0 \leqq t \leqq n - i_0$ we have

$$a_{i_0+t} = a_{i_0} + t + \lambda_t \geqq a_{i_0} + 2t > 2(i_0 + t) - 1,$$

the first inequality being derived from (5.13). Then $a_n > 2n - 1$ so that $N > 2n - 1$ and $n < \frac{1}{2}(N+1)$, a contradiction, so that (5.18) holds.

Since $x - r \notin \Sigma$ we have $x > 1$. Suppose $i_0 = 1$. Then $x = a_{i_0} \leqq 2i_0 - 1 = 1$, a contradiction. Then $i_0 > 1$, and hence by Lemma 3.3 $a_{i_0-1} = x - 1$, so that $k > 0$. We now have

Case 1. $k < i_0 - 1$. Then $a_{i_0-k-1} < x - r$ and so $\mu_k \leqq \lambda_{k+1}$; then $\mu_k \leqq k + 1 - 1 = k$ by (5.16).

Case 2. $k=i_0-1$. We then have $\mu_k+i_0=a_{i_0}\leqq 2i_0-1$ and hence $\mu_k\leqq i_0-1=k$. Thus

$$\mu_k\leqq k \quad \text{in general}.$$

Put now $N+1-x+r=\bar{a}_j$. We then have

$$\bar{a}_j-\bar{a}_{j_0}-2\mu_k=r-2\mu_k=k-\mu_k\geqq 0. \tag{5.19}$$

(5.14) and (5.19) together imply that $\varphi(\bar{\sigma})=\bar{\Sigma}$. This completes the proof of the lemma.

Lemma 5.3 admits of the following corollary.

Corollary 5.1. *If n is such that $\frac{1}{2}(N+1)\leqq n\leqq N$, then φ is univalent in S_n.*

Proof. Suppose $\Sigma\neq\Sigma'$ and $\varphi(\Sigma)=(\Sigma')=\sigma$. Then by Lemma 5.3 $\varphi(\bar{\sigma})=\bar{\Sigma}$ and $\varphi(\bar{\sigma})=\bar{\Sigma}'$, an obvious contradiction. This proves the corollary.

It should be noted that the restriction as to n is necessary both for the lemma and the corollary, as the following example shows: Choose $N=9$, $n=4$. Put $\Sigma=(2,4,6,8)$, $\Sigma'=(1,4,6,8)$. Then

$$\varphi(2,4,6,8)=(4,6,8)=\sigma; \qquad \bar{\sigma}=(1,3,5,7,8,9),$$

but

$$\varphi(\bar{\sigma})=(1,3,5,7,8)\neq(1,3,5,7,9)=\bar{\Sigma}.$$

As a counterexample for the corollary we have

$$\varphi(\Sigma)=\varphi(\Sigma')=\sigma.$$

Let $\alpha,\beta\in S_n$ for some $n<N$. If $\alpha\prec\beta$, then it is not generally true that $\bar{\alpha}\prec\bar{\beta}$, as the following example shows: Put $N=9$, $n=5$, $\alpha=(1,3,7,8,9)$, $\beta=(1,4,5,6,7)$; we then have $\bar{\alpha}=(4,5,6,8)$, $\bar{\beta}=(1,2,7,8)$, so that $\alpha\prec\beta$ but $\bar{\beta}\prec\bar{\alpha}$.

The following simple lemma may however be stated:

Lemma 5.4. *Let $\alpha,\beta\in S_n$ and let $|\alpha-\beta|=1$. Then $\alpha\prec\beta$ if and only if $\bar{\alpha}\prec\bar{\beta}$.*

Proof. Let $\alpha\prec\beta$. Put $\alpha-\beta=x$, $\beta-\alpha=y$. Then $x<y$. We have $x\notin\beta$ and $y\notin\alpha$. Put now $\bar{x}=N+1-x$, $\bar{y}=N+1-y$. It is easily seen that $\bar{x}=\bar{\beta}-\bar{\alpha}$ and $\bar{y}=\bar{\alpha}-\bar{\beta}$. $x<y$ implies $\bar{y}<\bar{x}$, so that $\bar{\alpha}\prec\bar{\beta}$, which proves the first part of the lemma. The second part is obviously obtained by interchanging the roles of the pairs (α,β) and $(\bar{\alpha},\bar{\beta})$ so that the lemma is proved.

We now have

Lemma 5.5. *Let $\sigma\in S(n-1,N)$, where $1<\frac{1}{2}(N+1)\leqq n$. Then $f(\sigma)=\Sigma$ implies $f(\bar{\Sigma})=\bar{\sigma}$.*

Proof. Suppose the lemma is false. Clearly $f(\sigma_1)=\Sigma_1$ and also $f(\bar{\Sigma}_1)=\bar{\sigma}_1$. Let s be the least integer for which $f(\sigma_s)=\Sigma_t$ and $f(\bar{\Sigma}_t)\neq\bar{\sigma}_s$.

Case 1. $f(\bar{\Sigma}_t)=\bar{\sigma}_i\prec\bar{\sigma}_s$. By Theorem 5.1 we have

$$\varphi(\bar{\sigma}_i)=\bar{\Sigma}_t. \tag{5.20}$$

$\bar{\sigma}_i, \bar{\sigma}_s \in S_{N+1-n}$ and $\bar{\Sigma}_t \subset \bar{\sigma}_i \cap \bar{\sigma}_s$ imply $|\bar{\sigma}_s - \bar{\sigma}_i| = 1$, so that $\bar{\sigma}_i \prec \bar{\sigma}_s$ implies $\sigma_i \prec \sigma_s$ by Lemma 5.4. Then $i<s$. Since $f(\sigma_s)=\Sigma_t$ we have $f(\sigma_i)=\Sigma_u \prec \Sigma_t$. $i<s$ implies $f(\bar{\Sigma}_u)=\bar{\sigma}_i$ and hence, by Theorem 5.1, $\varphi(\bar{\sigma}_i)=\bar{\Sigma}_u \neq \bar{\Sigma}_t$, a contradiction to (5.20).

Case 2. There exists $\bar{\Sigma}_v \prec \bar{\Sigma}_t$ such that $f(\sigma_s)=\Sigma_t$ and $f(\bar{\Sigma}_v)=\bar{\sigma}_s$. As in Case 1, Lemma 5.4 may be applied to $\bar{\Sigma}_v$ and $\bar{\Sigma}_t$, so that $\Sigma_v \prec \Sigma_t$. Since $\sigma_s \subset \Sigma_v \cap \Sigma_t$, there exists $\sigma_i \prec \sigma_s$ such that $f(\sigma_i)=\Sigma_v$. Then by hypothesis $f(\bar{\Sigma}_v)=\bar{\sigma}_i \neq \bar{\sigma}_s$, a contradiction. This proves Lemma 5.5.

Using the fact that $n<\frac{1}{2}(N+1)$ implies $N-n+1>\frac{1}{2}(N+1)$ we may now drop the restriction on n in Lemma 5.5.

Theorem 5.2. *Let* $\sigma \in S_{n-1}$, $n<N$ *for some* N. *Then* $f(\sigma)=\Sigma$ *if and only if* $f(\bar{\Sigma})=\bar{\sigma}$.

Proof. We assume $n<\frac{1}{2}(N+1)$ and prove

$$f(\sigma)=\Sigma \quad \text{implies} \quad f(\bar{\Sigma})=\bar{\sigma}.$$

Again we suppose the theorem false and choose s accordingly so that we have $f(\sigma_s)=\Sigma_t$ and $f(\bar{\Sigma}_t)\neq\bar{\sigma}_s$ with s minimal.

Case 1. There exists $\bar{\Sigma}_u \prec \bar{\Sigma}_t$ such that $f(\bar{\Sigma}_u)=\bar{\sigma}_s$. By Lemma 5.5 this implies $f(\sigma_s)=\Sigma_u \neq \Sigma_t$, a contradiction.

Case 2. $f(\bar{\Sigma}_t)=\bar{\sigma}_i \prec \bar{\sigma}_s$. Again by Lemma 5.5 $f(\sigma_i)=\Sigma_t$. By Lemma 5.4 $\sigma_i \prec \sigma_s$ and hence $f(\sigma_s)\neq\Sigma_t$, again a contradiction. This proves the theorem.

A symmetric theorem (without restriction) for the φ-function does not hold as was previously shown by a counterexample.

We may now state

Theorem 5.3. *If* $\Sigma \in S_n$, $n \geqq \frac{1}{2}(N+1)$, *then*

$$f\varphi(\Sigma)=\Sigma.$$

Proof. Suppose the theorem false. Let t be the least positive integer for which $f\varphi(\Sigma_t)\neq\Sigma_t$. Put $\varphi(\Sigma_t)=\sigma_s$.

Case 1. There exists $\sigma_i \prec \sigma_s$ such that $f(\sigma_i)=\Sigma_t$. By Theorem 5.2 we have $f(\bar{\Sigma}_t)=\bar{\sigma}_i$ so that by Theorem 5.1 $\varphi(\bar{\sigma}_i)=\bar{\Sigma}_t$. On the other hand, by Lemma 5.3, $\varphi(\Sigma_t)=\sigma_s$ implies $\varphi(\bar{\sigma}_s)=\bar{\Sigma}_t$. Since $\sigma_i \prec \sigma_s$, we have $\bar{\sigma}_i \prec \bar{\sigma}_s$ by Lemma 5.4. Put

$$\sigma_s - \sigma_i = x, \qquad \sigma_i - \sigma_s = y;$$

$$\bar{\sigma}_s - \bar{\sigma}_i = \bar{y}, \qquad \bar{\sigma}_i - \bar{\sigma}_s = \bar{x}.$$

We have $y<x$ and $\bar{x}=N+1-x$, $\bar{y}=N+1-y$, so that $\bar{x}<\bar{y}$. Suppose $\bar{y}=\bar{x}+1$. Then $x=y+1$. Also $x, y \in \Sigma_t$. But $\varphi(\Sigma_t)=\sigma_s=(\Sigma_t;\hat{y})$. This is a contradiction, since $x=y+1\in\Sigma_t$. Then $\bar{y}>\bar{x}+1$. If $\bar{y}-1\notin\bar{\Sigma}_t$, then $N+1-(\bar{y}-1)\in\Sigma_t$, so that $y+1\in\Sigma_t$; but $\varphi(\Sigma_t)=(\Sigma_t;\hat{y})$, a contradiction. Then $\bar{y}-1\in\bar{\Sigma}_t=\bar{\sigma}_i\cap\bar{\sigma}_s$.

Let $\bar{x}=\bar{a}_\lambda\in\bar{\sigma}_i$ and $\bar{y}=\bar{b}_\mu\in\bar{\sigma}_s$. Then $\bar{y}-1=\bar{a}_\mu\in\bar{\sigma}_i$. $\mu-\lambda$ indicates the number of elements of $\bar{\Sigma}_t$ between $\bar{x}$ and $\bar{y}$. The number of elements not in $\bar{\Sigma}_t$ which are between $\bar{x}$ and $\bar{y}$ is given by

$$\bar{y}-\bar{x}-1-(\mu-\lambda). \tag{5.21}$$

Since $\varphi(\bar{\sigma}_i)=(\bar{\Sigma}_t)=(\bar{\sigma}_i;\hat{\bar{x}})$, we have

$$\bar{y}-1-\bar{x}-2(\mu-\lambda)\geqq 0. \tag{5.22}$$

From (5.21) it follows that the number of elements of $\sigma_i\cup\sigma_s=\Sigma_t$ between x and y, not including x and y, is

$$\bar{y}-\bar{x}-1-(\mu-\lambda)=x-y-1-(\mu-\lambda).$$

Consider now x and y as elements of Σ_t. Let $x=b_{\lambda_x}$, $y=b_{\lambda_y}$.

$$\lambda_x-\lambda_y=x-y-1-(\mu-\lambda)+1=x-y-(\mu-\lambda).$$

$\varphi(\Sigma_t)=\sigma_s=(\Sigma_t;\hat{y})$ implies

$$x-y-2[x-y-(\mu-\lambda)]\geqq 0, \quad \text{or} \quad y-x+2(\mu-\lambda)\geqq 0.$$

This means

$$\bar{x}-\bar{y}+2(\mu-\lambda)\geqq 0. \tag{5.23}$$

Adding (5.22) and (5.23) brings us to a contradiction.

This proves case 1.

Case 2. $f(\sigma_s)=\Sigma_j\prec\Sigma_t$. Then, since $j<t$, we have $f\varphi(\Sigma_j)=\Sigma_j$; in other words $f\varphi f(\sigma_s)=f(\sigma_s)$. Since f is univalent by definition, this implies $\varphi f(\sigma_s)=\sigma_s$, or $\varphi(\Sigma_j)=\sigma_s$. This is a contradiction since $\varphi(\Sigma_t)=\sigma_s$ and φ is univalent by Corollary 5.1. This completes the proof of the theorem.

In conclusion let us make the following observation.

Let $\sigma=(a_1,,a_n)$. Define $\sigma+t=(b_1,,b_n)$ with $b_i=a_i+t$. The computation is of course performed modulo N. Using our previous results it is quite easy to establish the following

Theorem 5.4. *For any integer t we have*

$$\varphi(\sigma+t)=\varphi(\sigma)+t$$

and also

$$f(\sigma+t)=f(\sigma)+t.$$

(The meaning of the last equality being also that $f(\sigma+t)$ is defined if and only if $f(\sigma)$ is defined.)

6. An Illustrative Example

The following illustrative example might be of help.

Let $N=9$, $n=5$. Choose $\Sigma=(2,4,6,7,9)$. The associated matrix is then

$$A(\Sigma)=\begin{bmatrix} 0 & 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & -1 & -1 \\ -1 & -1 & 0 & -1 & -1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 & 0 \end{bmatrix}.$$

It follows immediately that $\varphi(2,4,6,7,9)=(2,4,6,9)$ and hence $f(2,4,6,9)=(2,4,6,7,9)$.

The complete table for $\varphi(\Sigma)$ and $f(\sigma)$ is given below. Σ is arranged lecicographically.

$\Sigma=f(\sigma)$	$\sigma=\varphi(\Sigma)$	$\Sigma=f(\sigma)$	$\sigma=\varphi(\Sigma)$	$\Sigma=f(\sigma)$	$\sigma=\varphi(\Sigma)$
12345	1234	13478	1378	23579	2579
12346	1236	13479	1379	23589	2589
12347	1237	13489	1389	23678	2367
12348	1238	13567	1356	23679	2679
12349	1239	13568	1358	23689	2689
12356	1235	13569	1359	23789	2789
12357	1257	13578	1357	24567	2456
12358	1258	13579	3579	24568	2458
12359	1259	13589	3589	24569	2459
12367	1267	13678	1367	24578	2457
12368	1268	13679	3679	24569	2479
12369	1269	13689	3689	24589	2489
12378	1278	13789	3789	24678	2467
12379	1279	14567	1456	24679	2469
12389	1289	14568	1458	24689	2468
12456	1245	14569	1459	24789	2478
12457	1247	14578	1457	25678	2567
12458	1248	14579	4579	25679	2569
12459	1249	14589	4589	25689	2568
12467	1246	14678	1467	25789	2578
12468	1468	14679	4679	26789	2678
12469	1469	14689	4689	34567	3456
12478	1478	14789	4789	34568	3458
12479	1479	15678	1567	34569	3459
12489	1489	15679	5679	34578	3457
12567	1256	15689	5689	34579	3479
12568	1568	15789	5789	34589	3489
12569	1569	16789	6789	34678	3467
12578	1578	23456	2345	34679	3469
12579	1579	23457	2347	34689	3468
12589	1589	23458	2348	34789	3478
12678	1678	23459	2349	35678	3567
12679	1679	23467	2346	35679	3569
12689	1689	23468	2368	35689	3568
12789	1789	23469	2369	35789	3578
13456	1345	23478	2378	36789	3678
13457	1347	23479	2379	45678	4567
13458	1348	23489	2389	45679	4569
13459	1349	23567	2356	45689	4568
13467	1346	23568	2358	45789	4578
13468	1368	23569	2359	46789	4678
13469	1369	23578	2357	56789	5678

References

1. Dilworth, R. P.: A decomposition theorem for partially ordered sets. Ann. of Math. **51**, 161–166 (1950).
2. Sperner, E.: Ein Satz über Untermengen einer endlichen Menge. Math. Z. **27**, 544–548 (1928).
3. Katona, G.: A theorem of finite sets. Theory of Graphs, Tihany Colloquium Proc. 1966. New York and London: Academic Press 1968.

Dr. M. Lewin
The Technion
Israel Institute of Technology
Haifa, Israel

(Received October 10, 1970)

Math. Z. 124, 37–46 (1972)

Hebbare Singularitäten quasikonformer Abbildungen und lokal beschränkter holomorpher Funktionen

ALFRED WOHLHAUSER

I. Einleitung

Sei E eine Punktmenge der endlichen Ebene. Wir betrachten alle Überdeckungen von E mit abzählbar vielen offenen Kreisscheiben, deren Durchmesser kleiner sind als eine positive Zahl d. Jeder solchen Überdeckung $\{K_1, K_2, \ldots\}$ ordnen wir die Zahl $\sum d_n^\alpha$ zu, wo $\alpha > 0$ ist und d_n den Durchmesser des Kreises K_n bezeichnet. Für $d \to 0$ nimmt die untere Grenze dieser Zahlen zu; der limes $\ell_\alpha^*(E) := \lim_{d \to 0} (\inf \sum d_n^\alpha)$ heißt das α-dimensionale Hausdorffsche äußere Maß von E. Die Einschränkung ℓ_α von ℓ_α^* auf solche Mengen, die in bezug auf dieses äußere Maß meßbar sind, nennen wir das α-dimensionale Hausdorffsche Maß. Sprechen wir von der Länge $\ell(E)$ einer Punktmenge E, so verstehen wir darunter $\ell_1(E)$.

G bezeichne ein Gebiet und E eine relativ abgeschlossene Teilmenge von G.

Dem zweiten Teil vorliegender Arbeit legen wir die Klasse $\mathscr{W}^K(G-E)$ zugrunde, die aus allen K-quasikonformen Homöomorphismen ($K \geqq 1$) von $G-E$ besteht, welche a priori bis auf höchstens eine Menge der Länge null topologisch in G sind. Im Mittelpunkt der Betrachtungen steht eine hinreichende Bedingung für die Hebbarkeit von E bez. $\mathscr{W}^K(G-E)$. Das Kriterium fordert σ-endliches E, d.h. die Darstellbarkeit von E als abzählbare Vereinigung von Mengen endlicher Länge. Es ist eine Erweiterung des Hebbarkeitskriteriums von Strebel (s. [8], p. 906) und wird wie dieses mit Hilfe von Modulbetrachtungen topologischer Quadrate bewiesen.

Aus dem Satz, den wir im dritten Teil dieser Arbeit beweisen, folgt, daß eine σ-endliche Punktmenge E auch hebbar ist bez. der Klasse aller in $G-E$ K-quasikonformen Abbildungen, welche in G bis auf höchstens eine Menge der Länge null stetig sind. Falls E total unzusammenhängend ist, folgt diese Verallgemeinerung bereits aus der Hebbarkeit einer σ-endlichen Menge E bez. $\mathscr{W}^K(G-E)$.

$\mathscr{W}^1(G-E)$ ist die Menge aller 1–1-konformen Abbildungen von $G-E$, die bis auf höchstens eine Menge der Länge null topologisch in G sind. Ist E hebbar bez. $\mathscr{W}^1(G-E)$, so folgt daraus die Hebbarkeit bez. $\mathscr{W}^K(G-E)$ und umgekehrt. Diese Äquivalenz der Hebbarkeit einer Punktmenge E führt uns darauf, eine σ-endliche Punktmenge auf Hebbarkeit zu prüfen bez. der Klasse $\mathscr{F}(G-E)$, welche aus allen in G lokal beschränkten und in $G-E$ holomorphen Funktionen besteht, die a priori bis auf höchstens eine Menge der Länge null stetig in G sind. Bei den Funktionen aus $\mathscr{F}(G-E)$ lassen wir also im Vergleich zu denjenigen aus $\mathscr{W}^1(G-E)$ die Forderung der Eineindeutigkeit fallen.

Im dritten Teil zeigen wir, daß eine σ-endliche Punktmenge E auch hebbar ist bez. $\mathscr{F}(G-E)$. Zum Beweis dieser Aussage verwenden wir die Cauchysche Integralformel.

Obiges Kriterium für die Hebbarkeit von E bez. $\mathscr{F}(G-E)$ steht in engem Zusammenhang mit dem Painlevéproblem, bei dem die Hebbarkeit von E bez. aller in $G-E$ holomorphen und in G lokal beschränkten Funktionen untersucht wird; über das Verhalten der Funktion auf E setzt man weiter nichts voraus. Ein bekanntes Resultat (s. Riemann surfaces by Ahlfors and Sario, Princeton University Press, 1960, p. 252) besagt, daß E hebbar ist bez. dieser Abbildungsklasse, falls $\ell(E)=0$ ist. Dolženko betrachtet in einer Arbeit (s. [4], p. 135) die Klasse der in $G-E$ holomorphen und in ganz G stetigen Funktionen. In Anlehnung an eine Arbeit von Carleson (s. [3]) zeigt er, daß E hebbar ist bez. dieser Abbildungsklasse, falls $\ell(E)<\infty$. Unser Resultat stellt demnach eine Verallgemeinerung dieser beiden Sätze dar.

Im vierten Teil schließlich geben wir eine sowohl notwendige als auch hinreichende Hebbarkeitsbedingung.

II. Ein Hebbarkeitskriterium für die Klasse $\mathscr{W}^K(G-E)$

G sei ein Gebiet und E eine relativ abgeschlossene Teilmenge von G.

Im Buch von Lehto und Virtanen (s. [5], p. 209) werden folgende Klassen von Abbildungen betrachtet:

$\mathscr{W}_1^K(G-E):=\{f \mid f$ quasikonforme Abbildung von G, $\sup\limits_{z\in G-E} F(z)\leqq K\}$.

$\mathscr{W}_2^K(G-E):=\{f \mid f$ orientierungserhaltender Homöomorphismus von G, $\sup\limits_{z\in G-E} F(z)\leqq K\}$,

$\mathscr{W}_3^K(G-E):=\{f \mid f\ K$-quasikonforme Abbildung von $G-E\}$.

F bezeichnet die lokale Maximaldilatation von f.

E heißt hebbar bez. $\mathscr{W}_1^K$, falls jedes $f\in\mathscr{W}_1^K$ K-quasikonform in G ist. Analog definieren wir die Hebbarkeit bez. $\mathscr{W}_2^K$. Im Fall der Klasse $\mathscr{W}_3^K$ nennen wir E hebbar, falls für jedes $f\in\mathscr{W}_3^K$ eine K-quasikonforme Fortsetzung in E existiert.

Offensichtlich ist $\mathscr{W}_1^K\subset\mathscr{W}_2^K\subset\mathscr{W}_3^K$.

Wir führen nun die Klasse $\mathscr{W}^K(G-E)$ ein. Sie soll aus allen K-quasikonformen Abbildungen von $G-E$ bestehen, die a priori bis auf höchstens eine Menge der Länge null topologisch in G sind.

Sei $G'-E'$ eine Teilmenge von $G-E$, wo E' eine relativ abgeschlossene Teilmenge von G' bezeichnet. Dann ist $\mathscr{W}^K(G-E)\subset\mathscr{W}^K(G'-E')$, m. a. W.: $\mathscr{W}^K(G-E)$ ist eine monotone Klasse von Abbildungen.

Sei ferner g eine $1-1$-konforme Abbildung von G auf G' und $f\in\mathscr{W}^K(G'-E')$, $E'=g(E)$. Da g Punktmengen der Länge null in ebensolche überführt, ist $f\circ g\in\mathscr{W}^K(G-E)$; $\mathscr{W}^K$ ist also in diesem Sinne konform invariant.

Es ist $\mathscr{W}_2^K(G-E)\subset\mathscr{W}^K(G-E)\subset\mathscr{W}_3^K(G-E)$. Die Inklusionen können echt sein, was für $\mathscr{W}^K(G-E)\subset\mathscr{W}_3^K(G-E)$ evident ist. Daß es auch für die andere Inklusion gilt, folgt aus der Arbeit [1] (insbesondere § 7, p. 124): wir gehen aus von einer total unzusammenhängenden abgeschlossenen Punktmenge $E\subset\hat{\mathbb{C}}$ mit vollkommen punktförmigen Komponenten (verallgemeinerte Cantormenge), wo $\hat{\mathbb{C}}$ die abgeschlossene komplexe Ebene sein soll, so daß $\hat{\mathbb{C}}-E$ endlichen Flächeninhalt besitzt. Dann ist z eine $1-1$-konforme Abbildung auf $\hat{\mathbb{C}}-E$ mit endlichem Dirichletintegral, und das impliziert die Existenz einer beschränkten $1-1$-konformen Abbildung $h(z)$ auf $\hat{\mathbb{C}}-E$. $z=\infty$ wird als einzige Komponente von E durch $h(z)$ verzerrt.

Ist $G=R$, wo R die Riemannsche Kugel bezeichnet, so nennen wir $f\in\mathscr{W}^1(R-E)$ meromorph fortsetzbar in E, falls zu f eine in R meromorphe Funktion $\tilde{f}$ existiert, so daß $\tilde{f}|_{R-E}=f$ ist. Wir sehen leicht, daß sich $f\in\mathscr{W}^1(R-E)$ genau dann meromorph fortsetzen läßt, und dies in eine beliebige abgeschlossene Teilmenge E von R ohne innere Punkte, wenn f eine lineare Transformation ist.

E heißt hebbar bez. $\mathscr{W}^K(G-E)$, wenn jedes $f\in\mathscr{W}^K$ K-quasikonform in G fortgesetzt werden kann. Offensichtlich ist E genau dann hebbar bez. $\mathscr{W}^K(G-E)$, falls zu jedem Punkt $p\in E$ eine Umgebung $U(p)$ existiert, so daß wir jedes $f\in\mathscr{W}^K(U(p)-E)$ in $E\cap U(p)$ K-quasikonform fortsetzen können.

Ein bekanntes Resultat von Ahlfors und Beurling sagt aus, daß sich jede $1-1$-konforme Abbildung von $G-E$ genau dann zu einer $1-1$-konformen Abbildung von G fortsetzen läßt, falls E eine 0_{AD}-Nullmenge ist, d.h. eine Menge, deren kompakte Teilmengen in ihrem Komplement keine nichtkonstanten und eindeutigen analytischen Funktionen mit beschränktem Dirichletintegral zulassen.

Ferner: E ist eine 0_{AD}-Nullmenge, sobald sie die Länge null hat.

Trivialerweise ist jede 0_{AD}-Nullmenge hebbar bez. $\mathscr{W}^1(G-E)$.

Es gilt (vgl. [5], p. 210) der

Satz II.1. *Ist E hebbar bez. $\mathscr{W}^1(G-E)$, so ist E hebbar bez. $\mathscr{W}^K(G-E)$ für jedes K, $1\leqq K<\infty$. Umgekehrt ist E hebbar bez. $\mathscr{W}^1(G-E)$, falls es mindestens ein K gibt, so daß E hebbar ist bez. $\mathscr{W}^K(G-E)$.*

Beweis. Sei E hebbar bez. $\mathscr{W}^1$ und $f\in\mathscr{W}^K$. Mit Hilfe der komplexen Dilatation $\kappa(z)$ von f definieren wir f.ü. in der Ebene eine meßbare Funktion:

$$\kappa_0(f(z)):=-\kappa(z)\,e^{2i\arg f_z(z)} \quad \text{für } z\in G-E,$$

$$\kappa_0(\zeta) \quad :=0 \quad \text{für } \zeta\notin f(G-E).$$

(Für die komplexe Dilatation $\kappa_{f^{-1}}$ der zu f inversen Abbildung f^{-1} gilt: $\kappa_{f^{-1}}(f(z))=-\kappa(z)\,e^{2i\arg f_z(z)}$.)

Nach dem Existenzsatz existiert nun eine K-quasikonforme Abbildung f_0 der ganzen Ebene, deren komplexe Dilatation f.ü. mit κ_0 übereinstimmt. $f_1:=f_0\circ f$ ist definiert in $G-E$ und Element von $\mathscr{W}^1$. Bezeichnen wir die Fortsetzung von f_1 mit f_2, so ist $f_0^{-1}\circ f_2$ K-quasikonform in G und $f_0^{-1}\circ f_2|_{G-E}=f$.

Die Umkehrung folgt daraus, daß jede 1-quasikonforme Abbildung auch K-quasikonform ist. q.e.d.

Sei $p\in E$. Falls eine Umgebung $U(p;\Gamma)$ von p existiert, deren Rand eine in $G-E$ gelegene rektifizierbare Jordankurve Γ ist, so betrachten wir $\mathscr{W}^K(U(p;\Gamma)-E)$. Diese Klasse von Abbildungen enthält die Menge der Restriktionen auf $U(p;\Gamma)-E$ der Abbildungen aus $\mathscr{W}^K(G-E)$ als echte Teilmenge.

Wir beweisen

Satz II.2. *Sei $p\in E$ und $U(p;\Gamma)$ eine Umgebung von p, die von einer in $G-E$ gelegenen rektifizierbaren Jordankurve Γ berandet wird.*

Falls E hebbar ist bez. $\mathscr{W}^K(G-E)$, so folgt daraus, daß jedes

$$f\in\mathscr{W}^K(U(p;\Gamma)-E)$$

zu einer K'-quasikonformen Abbildung von $U(p;\Gamma)$ fortgesetzt werden kann. (K' ist $\geqq K$ und von f unabhängig bestimmbar.)

Beweis. Da E relativ G abgeschlossen ist, gibt es ein Gebiet G' und eine rektifizierbare Jordankurve Γ', so daß G' ein in $U(p;\Gamma)-E$ gelegenes Ringgebiet ist mit Γ als einer der beiden Randkomponenten und $\Gamma'\subset G'$. Γ' soll p umschließen.

Sei $f\in\mathscr{W}^K(U(p;\Gamma)-E)$ beliebig. f ist auf G' K-quasikonform, und folglich existiert eine K''-quasikonforme Abbildung h der ganzen Ebene, die auf Γ' mit f übereinstimmt. (K'' ist eine nur von K, G' und Γ' abhängige obere Schranke für die maximale Dilatation von h; s. [5], p. 100.) Also können wir f zu einer K'-quasikonformen Abbildung $\tilde{f}$ von $G-\{\Gamma'\cup[U(p;\Gamma)-E]\}$ fortsetzen. ($K':=\max\{K,K''\}$.) Nach dem Hebbarkeitskriterium von Strebel (s. [8], p. 906) ist $\tilde{f}$ K'-quasikonform in $G-[U(p;\Gamma)\cap E]$ und demzufolge Element von $\mathscr{W}^{K'}(G-E)$. Wir haben Hebbarkeit von E bez. $\mathscr{W}^K(G-E)$ vorausgesetzt. Nach Satz II.1 ist dann E auch hebbar bez. $\mathscr{W}^{K'}(G-E)$. Daraus folgt die Behauptung.

Wir suchen nun nach Kriterien für die Hebbarkeit von E bez. $\mathscr{W}^K(G-E)$.

Eine notwendige Bedingung liefert der

Satz II.3. *Ist E hebbar bez. $\mathscr{W}^K(G-E)$, so besitzt jede kompakte Teilmenge von E das Flächenmaß null.*

Beweis. Sei E^* eine beliebige kompakte Teilmenge von E. Wir definieren die nach oben halbstetige Funktion

$$F(z):=\begin{cases}K+1 & \text{für } z\in E^*\\ K & \text{für } z\in G-E^*,\ K\geqq 1,\ \text{beliebig.}\end{cases}$$

Dann gibt es nach dem Existenzsatz eine quasikonforme Abbildung h von G mit der Eigenschaft, daß ihre lokale Maximaldilatation f.ü. in E^* mit $F(z)$ übereinstimmt und für kein $z\in G-E$ größer ist als K. Da wir E hebbar voraussetzen, ist h K-quasikonform in G. Folglich muß das Flächenmaß von E^* verschwinden. q.e.d.

Im Hinblick auf einen Eindeutigkeitssatz in Zusammenhang mit dem Kreisnormierungsproblem wurde in der Arbeit [6], p. 312 die Aussage bewiesen:

Eine total unzusammenhängende und relativ abgeschlossene Teilmenge E von G ist hebbar bez. $\mathscr{S}(G-E)$, wenn sie auf einer abzählbaren Menge von paarweise disjunkten, abgeschlossenen und rektifizierbaren Jordanbogen liegt. Dabei ist $\mathscr{S}(G-E)$ die Menge aller $1-1$-konformen Abbildungen von $G-E$, welche höchstens eine abzählbare Teilmenge von E in nicht punktförmige Randkomponenten des Bildgebietes überführen, und E heißt hebbar bez. $\mathscr{S}(G-E)$, wenn sich jedes $f\in\mathscr{S}$ holomorph in E fortsetzen läßt.

Diese Bedingung ist insofern eng, als es total unzusammenhängende Punktmengen endlicher Länge gibt, die sich nicht auf abzählbar viele rektifizierbare Jordanbogen legen lassen (s. [2], p. 426).

Nachstehend verallgemeinern wir das Hebbarkeitskriterium von Strebel (s. [8], p. 906). Dabei braucht E nicht total unzusammenhängend zu sein. Die Verallgemeinerung besteht darin, daß wir von unsern Funktionen nicht topologische Fortsetzbarkeit in ganz E fordern, sondern a priori nur bis auf eine Menge der Länge null.

Satz II.4. *Eine relativ abgeschlossene Teilmenge E von G mit σ-endlicher Länge ist hebbar bez. $\mathscr{W}^K(G-E)$.*

Beweis. Sei E eine relativ abgeschlossene Teilmenge von G mit σ-endlicher Länge. Dann hat auch $E\cap Q$, wo Q, $Q\subset G$, ein abgeschlossenes topologisches Quadrat bezeichnet, σ-endliche Länge. Dasselbe gilt für das Bild von $E\cap Q$ unter der kanonischen Abbildung von Q. Deshalb gehen wir aus von einem Rechteck $R(0\leqq x\leqq a,\ 0\leqq y\leqq b)$ in der z-Ebene.

Sei E^* eine Teilmenge von R mit $\ell(E^*)<\infty$. E^*_y bezeichne den Durchschnitt der Geraden $\operatorname{Im} z=y$ mit E^*. Die Länge der Menge aller y, für die E^*_y mindestens n Punkte enthält, ist $\leqq \ell/n$. Folglich verschwindet die Länge der Menge aller y, für welche E^*_y unendlich viele Punkte enthalten. Ist E eine Teilmenge von R σ-endlicher Länge, so hat die Menge aller y, für welche E_y aus überabzählbar vielen Punkten besteht, die Länge null.

Sei w eine K-quasikonforme Abbildung von $R-E$, die a priori höchstens in eine Teilmenge E_0 von E der Länge null nicht topologisch fortgesetzt werden kann. w bilde den Rand ∂R von R eckpunkttreu auf den Rand eines Rechtecks $(0\leqq u\leqq a',\ 0\leqq v\leqq b')$ der w-Ebene ab; auf ∂R sollen keine Punkte von E_0 liegen.

Die Menge der y, deren E_y Punkte von E_0 enthalten, hat die Länge null, da $\ell(E_0)=0$. Wir definieren $\ell_u(y)$ als die Vertikalprojektion auf $[0,a']$ der Häufungsmenge von $w(z)$ für $z\to E_y$.

Für eine σ-endliche Teilmenge E von R gilt also: $\ell_u(y)=0$ für fast alle $y\in[0,b]$, und daraus folgt, daß $\dfrac{a'}{b'}\leqq K\,\dfrac{a}{b}$.

Der Beweis dieser Ungleichung verläuft genau gleich wie der Beweis eines Lemmas von Strebel (s. [8], Lemma 2, p. 905), weshalb wir ihn an dieser Stelle weglassen. Wir weisen lediglich darauf hin, daß auch in unserem Fall die Menge $O_\varepsilon := \{y \mid \ell_u(y)<\varepsilon\}$ eine offene Menge der Länge b ist. Aus der Kompaktheit von E folgt nämlich, daß es zu jedem $y \in O_\varepsilon$ ein Intervall $y_1 < y < y_2$ und endlich viele Rechtecke R_i gibt mit Seiten auf $\operatorname{Im} z = y_1$ und $\operatorname{Im} z = y_2$ (wobei wir mit solchen Horizontalen auskommen, auf denen keine Punkte von E_0 liegen), so daß diese R_i für $y_1 < y < y_2$ jedes E_y enthalten und ihre w-Bilder eine Vertikalprojektion von totalem linearem Maß $< \varepsilon$ haben. O_ε ist somit offen und hat die Länge b.

Sei $z_0 \in G$ ein Punkt, in den wir w a priori nicht topologisch fortsetzen können. Wir betrachten einen konzentrischen Kreisring um z_0, auf dessen Rand w topologisch ist, und schneiden ihn längs einer radialen Strecke, auf der w topologisch sein soll, auf. Das entstehende Viereck und sein w-Bild bilden wir kanonisch ab. Mit Hilfe obiger Ungleichung sehen wir nun leicht, daß z_0 in einen einzigen Punkt übergeführt wird. Daraus schließen wir, daß w topologisch in z_0 fortgesetzt werden kann. Jetzt folgt aus dem Hebbarkeitskriterium von Strebel, daß w eine K-quasikonforme Erweiterung in G besitzt. Damit ist der Satz bewiesen. Er liefert das

Korollar. *Ein Kreisgebiet mit nur punktförmigen Randkomponenten, wobei die Menge der Randkomponenten σ-endliche Länge hat, läßt sich nur durch eine lineare Transformation auf ein anderes Kreisgebiet konform abbilden (vgl.* [7], *p.* 11*).*

Umgekehrt ist die Bedingung obigen Satzes kein notwendiges Kriterium für die Hebbarkeit von E bez. $\mathscr{W}^K(G-E)$. Um das zu sehen, wählen wir für E etwa das Cartesische Produkt zweier eindimensionaler Cantormengen E_x und E_y der Länge null. Dieses E ist nicht σ-endlich, denn $\dim E = \dim E_x + \dim E_y = 2$, und hebbar sogar bez. $\mathscr{W}_3^K(G-E)$.

III. Ein Hebbarkeitskriterium für die Klasse $\mathscr{F}(G-E)$

Im folgenden befassen wir uns mit hebbaren Singularitäten lokal beschränkter holomorpher Funktionen.

Sei E eine relativ abgeschlossene Teilmenge von G. $\mathscr{F}(G-E)$ bezeichne die Menge aller in $G-E$ holomorphen Funktionen, welche in G lokal beschränkt und bis auf eine Menge der Länge null stetig sind.

Es gilt

Satz III.1. *Jedes $f \in \mathscr{F}(G-E)$ kann holomorph in E fortgesetzt werden, falls E σ-endliche Länge hat (vgl.* [4], *p.* 135*).*

Läßt sich umgekehrt jedes $f \in \mathscr{F}(G-E)$ holomorph in E fortsetzen, so muß das $(1+\alpha)$-dimensionale Hausdorffsche Maß von E, $\ell_{1+\alpha}(E)$, wo $\alpha > 0$ beliebig, null sein.

Wir beweisen den hinreichenden Teil der Aussage. Der Beweisgedanke ist kurz folgender:

Zu jedem Punkt $z \in E$ gibt es eine offene Kreisscheibe K, so daß $z \in K$ und $\bar{K} \subset G$. Nach Voraussetzung hat E σ-endliche Länge. Folglich verschwindet das Flächenmaß von E, und wir können $\eta > 0$ so wählen, daß es Punkte $z \in K$ gibt, die nicht zur 3η-Umgebung $E'_{3\eta}$ von $E' := E \cap K$ gehören. Wir haben zu zeigen, daß sich jedes $f \in \mathscr{F}$ holomorph in E' fortsetzen läßt. (Das genügt für den Beweis, da $z \in E$ beliebig.) Dazu betrachten wir das Cauchyintegral

$$\frac{1}{2\pi i} \int_{\partial K} \frac{f(\zeta)}{\zeta - z} \, d\zeta,$$

das a priori nicht mit $f(z)$, $z \notin K$, $z \notin E'_{3\eta}$, übereinzustimmen braucht. Mit Hilfe einer geeigneten Darstellung der Differenz

$$f(z) - \frac{1}{2\pi i} \int_{\partial K} \frac{f(\zeta)}{\zeta - z} \, d\zeta,$$

bei der die σ-endliche Länge von E und die Eigenschaften von f verwendet werden, zeigen wir, daß diese Differenz verschwindet für $z \in K$, $z \notin E'_{3\eta}$. Daraus folgt die Behauptung.

Wir setzen vorerst den Fall, daß $\ell(E) < \infty$, Dabei darf $\ell(E) > 0$ vorausgesetzt werden, da die Aussage im Fall $\ell(E) = 0$ evident ist.

Wir nehmen ein beliebiges $f \in \mathscr{F}$ und führen den Begriff der Grenzschwankung $\omega_f(\zeta)$ von f im Punkte ζ ein:

$$\omega_f(\zeta) := \lim_{r \to 0} \sup \{ |f(\zeta') - f(\zeta'')| \, \zeta', \zeta'' \in K(\zeta, r) \},$$

wo $K(\zeta, r)$ die offene Kreisscheibe mit Mittelpunkt ζ und Radius r bezeichnet.

Sei K eine offene Kreisscheibe, deren abgeschlossene Hülle $\bar{K}$ in G liegt. Wir bestimmen $\eta > 0$ so, daß es eine offene Menge von Punkten $z \in K$ gibt, die nicht zur 3η-Umgebung $E'_{3\eta}$ von $E' := E \cap K$ gehört.

Unter $Q_\alpha(f) := \{\zeta \,|\, \zeta \in \bar{K}, \omega_f(\zeta) \geqq \alpha\}$, $\alpha > 0$, verstehen wir die Menge aller Punkte $\zeta \in \bar{K}$, in denen die Grenzschwankung $\geqq \alpha$ ist. Q_α ist abgeschlossen und Teilmenge von E_0.

Jetzt geben wir $\varepsilon > 0$ beliebig vor. Die Menge $\bar{K} \cap Q_\alpha$, wo $\alpha = \dfrac{\varepsilon \eta}{2 \cdot 2\ell(E)}$, hat die Länge null, da $Q_\alpha \subset E_0$. $\bar{K} \cap Q_\alpha$ läßt sich folglich durch endlich viele offene Kreisscheiben $\{K_n^{(Q)}\}$ mit Radien $r_n^{(Q)} < \eta$ überdecken, so daß $\sum r_n^{(Q)} < \dfrac{\ell(E)\,\alpha}{M}$, wo $|f(z)| \leqq M$ in $\bar{K}$. Diese Überdeckung bezeichnen wir mit $\mathfrak{U}(Q)$.

Durch $\mathfrak{U}(Q)$ werden die Punkte von $O := (E \cap \bar{K}) - \mathfrak{U}(Q)$ nicht überdeckt. O ist eine abgeschlossene Menge, und es ist $\ell(O) \leqq \ell(E)$. Wir wollen O durch endlich viele Kreisscheiben $\{K_n^{(O)}\}$ mit Radien $\{r_n^{(O)}\}$ so überdecken, daß 1) $\sum 2 \cdot r_n^{(O)} < 2\ell(E)$ und 2) es in jeder Kreisscheibe $K_n^{(O)}$ einen Punkt a_n gibt mit der Eigenschaft, daß $|f(\zeta) - f(a_n)| < \alpha$ für $\zeta \in \bar{K}_n^{(O)}$.

Die Forderung 1) läßt sich nach Definition der Länge einer Punktmenge ohne weiteres erfüllen. Um der Bedingung 2) zu genügen, haben wir die Radien $\{r_n^{(O)}\}$ genügend klein zu wählen. Wir bestimmen eine obere Schranke dieser Radien wie folgt: $\bar{K} - \mathfrak{U}(Q)$ ist eine abgeschlossene Teilmenge von $\bar{K}$. In jedem Punkt $\zeta \in \bar{K} - \mathfrak{U}(Q)$ ist $\omega_f(\zeta) < \alpha$. Das erlaubt uns, $\bar{K} - \mathfrak{U}(Q)$ durch endlich viele Kreisscheiben $\{K_n^{(\bar{K} - \mathfrak{U}(Q))}\}$ mit Radien $r_n^{(\bar{K} - \mathfrak{U}(Q))} < r' := \min \{r_n^{(Q)}\}$ so zu überdecken, daß $|f(\zeta') - f(\zeta'')| < \alpha$ für $\zeta', \zeta'' \in \bar{K}_n^{(\bar{K} - \mathfrak{U}(Q))}$, $n = 1, 2, \ldots, N$. Kommen in dieser Überdeckung Kreisscheiben vor, welche durch die restlichen überdeckt werden, so lassen wir sie weg. Die Radien $\{r_n^{(\bar{K}-\mathfrak{U}(Q))}\}$ werden kleiner als r' gewählt, damit die Kreisscheiben $\{K_n^{(Q)}\}$, welche $\bar{K}\cap Q_\alpha$ überdecken, alle zur nachfolgenden Konstruktion der G_{nk} beitragen. Die Kreise $\{K_n^{(\bar{K}-\mathfrak{U}(Q))}\}$ schneiden sich in endlich vielen (voneinander verschiedenen) Punkten $s_1, s_2, \ldots, s_T$. Wir setzen $r'':=\min\{|s_i-s_j|\,|\,i\neq j,\, i,j=1,\ldots,T\}$. Die Zahl $r''/2$ ist, wie man leicht überprüft, eine obere Schranke der oben erwähnten Art für die $\{r_n^{(O)}\}$.

Sei $\{K_n^{(O)}\}$ eine Überdeckung von O mit den Eigenschaften 1) und 2).

Wir betrachten $\{K_n^{(Q)}\}\cup\{K_n^{(O)}\}$ und lassen solche Kreisscheiben weg, welche durch die andern überdeckt werden. Die verbleibenden Elemente ordnen wir nach fallenden Radien und erhalten so eine endliche Folge K_1, $K_2, \ldots, K_P$ von Kreisscheiben, wobei für die zugehörigen Radien gilt: $r_1\geqq r_2\geqq\cdots\geqq r_P$.

Wir definieren (s. [4], p. 136)

$$G_1:=K\cap K_1,\qquad G_n:=K\cap\left(K_n-\bigcup_{i=1}^{n-1}\bar{K}_i\right),$$

wo $2\leqq n\leqq P$. Es ist

$$G_n\subset K_n,\qquad K\cap\left(\bigcup_{n=1}^{P}K_n\right)\subset\bigcup_{n=1}^{P}\bar{G}_n$$

und, weil $r_1\geqq r_2\geqq\cdots\geqq r_P$, $\ell(\partial G_n)\leqq 2\pi\, r_n$. Jedes G_n zerfällt in eine endliche Anzahl von disjunkten Gebieten G_{nk}, $k=1,2,\ldots$, wobei $\sum_k \ell(\partial G_{nk})\leqq\ell(\partial K_n)=2\pi\, r_n$.

Mit Hilfe der so definierten G_{nk} erhalten wir für $f(z)$, $z\in K$, $z\notin E'_{3\eta}$, folgende Darstellung:

$$f(z)=\frac{1}{2\pi i}\int_{\partial K}\frac{f(\zeta)}{\zeta-z}\,d\zeta-\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)}{\zeta-z}\,d\zeta.$$

Es ist $G_{nk}\subset G_n\subset K_n$, und jedem K_n, $1\leqq n\leqq P$, ordnen wir einen Punkt auf folgende Art zu: ist $K_n\in\{K_n^{(Q)}\}$, so darf a_n ein beliebiger aber fester Punkt aus der entsprechenden Kreisscheibe aus $\{K_n^{(Q)}\}$ sein; andernfalls, d.h. wenn $K_n\in\{K_n^{(O)}\}$, wählen wir als a_n einen Punkt in der entsprechenden Kreisscheibe aus $\{K_n^{(O)}\}$, für den gilt: $|f(\zeta)-f(a_n)|<\alpha$ für $\zeta\in\bar{K}_n^{(O)}$. (Das ist obige Forderung 2) an die Überdeckung $\{K_n^{(O)}\}$ von O.) Es ist

$$\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)}{\zeta-z}\,d\zeta=\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)-f(a_n)}{\zeta-z}\,d\zeta,$$

und wir haben die ganze Konstruktion daraufhin ausgerichtet, daß der letzte Ausdruck die Abschätzung erlaubt:

$$\begin{aligned}\left|\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)-f(a_n)}{\zeta-z}\,d\zeta\right|&<\frac{1}{2\pi\eta}\cdot 2M\cdot\sum 2\pi\, r_n^{(Q)}\\&\quad+\frac{1}{2\pi\eta}\cdot\alpha\cdot\sum 2\pi\, r_n^{(O)}<\frac{1}{2\pi\eta}\cdot 2M\cdot\frac{2\pi\,\ell(E)\,\alpha}{M}\\&\quad+\frac{1}{2\pi\eta}\cdot\frac{\varepsilon\eta}{2\cdot 2\ell(E)}\cdot 2\pi\,\ell(E)=\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

Daraus folgt die Behauptung für den Fall, daß $\ell(E)<\infty$.

Sei E σ-endlich, d.h. $E=\bigcup_{i=1}^{\infty} E_i$, wo $\ell(E_i)<\infty$ für $i=1,2,\ldots$; o.B.d.A. darf $\ell(E_i)>0$ vorausgesetzt werden. Mit Q_i, $i=1,2,\ldots$, bezeichnen wir die Menge derjenigen Punkte von $\bar{K}$, in denen die Grenzschwankung $\geqq \alpha_i$ ist, mit $\alpha_i=\frac{\varepsilon\eta}{2\cdot 2\ell(E_i)\,2^i}$. Es ist $\ell(Q_i)=0$ für jedes i. Wir überdecken Q_i durch endlich viele offene Kreisscheiben $\{K_n^{(Q_i)}\}$ mit Radien $r_n^{(Q_i)}<\eta$, so daß

$$\sum_n r_n^{(Q_i)}<\frac{\alpha_i\,\ell(E_i)}{M}.$$

Diese Überdeckung bezeichnen wir mit $\mathfrak{U}(Q_i)$. Mit den Punkten von $O_i:=(E_i\cap\bar{K})-\mathfrak{U}(Q_i)$, das sind diejenigen Punkte von E_i in $\bar{K}$, welche durch $\mathfrak{U}(Q_i)$ nicht überdeckt werden, verfahren wir gleich wie im endlichen Fall $(\ell(E)<\infty)$ mit O, d.h. wir überdecken O_i durch (evtl. abzählbar viele) offene Kreisscheiben $\{K_n^{(O_i)}\}$ mit Radien $\{r_n^{(O_i)}\}$ mit den Eigenschaften: 1) $r_n^{(O_i)}<\min\{r_n^{(Q_i)}\}$, 2) $\sum_n 2r_n^{(O_i)})<2\ell(E_i)$ und 3) in jeder Kreisscheibe $K_n^{(O_i)}$ gibt es einen Punkt $a_n^{(i)}$, so daß $|f(\zeta)-f(a_n^{(i)})|<\alpha_i$ für $\zeta\in\bar{K}_n^{(O_i)}$. Mit $\{K_n^{(O_i)}\}\cup\mathfrak{U}(Q_i)$ haben wir für jedes i eine Überdeckung von $E_i\cap\bar{K}$, die wir mit $\mathfrak{U}(E_i\cap\bar{K})$ bezeichnen wollen. Es ist $(E\cap\bar{K})\subset\bigcup_{i=1}^{\infty}\mathfrak{U}(E_i\cap\bar{K})$, und aus $\bigcup_{i=1}^{\infty}\mathfrak{U}(E_i\cap\bar{K})$ wählen wir endlich viele Kreisscheiben $\{K_n\}$ aus, welche $E\cap\bar{K}$ überdecken. Mit Hilfe dieser $\{K_n\}$ bestimmen wir die $\{G_{nk}\}$ und ordnen jedem K_n auf dieselbe Art wie im Fall $\ell(E)<\infty$ einen Punkt a_n zu. Dann gilt für $z\in K$, $z\notin E'_{3\eta}$:

$$f(z)=\frac{1}{2\pi i}\int_{\partial K}\frac{f(\zeta)}{\zeta-z}\,d\zeta-\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)-f(a_n)}{\zeta-z}\,d\zeta,$$

wobei

$$\left|\frac{1}{2\pi i}\sum\int_{\partial G_{nk}}\frac{f(\zeta)-f(a_n)}{\zeta-z}\,d\zeta\right|<\frac{1}{2\pi\eta}\cdot 2M\cdot\sum_i\sum_n 2\pi\, r_n^{(Q_i)}$$
$$+\frac{1}{2\pi\eta}\sum_i\alpha_i\,2\pi\,\ell(E_i)<\frac{1}{2\pi\eta}\sum_i 2M\,\frac{2\pi\,\varepsilon\,\mu\,\ell(E_i)}{2\cdot 2\ell(E_i)\,2^i M}$$
$$+\frac{1}{2\pi\eta}\sum_i\frac{\varepsilon\eta}{2\cdot 2\ell(E_i)\,2^i}\cdot 2\pi\,\ell(E_i)<\varepsilon.$$

Damit ist der hinreichende Teil des Satzes bewiesen.

Der notwendige Teil der Aussage folgt aus [4], p. 138. Dort wird gezeigt: ist $\ell_{1+\alpha}(E)>0$, wo $0<\alpha<1$, so gibt es eine Funktion f, die in G einer Lipschitz-bedingung der Ordnung α genügt und in $G-E$ holomorph ist, sich aber nicht holomorph in E fortsetzen läßt. Folglich muß $\ell_{1+\alpha}(E)$ verschwinden für $0<\alpha<1$. Daraus folgt, daß $\ell_{1+\alpha}(E)=0$ für beliebiges $\alpha>0$. Wäre nämlich $\ell_{1+\alpha}(E)\neq 0$ für mindestens eine Zahl $\alpha\geqq 1$, so müßte $\ell_{1+\alpha}(E)=\infty$ sein für $0<\alpha<1$ (s. [5], p. 121).

IV. Ein notwendiges und hinreichendes Hebbarkeitskriterium

Sei E eine relativ abgeschlossene Teilmenge von G mit Flächenmaß null und f eine auf $G-E$ holomorphe Funktion, die in G lokal beschränkt ist.

Sei E^* eine beliebige kompakte Teilmenge von E und $\{K_n(\zeta_n, r_n)\}$ eine Überdeckung von E^*, wo $K_n(\zeta_n, r_n)$ die offene Kreisscheibe mit Zentrum ζ_n und Radius r_n bezeichnet. Mit

$$\omega_f(\bar{K}_n) := \sup\{|f(z')-f(z'')|\,|\,z', z'' \in \bar{K}_n(\zeta_n, r_n)\}$$

bezeichnen wir die Schwankung von f im Kreis $\bar{K}_n(\zeta_n, r_n)$.

Wir setzen $\ell_f^{(S)}(E^*) := \lim_{r\to 0} \inf_{r_n \leqq r} \sum r_n\, \omega_f(\bar{K}_n)$.

Ist $\ell_f^{(S)}(E^*)=0$ für jede kompakte Menge $E^* \subset E$, so kann f höchstens auf einer Menge der Länge null (Teilmenge von E) unstetig sein. Wir zeigen

Satz IV.1. *Eine in G lokal beschränkte und in $G-E$ holomorphe Funktion f läßt sich genau dann holomorph in E fortsetzen, falls $\ell_f^{(S)}(E^*)=0$ für jede kompakte Menge $E^* \subset E$.*

Beweis. Sei f holomorph fortsetzbar. Dann gilt für eine beliebige kompakte Teilmenge E^* von E:

$$\ell_f^{(S)}(E^*) = \lim_{r\to 0} \inf_{r_n \leqq r} \sum r_n\, \omega_f(\bar{K}_n) \leqq \lim_{r\to 0} \inf_{r_n \leqq r} C(E^*) \sum r_n^2 = 0,$$

wo $C(E^*)$ eine von E^* abhängige Konstante bezeichnet.

Sei umgekehrt f eine in G lokal beschränkte und in $G-E$ holomorphe Funktion, so daß $\ell_f^{(S)}(E^*)=0$ für jedes kompakte $E^* \subset E$. Wir nehmen einen beliebigen Punkt $z \in E$ und bestimmen eine Kreisumgebung K von z, deren abgeschlossene Hülle $\bar{K}$ in G liegt. Mit Hilfe der Beweismethode von Satz III.1. sehen wir, daß sich f in K durch das Cauchyintegral über ∂K darstellen läßt; somit kann f holomorph in E fortgesetzt werden. q.e.d.

Literatur

1. Ahlfors, L., Beurling, A.: Conformal invariants and function—theoretic null—sets. Acta math. **83**, 101—129 (1950).
2. Besicovitch, A.S.: On the fundamental geometrical properties of linearly measurable plane sets of points. Math. Ann. **98**, 422—464 (1926).
3. Carleson, L.: On null—sets for continuous analytic functions. Ark. Mat. 1 Nr. **22**, 311—318 (1949—1951).
4. Dolženko, E.: Hebbarkeit von Singularitäten analytischer Funktionen [in russischer Sprache]. Uspehi Mat. Nauk 18, Vol. **4**, 135—142 (1963).
5. Lehto, O., Virtanen, K.I.: Quasikonforme Abbildungen. Grundlehren der math. Wiss., Band 126. Berlin-Heidelberg-New York: Springer 1965.
6. Meier-Solfrian, W.: Zur Eindeutigkeit konformer Abbildungen von Gebieten unendlichen Zusammenhangs. Diss., Commentarii Math. Helvet. **43**, 311—330 (1968).
7. Strebel, K.: Über das Kreisnormierungsproblem der konformen Abbildung. Diss., Ann. Acad. Sci. Fennicae **101**, 5—21 (1951).
8. — On the maximal dilation of quasiconformal mappings. Proc. Amer. Math. Soc. **6**, 903—909 (1955).

Dr. Alfred Wohlhauser
Mathematisches Institut
der Universität Zürich
CH-8032 Zürich, Freiestr. 36
Schweiz

(Eingegangen am 4. Januar 1971)

Math. Z. 124, 47–50 (1972)

Essential Selfadjointness of a Schrödinger Operator with Strongly Singular Potential

UPKE-WALTHER SCHMINCKE

Writing $\mathbb{R}^n_+ := \mathbb{R}^n \setminus \{o\}$ $(n \geqq 2)$ we consider in the Hilbert space $H = L^2(\mathbb{R}^n)$ the Schrödinger operator T defined by

$$Tu = -\Delta u + qu, \qquad D(T) = C_0^\infty(\mathbb{R}^n_+),$$

where $q \in L^2_{\mathrm{loc}}(\mathbb{R}^n_+)$ is a real-valued function.

Dealing with general singular elliptic operators Kalf and Walter derive in [6] selfadjointness criteria which for the special case of T reduce to the following

Theorem (Kalf-Walter [6]). *Assume that q can be expressed as $q = q_1 + q_2$ with $q_1 \in L^\infty(\mathbb{R}^n)$, $q_2 \in Q^\alpha_{\mathrm{loc}}(\mathbb{R}^n_+)$ (here $Q^\alpha_{\mathrm{loc}}(\mathbb{R}^n_+)$ denotes the class of functions satisfying a local Stummel condition in $\mathbb{R}^n_+$; see e.g. Jörgens [4]) and*

$$\frac{\beta}{|x|^2} \leqq q_2(x), \qquad x \in \mathbb{R}^n_+,$$

β being a constant. If $\beta > \beta_0 := 1 - \left(\frac{n-2}{2}\right)^2$ then T is essentially selfadjoint. The constant β_0 is the best possible.

This assertion is proved by means of a criterion due to Walter [9] involving a certain lower estimate for the operator and an inequality of Hardy's type which seems to be the most appropriate in this connection. In this way it is possible to cover cases which could previously be treated by separation only.

Now the question arises whether $\beta = \beta_0$ can be admitted in the statement above or not.

In this paper we shall give a partial answer to this question. Using a perturbation argument and an inequality due to Rellich which is closely related to the inequalities of Hardy's type we show that under certain restrictions for the growth of the potential as $|x| \to 0$ and $|x| \to \infty$ the answer is affirmative.

Theorem. *Assume q to have the form $q = q_1 + q_2 + q_3$, where the q_j are measurable real-valued functions satisfying the conditions*

$$\left[1 - \left(\frac{n-2}{2}\right)^2\right] \frac{1}{|x|^2} \leqq q_1(x) \leqq \frac{\gamma}{|x|^2}, \qquad x \in \mathbb{R}^n_+, \tag{1}$$

$$0 \leqq q_2(x) \leqq f(x), \qquad x \in \mathbb{R}^n_+, \tag{2}$$

$$q_3 \in L^\infty(\mathbb{R}^n), \tag{3}$$

γ being a constant with $\gamma > 1 + (n-2)^2/4$ and $f \in C^0(\mathbb{R}^n) \cap C^2(\mathbb{R}^n_+)$ a positive-valued function with

$$\alpha_1 f + \alpha_2 \geqq \Delta f, \quad x \in \mathbb{R}^n_+ \tag{4}$$

(α_1, α_2 suitable positive constants). Then T is essentially selfadjoint.

Remark. One can choose $f(x) = |x|^2 e^{c|x|}$, $c > 0$, for instance.

For the proof of this theorem we need two lemmata.

Lemma 1. *Let H be a Hilbert space, A and B operators in H. Let A be essentially selfadjoint and let B be symmetric with $D(B) \supseteq D(A)$ and*

$$\|Bu\|^2 \leqq c\|u\|^2 + \|Au\|^2, \quad u \in D(A), \tag{5}$$

where c is a suitable constant. Then $A+B$ is essentially selfadjoint.

Proof. See Kato [7], p. 289. For a stronger statement see Wüst [10] who shows that the squares in (5) can be omitted.

Lemma 2. *Suppose $u \in C_0^\infty(\mathbb{R}^n_+)$ and $s \in \left[-\frac{n(n-4)}{2}, \infty\right)$. Then*

$$\int_{\mathbb{R}^n} |\Delta u|^2\, dx \geqq -s \int_{\mathbb{R}^n} \frac{|\nabla u|^2}{|x|^2}\, dx + \frac{(n-4)^2}{16}(n^2+4s) \int_{\mathbb{R}^n} \frac{|u|^2}{|x|^4}\, dx. \tag{6}$$

Proof. See Rellich [8] and for generalizations Heinz [3], Cordes [2], and Aronszajn [1]. We may sketch a simple proof. Excluding the trivial case we assume $n \neq 4$. Let L, G and C denote the integrals in (6) successively. Applying the identity $1/|x|^4 = (1/(8-2n))\,\Delta(1/|x|^2)$, Green's formula and Schwarz's inequality we find (integration extends over the whole $\mathbb{R}^n$ here as well as in the following):

$$C = \frac{1}{2(4-n)} \int |u|^2 \Delta\left(\frac{1}{|x|^2}\right) dx = \frac{1}{2(4-n)} \int \frac{1}{|x|^2} \Delta(|u|^2)\, dx$$

$$\leqq \frac{1}{|4-n|}\sqrt{CL} + \frac{1}{4-n} G.$$

Hence for all $\varepsilon > 0$, $\delta \geqq 0$

$$|n-4|\, C \leqq \frac{\varepsilon}{2} C + \frac{1}{2\varepsilon} L + \operatorname{sgn}(4-n) \cdot G + \delta G - \frac{(n-4)^2}{4} \delta C$$

if we use the inequality $4G \geqq (n-4)^2 C$ which is of Hardy's type (see [6]). Choosing $\varepsilon = (n|n-4|)/4$, $\delta = (2s + n(n-4))/(n|n-4|)$ we obtain the desired result.

Proof of the Theorem. We may assume $q_3 \equiv 0$ for we can always add a bounded symmetric operator to T without perturbing the essential selfadjointness of T. Putting $a := (\gamma + \beta_0)/2$, $b := (\gamma - \beta_0)/2$ we split T into two parts:

$$A := -\Delta + q_{\mathrm{I}}, \quad B := q_{\mathrm{II}}, \quad D(A) = D(B) = D(T),$$

with

$$q_{\mathrm{I}} := \frac{a}{|x|^2} + \frac{1}{2} f, \qquad q_{\mathrm{II}} := q_1 + q_2 - q_{\mathrm{I}} \qquad (x \in \mathbb{R}^n_+).$$

As an easy consequence of (1) and (2) we have

$$q_{\mathrm{II}}^2 \leqq \frac{b^2}{|x|^4} + \frac{b}{|x|^2} f + \frac{1}{4} f^2, \qquad x \in \mathbb{R}^n_+. \tag{7}$$

In order to apply Lemma 1 we show that for sufficiently large α there is a constant c such that

$$\|(A+\alpha E)u\|^2 - \|Bu\|^2 \geqq -c\|u\|^2 \tag{8}$$

holds for all $u \in D(A)$. Let α be an arbitrary positive number and suppose $u \in D(A)$. Green's formula together with the relation $\bar{u}\Delta u + u\Delta\bar{u} = \Delta(|u|^2) - 2|\nabla u|^2$ yields (note that the boundary terms vanish):

$$\begin{aligned}\|(A+\alpha E)u\|^2 - \|Bu\|^2 = &\int |\Delta u|^2\,dx + 2\int q_{\mathrm{I}} |\nabla u|^2\,dx + 2\alpha \int |\nabla u|^2\,dx \\ &+ \int \{q_{\mathrm{I}}^2 + 2\alpha q_{\mathrm{I}} - \Delta q_{\mathrm{I}} - q_{\mathrm{II}}^2\} |u|^2\,dx + \alpha^2 \int |u|^2\,dx.\end{aligned}$$

Since we have ($r := |x|$)

$$\Delta q_{\mathrm{I}} = -\frac{2a(n-4)}{r^4} + \frac{1}{2}\Delta f, \qquad x \in \mathbb{R}^n_+,$$

$$\int |\nabla u|^2\,dx \geqq \frac{(n-2)^2}{4} \int \frac{|u|^2}{r^2}\,dx$$

(this is Hardy's inequality) and

$$\int \frac{f}{r^2} |u|^2\,dx \leqq \max_{|y| \leqq 1} f(y) \int \frac{|u|^2}{r^2}\,dx + \int f |u|^2\,dx \tag{9}$$

we get using the same notations as in the proof of Lemma 2 and regarding (7):

$$\begin{aligned}&\|(A+\alpha E)u\|^2 - \|Bu\|^2 \\ &\quad \geqq L + 2aG + \int f|\nabla u|^2\,dx + \alpha\frac{(n-2)^2}{2}\int\frac{|u|^2}{r^2}\,dx + \{a^2 + 2a(n-4) - b^2\}\,C \\ &\quad + \int\left\{\frac{a}{r^2} f + 2\alpha\frac{a}{r^2} + \alpha f - \frac{1}{2}\Delta f - \frac{b}{r^2} f\right\}|u|^2\,dx + \alpha^2\|u\|^2.\end{aligned}$$

If we take account of (4) and (9) we see that α can be chosen so large that the inequality

$$\|(A+\alpha E)u\|^2 - \|Bu\|^2 \geqq L + 2aG + \{a^2 + 2a(n-4) - b^2\}\,C - c\|u\|^2$$

holds for all $u \in D(A)$ with a suitable constant c. Application of Lemma 2 with $s = 2a$ now leads us to (8).

Noting finally that A as well as $A+\alpha E$ are essentially selfadjoint (see e.g. Jörgens [4]) and applying Lemma 1 we obtain the assertion of our theorem.

References

1. Aronszajn, N.: A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order. J. Math. Pur. Appl. (9) **36**, 235–249 (1957).
2. Cordes, H.O.: Über die eindeutige Bestimmtheit der Lösungen elliptischer Differentialgleichungen durch Anfangsvorgaben. Nachr. Akad. Wiss. Göttingen II a, 239–258 (1956).
3. Heinz, E.: Über die Eindeutigkeit beim Cauchyschen Anfangswertproblem einer elliptischen Differentialgleichung zweiter Ordnung. Nachr. Akad. Wiss. Göttingen II a, 1–12 (1955).
4. Jörgens, K.: Wesentliche Selbstadjungiertheit singulärer elliptischer Differentialoperatoren zweiter Ordnung in $C_0^\infty(G)$. Math. Scandinav. **15**, 5–17 (1964).
5. Kalf, H.: Über die Selbstadjungiertheit halbbeschränkter gewöhnlicher oder elliptischer Differentialoperatoren mit stark singulärem Potential. Dissertation RWTH Aachen 1971.
6. — Walter, J.: Strongly singular potentials and essential selfadjointness of singular elliptic operators in $C_0^\infty(\mathbb{R}^n \setminus \{o\})$. J. Functional Analysis (to appear).
7. Kato, T.: Perturbation theory for linear operators. Berlin-Heidelberg-New York: Springer 1966.
8. Rellich, F.: Perturbation theory of eigenvalue problems. New York-London-Paris: Gordon and Breach 1969.
9. Walter, J.: Note on a paper by Stetkær-Hansen concerning essential selfadjointness of Schrödinger operators. Math. Scandinav. **25**, 94–96 (1969).
10. Wüst, R.: Generalisations of Rellich's theorem on perturbation of (essentially) selfadjoint operators. Math. Z. **119**, 276–280 (1971).

Dr. Upke-Walther Schmincke
Institut für Mathematik
Technische Hochschule Aachen
D-5100 Aachen, Templergraben 55
Deutschland

(Received May 12, 1971)

Math. Z. 124, 51-63 (1972)

Groups with Automorphisms Inverting most Elements

HANS LIEBECK* and DESMOND MACHALE**

1. Introduction

It is well known that a group is Abelian if and only if it has an automorphism inverting *all* its elements. In this paper we study finite non-Abelian groups in which *most* (i.e. more than half) of the elements are inverted by some automorphism. It might be expected that such groups are close to being Abelian, and this is indeed the case. We prove that either they possess an Abelian subgroup of index 2 (with no further restriction) or they are nilpotent of class 2 of a special type, as detailed in Theorem 4.13. This structure does not extend to groups with an automorphism inverting exactly half its elements. For example, the alternating group on four symbols has such an automorphism.

Our problem has attracted the attention of a number of authors in the early part of this century. Miller first showed that there are no non-Abelian groups with an automorphism inverting more than $\frac{3}{4}$ of its elements. Manning [2] proved that a group G has an automorphism inverting exactly $\frac{3}{4}$ of its elements (which we call a $\frac{3}{4}$-automorphism) if and only if the centre has index 4 in G. He also showed that if a group has a k-automorphism with $k \geqq \frac{5}{8}$ then $k = \frac{5}{8}, \frac{2}{3}, \frac{3}{4}$, or 1. Miller [5, 6] obtained similar results and considered more generally the properties of groups with a k-automorphism, $k > \frac{1}{2}$.

Manning's restriction on k fits into a pattern which is part of our Structure Theorem (4.13): we show that groups with a k-automorphism, $\frac{1}{2} < k < 1$, exist if and only if k has the form $\dfrac{q+1}{2q}$ for some integer q. It turns out, rather surprisingly, that a group cannot have both a k_1-automorphism and a k_2-automorphism for different values of k_1 and k_2 exceeding $\frac{1}{2}$ (Corollary (3.10)).

As a special case of Theorem (4.13) we obtain the structure of groups in which the identity automorphism inverts more than half the elements. These are precisely the groups in which at least half the elements are involutions. There exist groups of arbitrary even order with this property. The problem of constructing such groups also has a long history. Miller considered it in a series of papers covering a period of 15 years. The first and last of these are [3] and [4]. He finally arrived at a classification which is not too easy to follow. Recently Wall [7] gave a neat classification. Professor Wall used character theory in his analysis. We present an alternative treatment by "elementary" methods.

---

* This paper was completed while the author enjoyed the hospitality of the Mathematical Institute, University of Oxford, which he gratefully acknowledges.

** The author is holder of a studentship at the University of Keele.

2. Notation

G denotes a finite non-Abelian group, and H and A are subgroups of G which are assigned special properties as required. A finite set T (of group elements) has $|T|$ distinct members. If α is an automorphism of G we denote by S_α the set of elements in G which are inverted under α. We define

$$l(\alpha)=\frac{|S_\alpha|}{|G|}$$

and

$$l(G)=\max_{\gamma\in\operatorname{Aut}G} l(\gamma).$$

Our aim is to classify groups G with $l(G)>\frac{1}{2}$.

If $l(\alpha)=k$, we call α a k-automorphism. If $k>\frac{1}{2}$ we also say that α is a $>\frac{1}{2}$-automorphism. A $>\frac{1}{2}$-group is a group with a $>\frac{1}{2}$-automorphism.

Given $x_i\in G$ $(i=1,\dots,r)$ and subset $T\subseteq G$, $\langle x_1,\dots,x_r,T\rangle$ denotes the group generated by $x_1,\dots,x_r$ and the elements of T. The index of a subgroup H in G is $(G:H)$.

3. Preliminary Theorems

(3.1) **Lemma.** *Given $\alpha\in\operatorname{Aut}G$ and $s\in S_\alpha$, let I_s be the inner automorphism defined by $g I_s=s^{-1}g s$, $g\in G$. Then the set S_β of elements of G inverted by the automorphism $\beta=I_s\alpha$ is given by*

$$S_\beta=S_\alpha s^{-1}=s S_\alpha.$$

Thus $l(\beta)=l(\alpha)$.

Proof. $g(I_s\alpha)=g^{-1}$ iff $(s^{-1}g s)\alpha=g^{-1}$ iff $(g s)\alpha=(g s)^{-1}$ iff $g s\in S_\alpha$. Therefore $S_\beta=S_\alpha s^{-1}$. The second equality follows similarly.

(3.2) **Subgroup Theorem.** *Let H be a subgroup of a $>\frac{1}{2}$-group G. Then there is a $>\frac{1}{2}$-automorphism of G that inverts more than half the elements of H (and so maps H onto itself). Moreover, $l(H)\geqq l(G)$.*

Proof. Let α be a $>\frac{1}{2}$-automorphism of G. Since $|S_\alpha|=l(\alpha)|G|$, it follows that some coset of H in G, Hs say, has at least $l(\alpha)|Hs|$ elements in S_α. We may clearly choose the coset representative s in S_α. From the inequality

$$|Hs\cap S_\alpha|\geqq l(\alpha)|Hs|$$

it follows that

$$|H\cap S_\alpha s^{-1}|\geqq l(\alpha)|H|.$$

Put $\beta=I_s\alpha$. Then, by Lemma (3.1), $l(\beta)=l(\alpha)$ and

$$|H\cap S_\beta|\geqq l(\beta)|H|.$$

Thus β inverts more than half the elements of H. Denote by $\beta|H$ the automorphism of H induced by β. Then

$$l(\beta|H)=\frac{|H\cap S_\beta|}{|H|}\geqq l(\beta).$$

It follows that if we choose α such that $l(\alpha)=l(G)$, then $l(\beta|H)\geqq l(\beta)=l(G)$ and so $l(H)\geqq l(G)$.

The following corollary is an immediate consequence of the subgroup theorem.

(3.3) **Corollary.** *If H is an Abelian subgroup of G (with $l(G)>\frac{1}{2}$) then there is a $>\frac{1}{2}$-automorphism of G that inverts H elementwise.*

Abelian subgroups play a fundamental role in the structure of $>\frac{1}{2}$-groups, and we proceed to study them in some detail.

(3.4) **Lemma.** *Let β be a $>\frac{1}{2}$-automorphism of G that inverts an Abelian subgroup H elementwise. Suppose that the coset Hg has non-trivial intersection with S_β. Then the number of elements in Hg that are inverted by β is $|C_H(g)|$, the order of the centralizer of g in H.*

Proof. By hypothesis $Hg\cap S_\beta$ is not empty, so we may choose $s\in S_\beta$ such that $Hg=Hs$. Now for $h\in H$, $(hs)\alpha=(hs)^{-1}$ if and only if $h^{-1}s^{-1}=s^{-1}h^{-1}$. Hence

$$Hg\cap S_\beta=(C_H(s))\,s.$$

But $C_H(s)=C_H(g)$, and the lemma follows.

(3.5) **Transversal Theorem.** *Let β be a $>\frac{1}{2}$-automorphism of G that inverts the Abelian subgroup H elementwise. If H is a maximal subgroup of S_β (that is, H is not contained properly in a subgroup of G lying entirely in S_β) then there exists a decomposition of G relative to cosets of H*

$$G=Hs_1\cup Hs_2\cup\cdots\cup Hs_n,\qquad s_1=1,$$

such that $s_i\in S_\alpha$ $(i=1,\ldots,n)$.

Proof. Suppose that Hg contains an element $s\in S_\beta$; then $Hg=Hs$, and s may be chosen as coset representative. Now if $Hs\neq H$ then $C_H(s)$ must be a proper subgroup of H, for otherwise $\langle s, H\rangle$ lies in S_β and properly contains H, contrary to hypothesis. By Lemma (3.4), at most half the elements of Hg belong to S_β. Hence if some coset of H in G contained no elements of S_β then $|S_\beta|$ could not exceed $\frac{1}{2}|G|$. Since β is a $>\frac{1}{2}$-automorphism every coset must contain an element of S_β and the theorem is proved.

(3.6) **Centralizer Theorem.** *Let H be a maximal subgroup of S_β, where β is a $>\frac{1}{2}$-automorphism of G. Let*

(3.7) $$G=H\cup Hg_2\cup\cdots\cup Hg_n$$

be a decomposition of G into a union of disjoint cosets of H. Put

$$q_i=(H:C_H(g_i)).$$

Then

(3.8) $$|S_\beta|=|H|+\sum_{i=2}^{n}|C_H(g_i)|,\qquad \text{where } q_i\geqq 2\ (i=2,\ldots,n).$$

Moreover, the following inequality on the indices q_i holds:

(3.9)
$$\sum_{i=2}^{n} \frac{1}{2} - \frac{1}{q_i} < \frac{1}{2}.$$

Proof. Formula (3.8) follows from Lemma (3.4) and Theorem (3.5). The inequality (3.9) is a consequence of the fact that $|S_\beta| > \frac{1}{2}n|H|$.

(3.10) **Corollary.** *An automorphism of a $>\frac{1}{2}$-group G either inverts $l(G)|G|$ elements or it inverts not more than half the elements.*

Proof. Let α be a $>\frac{1}{2}$-automorphism of G, and let H be an arbitrary maximal Abelian subgroup of G. By the proof of the Subgroup Theorem (3.2) and Corollary (3.3), there exists an automorphism β that inverts H elementwise and such that $l(\beta) = l(\alpha)$. Since H is maximal Abelian in G it is certainly a maximal subgroup of S_β. Therefore, by (3.8),

(3.11)
$$|S_\alpha| = |S_\beta| = |H| + \sum_{i=2}^{n} |C_H(g_i)|,$$

where we suppose that G admits a coset decomposition (3.7). But the number given by the right hand side of (3.11) is independent of α; so any $>\frac{1}{2}$-automorphism must invert this number of elements.

From the above proof we obtain a formula which relates this number to the Abelian subgroup structure of G: *if H is an arbitrary maximal Abelian subgroup of the $>\frac{1}{2}$-group G, then*

$$l(G)|G| = |H| + \sum_{i=2}^{n} |C_H(g_i)|,$$

where G admits a coset decomposition (3.7).

As a consequence of the centralizer inequality (3.9) certain conditions are imposed on the maximal Abelian subgroups of a $>\frac{1}{2}$-group G. Suppose a maximal Abelian subgroup H has index n in G. Then relative to a suitable ordering of the cosets of H in G only the following cases can possibly arise:

(3.12)
- I $n=2$;
- II $n \geqq 3$, $q_i = 2$ $(i=2, \ldots, n)$;
- III $n \geqq 3$, $q_2 \geqq 3$, $q_i = 2$ $(i=3, \ldots, n)$;
- IV $n \geqq 3$, $q_2 = 3$, $q_3 = 4$ or 5, $q_i = 2$ $(i=4, \ldots, n)$;
- V $n \geqq 3$, $q_2 = q_3 = 3$, $q_i = 2$ $(i=4, \ldots, n)$.

It turns out that some of these cases do not occur in $>\frac{1}{2}$-groups. In order to show this we require some further results concerning the Abelian subgroup structure of such groups.

(3.13) **The Squares Theorem.** *Let H be a subgroup of maximum order in S_β, where β is a $>\frac{1}{2}$-automorphism of G. Then the square of every element in S_β belongs to H.*

Proof. Suppose $s \in S_\beta$ but $s \notin H$. We must show that $s^2 \in H$. If $(G:H)=2$ the result is clear. So we may suppose that $(G:H)=n \geqq 3$. We first consider the case $(H:C_H(s))=2$. The group $H_1=\langle s, C_H(s)\rangle$ is contained in S_β. If s^2 does not belong to $C_H(s)$ then $|H_1| \geqq 3|C_H(s)|$; but $|H|=2|C_H(s)|$ and so H has not maximum order in S_β, contrary to hypothesis. We conclude that $s^2 \in C_H(s) \subset H$.

We next consider the case $q=(H:C_H(s)) \geqq 3$. If s^2 does not belong to H then Hs and Hs^{-1} are distinct cosets. Since $C_H(s)=C_H(s^{-1})$, the only possible structure of G is subject to condition (3.12)V. In particular

$$(H:C_H(s))=(H:C_H(s^{-1}))=3. \tag{3.14}$$

We may suppose that s^3 belongs to $C_H(s)$, for otherwise $\langle s, C_H(s)\rangle$, which belongs to S_β, has order greater than that of H.

We now put $G_1=\langle s, H\rangle$. Two possibilities arise:

Case (i) H is normal in G_1; here β inverts $s^{-1}hs$ for all $h \in H$, that is,

$$s^{-1}h^{-1}s=(s^{-1}hs)^{-1}=(s^{-1}hs)\beta=sh^{-1}s^{-1}.$$

Hence s^2 commutes elementwise with H, and so $C_H(s^{-1})=H$, which contradicts (3.14).

Case (ii) H is not normal in G; in this case $s^{-1}Hs \neq H$. We choose an element $h \in H$ such that $h \notin C_H(s)$. Then by (3.14),

$$H=\langle h, C_H(s)\rangle, \qquad h^3 \in C_H(s).$$

Now

$$C_H(s)=C_H(sh)=C_H(sh^2),$$

and so each of these centralizers has index 3 in H. Since G must satisfy condition (3.12)V, two of the three cosets Hs, Hsh, Hsh^2 must be equal. But it is easy to see that this implies that H is normal in G, contrary to hypothesis.

We conclude that s^2 must belong to H, and the proof is complete.

(3.15) **The Index 2 Theorem.** *Let A be an Abelian subgroup of maximum order in G. Let β be a $>\frac{1}{2}$-automorphism of G that inverts A elementwise. Then for every $s \in S_\beta$, $s \notin A$, A has index 2 in $\langle s, A\rangle$.*

Proof. Put $G_1=\langle s, A\rangle$ and $(G_1:A)=n$. Two cases arise.

Case (i) $(A:C_A(s))=2$. The centre of G_1 is $Z=C_A(s)$, and G_1/Z has order $2n$ with subgroup A/Z of order 2. Now if G_1/Z contains a coset bZ of order $m>2$, then $B=\langle b, Z\rangle$ is Abelian and $|B|=m|Z|>2|Z|=|A|$, which contradicts the definition of A. Therefore G_1/Z is elementary Abelian, and so A is a normal subgroup of G_1. By the Squares Theorem, $s^2 \in A$. Therefore $n=2$.

Case (ii) $q=(A:C_A(s)) \geqq 3$. Suppose by way of contradiction that $n>2$. Then besides A and As there exists a third coset of A in G_1, and, since $As^2=A$, it must be of the form Asa for some $a \in A$, $a \notin C_A(s)$. Now $C_A(s)=C_A(sa)$, and so G_1 must satisfy condition (3.12)V. Therefore $q=3$ and

$$A=\langle a, C_A(s)\rangle, \qquad a^3 \in C_A(s).$$

But we also have $C_A(s)=C_A(s\,a^2)$ and so, according to (3.12)V, two of the three cosets $A\,s$, $A\,s\,a$, $A\,s\,a^2$ are equal. We deduce that A is normal in G_1. But $s^2 \in A$ and so $n=2$. This contradiction completes the proof.

(3.16) **Remark.** In the last theorem we required A be Abelian of maximum order in G. The assumption that A be of maximum order in S_β (as was required in the Squares Theorem) is not sufficient. For example, in the symmetric group on three symbols, the identity automorphism, ι, is a $\frac{2}{3}$-automorphism, and $H=\langle(12)\rangle$ is a subgroup of maximum order in S_ι. The permutation $s=(13)$ belongs to S_ι, but H has index 3 in $\langle s, H\rangle$.

We reserve the letter A to denote an Abelian subgroup of maximum order in G.

4. The Structure of Non-Abelian $>\frac{1}{2}$-Groups

The results of Section 3 relate the subgroup structure of a $>\frac{1}{2}$-group G to sets of elements inverted under a $>\frac{1}{2}$-automorphism. We are now able to develop properties of G that do not refer to specific automorphisms. The Abelian subgroups of maximum order play an important role in the development of the structure of G. Finally, in Theorem (4.13) we obtain a classification of all non-Abelian $>\frac{1}{2}$-groups.

Throughout this section G denotes a $>\frac{1}{2}$-group, and A is an Abelian subgroup of maximum order in G.

(4.1) **Theorem.** *The subgroup A is normal in G and G/A is an elementary Abelian 2-group.*

Proof. By Corollary (3.3) there is a $>\frac{1}{2}$-automorphism β that inverts A elementwise. By the Transversal Theorem (3.5), every element $g \in G$ is expressible in the form $g=a\,s$ for some $a \in A$, $s \in S_\beta$. Clearly $g^{-1}A\,g=s^{-1}A\,s$. Now suppose $g \notin A$. Put $G_1=\langle s, A\rangle$. By Theorem (3.15), $(G_1:A)=2$. Thus A is normal in G_1 and so $g^{-1}A\,g=s^{-1}A\,s=A$. Since g is arbitrary, it follows that A is normal in G. Moreover,

$$g^2=(a\,s)^2=a\,s^2\,(s^{-1}\,a\,s),$$

and so, by the Squares Theorem (3.13), $g^2 \in A$. Thus G/A is an elementary Abelian 2-group.

(4.2) **Centralizer Structure Theorem.** *Let $G=A \cup A\,g_2 \cup \dots \cup A\,g_n$ be a decomposition of G into a union of disjoint cosets of A, and put $q_i=(A: C_A(g_i))$ $(i=2, \dots, n)$. Then, relative to a suitable ordering of the cosets, one of the following conditions must hold (the corresponding values of $l(G)$ are indicated in brackets):*

I* $\quad n=2;\quad \left(l(G)=\dfrac{q_2+1}{2q_2}\right)$.

II* $\quad n=2^k\ (k \geqq 2),\quad q_i=2\ (i=2, \dots, 2^k);\quad \left(l(G)=\dfrac{2^k+1}{2^{k+1}}\right)$.

III* $\quad n=4,\quad q_2=4,\quad q_3=q_4=2;\quad \left(l(G)=\frac{9}{16}\right)$.

Proof. We have already established conditions (3.12) and it remains to show that some of the cases listed there are impossible in $>\frac{1}{2}$-groups. Firstly, according to Theorem (4.1), the index n must be a power of 2. Next we rule out conditions (3.12) IV and V. For suppose that $q_2=(A:C_A(g_2))=3$ and $q_3=(A:C_A(g_3))=3$, 4 or 5. Since G/A is elementary Abelian, the cosets Ag_2, Ag_3 and Ag_2g_3 are distinct. Assuming that condition IV or V holds, we must have $(A:C_A(g_2g_3))=2$. Now, since $g_2^2\in A$,

$$C_A(g_3)=C_A(g_2^2g_3)\supseteq C_A(g_2)\cap C_A(g_2g_3)=B, \quad \text{say}.$$

Clearly B has index 6 in A and so $C_A(g_3)$ cannot have index 4 or 5 in A. So we may assume $q_2=q_3=3$. But then

$$C_A(g_2g_3)\supseteq C_A(g_2)\cap C_A(g_3)=C, \quad \text{say},$$

where C has index 3 or 9 in A. Since C cannot be contained in a subgroup of index 2 in A, the case $q_2=q_3=3$ is ruled out.

Next we consider condition III. Firstly, from the condition $q_i=2$ $(i\geqq 3)$ it follows that $q_2=2$ or 4, for

$$C_A(g_2)=C_A(g_2g_3^2)\supseteq C_A(g_2g_3)\cap C_A(g_3),$$

and since g_2g_3 may be taken as the fourth coset representative, this last intersection has index 2 or 4 in A.

The case $n=4$, $q_2=4$, $q_3=q_4=2$ arises in $>\frac{1}{2}$-groups, but $n=2^k$ $(k>2)$, $q_2=4$, $q_i=2$ $(i=3,\ldots,2^k)$ is impossible, as we shall now prove.

It is convenient at this point to introduce a notation that exhibits G/A as an elementary Abelian 2-group. Suppose G/A has order 2^k $(k>2)$ and is generated by $x_1A, x_2A, \ldots, x_kA$. We select x_1 such that $(A:C_A(x_1))=4$ and assume that $(A:C_A(x))=2$ for all $x\notin A\cup x_1A$. Our proof is based on the following observation:

(4.3) *Suppose* $xA\neq yA$ *and* $C_A(x)$, $C_A(y)$ *both have index* 2 *in* A. *Then*

$$C_A(x)=C_A(y)\Rightarrow C_A(xy)=C_A(x).$$

For, under the hypothesis, $C_A(xy)\supseteq C_A(x)\cap C_A(y)=C_A(x)$, and, since $xy\notin A$, the possibility $C_A(xy)=A$ is ruled out.

Now the elements x_2, x_3, x_2x_3, x_1x_2, x_1x_3, and $x_1x_2x_3$ belong to distinct cosets, and their centralizers in A have index 2 in A. Moreover

$$C_A(x_1)=C_A(x_1x_2^2)\supseteq C_A(x_1x_2)\cap C_A(x_2)$$

and since $C_A(x_1)$ has index 4 in A, we have equality here. By an extension of this argument we obtain

$$C_A(x_1)=C_A(x_1x_2)\cap C_A(x_2)=C_A(x_1x_3)\cap C_A(x_3)=C_A(x_1x_2x_3)\cap C_A(x_2x_3).$$

Thus the centralizers in A of each of the 6 listed elements contains $C_A(x_1)$ and

(4.4) $$C_A(x_1x_2)\neq C_A(x_2), \qquad C_A(x_1x_3)\neq C_A(x_3), \qquad C_A(x_1x_2x_3)\neq C_A(x_2x_3).$$

Now $A/C_A(x_1)$ is elementary Abelian of order 4 and so contains 3 subgroups of index 2.

Therefore the centralizers of the 6 listed elements are distributed among a set of 3 subgroups of index 2 in A.

By (4.3), either $C_A(x_2)$, $C_A(x_3)$, and $C_A(x_2 x_3)$ are all equal or they are all different. Both cases lead to a contradiction of (4.4) by application of (4.3). We omit the verification, which completes the proof of the theorem.

With the help of Theorem (4.2) we will be able to obtain a complete classification of $>\frac{1}{2}$-groups. We shall see that each of the cases I*, II*, and III* leads to a class of such groups, which we call groups of type I*, II*, III* respectively.

Firstly, *every group G of type* I* *has a $>\frac{1}{2}$-automorphism.* G contains an Abelian subgroup A of index 2 in G. Let $G = A \cup Ax$. The mapping

$$a \to a^{-1}, \quad ax \to a^{-1}x^{-1}, \quad a \in A$$

defines a $\dfrac{q+1}{2q}$-automorphism of G, where $q = (A: C_A(x))$.

To determine the $>\frac{1}{2}$-groups of type II* and III* we require further analysis.

(4.5) **Theorem.** *Let G be a $>\frac{1}{2}$-group of type* II* *or* III*. *If x and y lie in different cosets of A in G, then $C_A(x) \neq C_A(y)$.*

Proof. Assume to the contrary that $Ax \neq Ay$, but $C_A(x) = C_A(y)$. Then neither x nor y belongs to A, and $(A: C_A(x)) = 2$. Let $G_1 = \langle x, y, A \rangle$. By Theorem (3.2), G_1 is a $>\frac{1}{2}$-group and it is clearly of type II* or III*. Now

$$C_A(xy) \supseteq C_A(x) \cap C_A(y) = C_A(x).$$

But $xy \notin A$, and so $C_A(xy) = C_A(x)$.

Thus we have

$$C_A(x) = C_A(y) = C_A(xy) = Z, \quad \text{say},$$

where $(A:Z) = 2$. This is a contradiction if G_1 is of type III* and since in that case $G_1 = G$, we may assume for the rest of the proof that G and G_1 are of type II*.

Clearly Z is the centre of G_1 and since it has index 2 in A we may put $A = \langle a_1, Z \rangle$ with $a_1^2 \in Z$. We observe that $C_G(a_1) = A$, for A must be equal to its own centralizer. In particular,

$$x^{-1} a_1 x = a_1 z_1, \quad y^{-1} a_1 y = a_1 z_2, \tag{4.6}$$

where z_1 and z_2 are distinct elements of order 2 in Z. (They cannot be equal, for yx^{-1} does not commute with a_1.)

We note further that

$$xy \neq yx \tag{4.7}$$

for if x and y commute then $\langle x, y, Z\rangle$ is Abelian and has order $2|A|$ which contradicts the definition of A.

Let α be a $>\frac{1}{2}$-automorphism of G_1 that inverts A elementwise. By the Transversal Theorem (3.5) we may suppose that α inverts both x and y, and so

$$(4.8) \qquad (a\,x\,y)\,\alpha = a^{-1}x^{-1}y^{-1} \quad \text{for all } a \in A.$$

Now α inverts half the elements of the coset $A\,x\,y$ and since by (4.7) and (4.8) no element of $Z\,x\,y$ is inverted, we must have

$$(a_1\,x\,y)^{-1} = (a_1\,x\,y)\,\alpha = a_1^{-1}x^{-1}y^{-1}.$$

It follows from (4.6) that

$$a_1\,x\,y = y\,x\,a_1 = a_1\,z_1\,z_2\,y\,x,$$

and so

$$[x, y] = z_1\,z_2.$$

Now

$$[x\,a_1, y\,a_1] = [x, y]\,[x, a_1]\,[a_1, y] = (z_1\,z_2)\,z_1\,z_2 = 1.$$

Therefore $\langle x\,a_1, y\,a_1, Z\rangle$ is Abelian and has order $2|A|$. We have obtained a contradiction, and the proof is complete.

(4.9) **Corollary.** *Suppose G is a $>\frac{1}{2}$-group of type* II* *or* III* *such that G/A is elementary Abelian of order 2^k $(k \geqq 2)$ and*

$$G/A = \langle x_1 A, x_2 A, \ldots, x_k A\rangle.$$

Put $Z = C_A(x_1) \cap C_A(x_2) \cap \cdots \cap C_A(x_k)$. Then Z is the centre of G and the index of Z in A is 2^k. Moreover, A/Z is elementary Abelian.

Proof. It is clear that Z is the centre of G and that $|A/Z| \leqq 2^k$. Consider first G of type II*. By Theorem (4.5) A/Z has $2^k - 1$ distinct subgroups of index 2. Therefore it must be elementary Abelian of order 2^k.

Finally if G has type III* then two of the centralizers $C_A(x_1)$, $C_A(x_2)$, $C_A(x_1 x_2)$ have index 2 in A and intersect in the third, which is equal to Z. Thus in this case A/Z has order 4 and is elementary Abelian.

Our next result concerns the action of G/A on A.

(4.10) **Lemma.** *Let G and G/A be defined as in Corollary* (4.9), *and suppose that if G is of type* III*, *generators x_1, x_2 are so chosen that $C_A(x_1)$ and $C_A(x_2)$ have index* 2 *in A. Then for $i = 1, \ldots, k$ there exists $a_i \in A$ and $z_i \in Z$ such that*

$$[a_i, x_i] = z_i, \qquad [a_i, x_j] = 1, \qquad (j \neq i)$$

with $z_i \neq 1$, $z_i^2 = 1$.

Moreover, if G has type II* *then $z_1 = z_2 = \cdots = z_k$, whereas if G has type* III* *then $z_1 \neq z_2$.*

Proof. Put

$$D_i = C_A(x_1) \cap \cdots \cap C_A(x_{i-1}) \cap C_A(x_{i+1}) \cap \cdots \cap C_A(x_k).$$

Now $(D_i:Z)=2$ and $C_A(x_i)\cap D_i=Z$, so we may choose $a_i\in D_i\setminus C_A(x_i)$, with the property that $[a_i,x_j]=1$ when $j\neq i$, and $[a_i,x_i]\neq 1$. We now show that $[a_i,x_i]\in Z$.

From a well-known commutator identity ([1], p. 150) we obtain

$$[x_i^{-1},x_j^{-1},a_i]^{x_j}[x_j,a_i^{-1},x_i^{-1}]^{a_i}[a_i,x_i,x_j]^{x_i^{-1}}=1,$$

and since $[x_i^{-1},x_j^{-1}]\in A$ and $[x_j,a_i^{-1}]=1$ $(i\neq j)$, it follows that $[a_i,x_i]\in D_i$. A simple calculation shows that $[a_i,x_i]\in Z$ and leads to the first statement of the lemma.

Suppose now that G has type II*. Consider the centralizer $C_A(x_i x_j)$ for $i\neq j$. By what has already been proved this group contains a_m $(m\neq i$ or $j)$ but does not contain a_i or a_j. Since it has index 2 in A, we are forced to the conclusion that $a_i a_j\in C_A(x_i x_j)$. Thus

$$1=[x_i x_j, a_i a_j]=z_i z_j,$$

and so $z_i=z_j$.

On the other hand, if G has type III* then $C_A(x_1 x_2)$ has index 4 in A, and so $z_1 z_2\neq 1$. This completes the proof.

An immediate consequence of Lemma (4.10) is

(4.11) **Corollary.** *For $>\frac{1}{2}$-groups G not of type* I*, $[G,A]$ *lies in the centre of G, and has order* 2 *or is non-cyclic of order* 4 *according as G has type* II* *or* III*.

We need one further result before we can give the structure of $>\frac{1}{2}$-groups.

(4.12) **Lemma.** *With the notation of Corollary* (4.9) *and Lemma* (4.10), *the elements $x_1,x_2,\ldots,x_k$ can be chosen to commute pairwise.*

Proof. Consider first G of type II*. By Corollary (4.11), $[G,A]$ is generated by an element z, say, which belongs to Z and has order 2. We show first that for all i,j, $[x_i,x_j]=z$ or 1. For consider $A_j=\langle x_j, C_A(x_j)\rangle$. This is an Abelian subgroup of maximum order in G, for clearly $|A_j|=|A|$. By Theorem (4.1), A_j is normal in G and G/A_j is elementary Abelian, generated by a_jA_j, and the cosets x_iA_j for all $i=1,\ldots,k$ except $i=j$. By Corollary (4.11), $[G,A_j]$ has order 2, and since it contains $[a_j,x_j]=z$, we conclude that $[G,A_j]=\langle z\rangle$. Therefore $[x_i,x_j]=z$ or 1.

We can now prove by induction that the coset representatives $x_1,\ldots,x_k$ of $Ax_1,\ldots,Ax_k$ can be so chosen that for $i=1,\ldots,k$, x_i commutes with x_j $(j=1,\ldots,k)$. Consider the case $i=1$. If $[x_1,x_j]=z$, then we replace x_j by a_1x_j as coset representative of Ax_j, and obtain $[x_1,a_1x_j]=[x_1,a_1][x_1,x_j]=z^2=1$. We note that the elements $a_1,\ldots,a_k$ constructed in Lemma (4.10) satisfy the same commutator relations with the new coset representatives as they did with the old.

Now suppose that we have already chosen $x_1,\ldots,x_k$ such that $x_1,x_2,\ldots,x_{i-1}$ commute with x_j $(j=1,\ldots,k)$. In particular then x_i commutes with each of $x_1,\ldots,x_{i-1}$. For every $j>i$ such that $[x_i,x_j]=z$ we replace x_j by a_ix_j, and obtain $[x_i, a_i x_j] = z^2 = 1$. Thus we construct new coset representatives which commute with x_i, and since they clearly commute with each of $x_1, \ldots, x_{i-1}$, the proof by induction is complete.

If G has type III*, we form $A_1 = \langle x_1, C_A(x_1) \rangle$ and by an argument similar to the above we conclude that $[x_1, x_2] = 1$, z_1, z_2 or $z_1 z_2$. Depending on which of these cases occurs we find that one of the following pairs of coset representatives commute: $\{x_1, x_2\}$, $\{x_1, a_1 x_2\}$, $\{a_2 x_1, x_2\}$, $\{a_2 x_1, a_1 x_2\}$. This completes the proof.

We summarise our findings in the following theorem.

(4.13) **Structure Theorem.** *A non-Abelian $>\frac{1}{2}$-group G is one of the following types.*

Type I*. *G has an Abelian subgroup A of index 2 in G. For every such group if $G = A \cup Ax$ then the map α: $a \to a^{-1}$, $ax \to a^{-1} x^{-1}$ $(a \in A)$ defines a $>\frac{1}{2}$-automorphism of G with*

$$l(\alpha) = l(G) = \frac{q+1}{2q}, \qquad \text{where } q = (A : C_A(x)).$$

Type II*. *G is nilpotent of class 2. It has commutator subgroup $\langle z \rangle$ of order 2. Its centre Z has index 2^{2k} $(k \geqq 2)$ in G, and G/Z is an elementary Abelian 2-group, generated by $x_1 Z, x_2 Z, \ldots, x_k Z$, $a_1 Z, a_2 Z, \ldots, a_k Z$ subject to the following commutator relations:*

$$\begin{aligned} &[x_i, x_j] = [a_i, a_j] = 1 \qquad \text{for all } i, j = 1, \ldots, k. \\ &[a_i, x_j] = 1 \quad (i \neq j), \\ &[a_i, x_i] = z. \end{aligned}$$

Every such group has a $>\frac{1}{2}$-automorphism α defined by the map

$$a x_1^{\varepsilon_1} x_2^{\varepsilon_2} \ldots x_k^{\varepsilon_k} \to a^{-1} x_1^{-\varepsilon_1} x_2^{-\varepsilon_2} \ldots x_k^{-\varepsilon_k}$$

for all $a \in A = \langle a_1, \ldots, a_k, Z \rangle$ and $\varepsilon_i = 0$ or 1 $(i = 1, \ldots, k)$. Moreover,

$$l(\alpha) = l(G) = \frac{2^k + 1}{2^{k+1}}.$$

Type III*. *G is nilpotent of class 2. It has elementary Abelian commutator subgroup $\langle z_1, z_2 \rangle$ of order 4. Its centre Z has index 2^4 in G and G/Z is an elementary Abelian 2-group, generated by $x_1 Z$, $x_2 Z$, $a_1 Z$, $a_2 Z$ subject to the commutator relations*

$$\begin{gathered} [x_1, x_2] = [a_1, a_2] = [a_1, x_2] = [a_2, x_1] = 1, \\ [a_1, x_1] = z_1, \qquad [a_2, x_2] = z_2. \end{gathered}$$

Every such group has a $>\frac{1}{2}$-automorphism α defined by the map

$$a x_1^{\varepsilon_1} x_2^{\varepsilon_2} \to a^{-1} x_1^{-\varepsilon_1} x_2^{-\varepsilon_2}$$

for all $a \in A = \langle a_1, a_2, Z\rangle$ *and* $\varepsilon_i = 0$ *or* 1 $(i=1,2)$. *Moreover,*

$$l(\alpha) = l(G) = \tfrac{9}{16}.$$

Note that we do not specify the orders of the x_i's and the a_i's in groups of type II* and III*. The only condition imposed on powers of these elements is that their squares must lie in the centre Z of G, for clearly $[a_i^2, x_j] = [a_i, x_j^2] = [a_i, x_j]^2 = 1$ for all i, j. It follows immediately from the structure theorem that in a $>\frac{1}{2}$-group G of type II* or III* the elements of odd order form a subgroup Z_0 of the centre and G splits into a direct product of Z_0 and a 2-group.

5. Groups Consisting Mostly of Involutions

In [7] Wall constructed all groups of a given even order that possess the largest possible number of involutions. We obtain a solution of Wall's problem by classifying all groups in which the identity automorphism is a $>\frac{1}{2}$-automorphism. It turns out that groups of arbitrary even order exist with this property. Clearly such groups are non-Abelian, except for the trivial case of elementary Abelian 2-groups all of whose non-identity elements are involutions.

Throughout this section we let G denote a non-Abelian group such that the identity automorphism, ι, is a $>\frac{1}{2}$-automorphism. Thus G has precisely $l(G)\,|G| - 1 \geqq \frac{1}{2}|G|$ involutions.

Type I*. Two cases arise. If A is an Abelian subgroup of maximum order in G then

either (i) every element of A is mapped by ι onto its inverse, and so A is an elementary Abelian 2-group;

or (ii) A is not elementary Abelian. In this case, by the proof of Theorem (3.2) with $\alpha = \iota$, there is an inner automorphism I_x $(x \in S_\iota)$ that inverts A elementwise. In other words, there exists an involution x such that

$$x^{-1} a x = a^{-1} \quad \text{for all } a \in A,$$

and so

$$(x a)^2 = 1 \quad \text{for all } a \in A.$$

In either case the centre Z of G is an elementary Abelian 2-group.

We now employ Theorem (4.13) to obtain the structure of G.

In case (i), G has an elementary Abelian subgroup A of index 2 and an involution $x \notin A$ which induces by conjugation an arbitrary automorphism of order 2 in A. These groups appear in class IV of Wall's classification.

In case (ii), G has an Abelian subgroup A of index 2 and arbitrary order which is not elementary Abelian, and an involution $x \notin A$ which induces in A the automorphism that inverts A elementwise. These groups appear in class I of Wall's classification.

Clearly in this case G may be chosen to have arbitrary even order $\geqq 6$.

Type II*. The centre Z of G may be written as a direct product $Z = \langle z\rangle \times E$ where z generates the commutator subgroup of G, and E is an elementary

Abelian 2-group (if non-trivial). Clearly E splits from G for otherwise at least half the elements of G have order >2. So $G=G_0\times E$, where G_0 is a type II* group with centre $(G_0)'$ of order 2. We show that G_0 has an Abelian subgroup of maximum order that is elementary Abelian. For the square of every element of G_0 is equal to z or 1 and so one of a_i, x_i, $a_i x_i$ has order 2. We select such an element for each i ($i=1, \ldots, k$). They generate with z the required elementary Abelian subgroup.

It follows that G_0 has the presentation

$$G_0=\langle z, x_1, \ldots, x_k, a_1, \ldots, a_k | z^2=x_i^2=a_i^2=1,$$
$$\textit{all pairs of generators commute except } [a_i, x_i]=z \ \ (i=1, \ldots, k)\rangle.$$

These are Wall's class III groups. As he points out, they are the product of k dihedral groups of order 8 with the centres amalgamated.

Type III*. Here $G=G_0\times E$, where G_0 is a type III* group whose centre $(G_0)'$ is a four-group. There is only one group G_0 of this type that admits ι as a $>\frac{1}{2}$-automorphism. It is Wall's class II group: the direct product of two dihedral groups of order 8. The analysis is similar to the above, and we omit it.

References

1. Hall, M.: The theory of groups. New York: MacMillan 1959.
2. Manning, W.A.: Groups in which a large number of operators may correspond to their inverses. Trans. Amer. Math. Soc. **7**, 233–240 (1906).
3. Miller, G.A.: Groups containing the largest possible number of operators of order two. Amer. Math. Monthly **12**, 149–151 (1905).
4. — Groups of order 2^m which contain a relatively large number of operators of order two. Amer. J. Math. **42**, 1–10 (1920).
5. — Non-Abelian groups admitting more than half inverse correspondences. Proc. Nat. Acad. Sci. **16**, 168–172 (1930).
6. — Groups which admit five-eights automorphisms. Proc. Nat. Acad. Sci. **17**, 39–43 (1931).
7. Wall, C.T.C.: On groups consisting mostly of involutions. Proc. Cambridge Philos. Soc. **67**, 251–262 (1970).

Dr. Hans Liebeck
Mr. Desmond MacHale
Department of Mathematics
University of Keele
Staffs, ST5 5BG
England

(Received May 3, 1971)

Math. Z. 124, 64–66 (1972)

A Note on the Covering Groups of Σ_n, $n \geqq 6$

FREDRIK L. SMITH

Recently, Thompson brought forth the problem to determine all finite simple groups in which the centralizer of an involution is isomorphic to one of the covering groups of Σ_n, the symmetric group on n letters.

Since the outer automorphism group of the new Lyons' simple group Ly is trivial [3], it seemed possible that Ly might be imbedded in some larger simple group which possessed the above property.

Since the answer to this question is known for $n=4$ and 5, we have restricted ourselves to the case that $n \geqq 6$ and we have shown that there are no simple groups with this property when $n \geqq 6$.

Our notation is standard (see [2]) and includes the "bar convention" for denoting homomorphic images.

Theorem. *Let G be a finite group of even order and let z be an involution in G. Let $H=C_H(z)$ and assume that $H/\langle z\rangle \cong \Sigma_n$, $n \geqq 6$, and $z \in H'$. Then G is not simple.*

Proof. We assume, by way of contradiction, that G is a simple group. If $K=H'$, then K is a perfect group, $K/\langle z\rangle \cong A_n$, and the structure of K is uniquely determined [4]. If T is an S_2-subgroup of H, then $Z(T)$ is cyclic and so T is an S_2-subgroup of G.

By the results of Schur [4], there are two possibilities for the structure of H and these are non-isomorphic for $n>6$, however they are isomorphic for $n=6$. We now proceed to give these structures. If π is a permutation in Σ_n, then $\bar{\pi}$ will denote the coset in $\bar{H}=H/\langle z\rangle$ which corresponds to π and a_π will denote an element in $\bar{\pi}$. We then have for suitable choices of a_π:

1) $H=\langle a_{(12)}, a_{(23)}, \ldots, a_{(n-1\cdot n)} \mid 1=(a_{(i\cdot i+1)})^2=z^2,$
$i=1,2,\ldots,n-1;$
$(a_{(i\cdot i+1)}\, a_{(i+1\cdot i+2)})^3=1,\ [a_{(i\cdot i+1)}, a_{(j\cdot j+1)}]=z,$
$i=1,2,\ldots,n-2, j>i+1, n-1 \geqq j\rangle.$

2) $H=\langle a_{(12)}, \ldots, a_{(n-1\cdot n)} \mid (a_{(i\cdot i+1)})^2=z,\ z^2=1,$
$i=1,2,\ldots,n-1;$
$(a_{(i\cdot i+1)}\, a_{(i+1\cdot i+2)})^3=z,\ [a_{(i\cdot i+1)}, a_{(j\cdot j+1)}]=z,$
$i=1,2,\ldots,n-2, j>i+1, n-1>j\rangle.$

Suppose there are involutions in $K-\langle z\rangle$ and that x is one. Then we must have $n \geqq 8$ and a conjugate x_0 of x is contained in

$$\bar{x}_0=\overline{(12)}\,\overline{(34)}\,\overline{(56)}\,\overline{(78)}\ldots\overline{(8r-1\cdot 8r)}, \qquad r \geqq 1.$$

We also have $x_0^{a_{(12)}} = x_0 z$. If

$$\bar{w} = \overline{(1324)}\,\overline{(5768)} \dots \overline{(8r-3 \cdot 8r-1 \cdot 8r-2 \cdot 8r)},$$

then $\bar{w}^2 = \bar{x}_0$ and so x_0 has a square root in K. It follows that every involution in $K - \langle z \rangle$ has a square root in K.

We also have $a_{(12)} a_{(34)}$ centralizes x_0 and $z = (a_{(12)} a_{(34)})^2$ and hence, z has a square root in $C_K(x)$ for every involution x in $K - \langle z \rangle$.

Set $a = a_{(12)}$ and $b = a_{(12)} a_{(34)} a_{(56)}$. Then $a^2 = 1$ if H is given by 1) and $b^2 = 1$ if H is given by 2). We see that $a_{(34} a_{(56)}$ centralizes both a and b in either case and so z has a square root in $C_H(a)$ and in $C_H(b)$ in either case.

Suppose $z \underset{G}{\sim} a$ or $z \underset{G}{\sim} b$ and let $H_\alpha = C_G(\alpha)$, $K_\alpha = H'_\alpha$ where $\alpha = a$ or b and $z \underset{G}{\sim} \alpha$. Since z has a square root in $C_H(\alpha)$, we have $z \in K_\alpha - \langle \alpha \rangle$. By the above remarks, we then have that α has a square root in $C_{K_\alpha}(z)$ and hence, in H. This however is a contradiction.

By Lemma 5.38 of [5], we see that in any case there are involutions in $K - \langle z \rangle$, since in 1) a is conjugate to an involution in K and in 2) b has this property.

Suppose we are in case 1) so that $a^2 = 1$ and $a \underset{G}{\sim} x \in K - \langle z \rangle$. Let T_0 be an S_2-subgroup of $C_H(a)$ and set $T_1 = T_0 \cap K$. Since $a^{a_{(34)}} = a z$, we see that $C_H(a) = \langle a \rangle \times C_K(a)$ and $C_K(a)$ is isomorphic to the covering group of A_{n-2}. It follows that $T_0 = \langle a \rangle \times T_1$. Since x has a square root in H, we conclude that T_0 is not an S_2-subgroup of $C_G(a)$. Since $z \nsim a$, we have $\Omega_1(Z(T_1)) \supsetneqq \langle z \rangle$. Let u be an involution in $Z(T_1) - \langle z \rangle$. We can assume without loss of generality that

$$\bar{u} = \overline{(34)}\,\overline{(56)}\,\overline{(78)}\,\overline{(9 \cdot 10)} \dots \overline{(8r+1 \cdot 8r+2)}$$

for some $r \geqq 1$. Then if $\bar{\pi} = (35)(46)$, $\bar{\pi}$ centralizes $\bar{u}$, a_π centralizes a, but a_π does not centralize u. This however contradicts the fact that u is central in an S_2-subgroup of $C_K(a)$. This rules out case 1).

We now assume that we are in case 2) and so $b^2 = 1$. If $\rho = (13)(24)$, then $b^{a_\rho} = b z$. Again, we see that $C_H(b) = \langle b \rangle \times C_K(b)$. If T_0 is an S_2-subgroup of $C_H(b)$ and $T_1 = T_0 \cap K$, then $T_0 = \langle b \rangle \times T_1$ and since $b \underset{G}{\sim} x$, $x \in K$, we have that T_0 is not an S_2-subgroup of $C_G(b)$. It follows as in the preceding case that $\Omega_1(Z(T_1)) \supsetneqq \langle z \rangle$. If $n \leqq 9$, then $K - \langle z \rangle$ contains a single class of involutions. Since $z \in \Omega_1(Z(T_0)) \cap \Phi(T_0)$ and $\Phi(T_0) \subseteq K$, we must have that z is conjugate in $C_G(b)$ to an involution in K. Since $z \underset{G}{\nsim} b$, we have a contradiction and so $n \geqq 10$. Thus we see that $\overline{C_K(b)} = (\bar{X}_1 \times \bar{X}_2) \langle \bar{r} \rangle$ where $X_1 = \langle (12)(34), (12)(56), (135)(246) \rangle$, X_2 is the alternating group on $\{7, 8, 9, \dots, n\}$, and $r = (13)(24)(56)(78)$. If u is an involution in $Z(T_1) - \langle z \rangle$, then we can assume without loss of generality that

$$\bar{u} = \overline{(12)}\,\overline{(34)}\,\overline{(78)}\,\overline{(9 \cdot 10)} \dots \overline{(8r+1 \cdot 8r+2)}, \qquad r \geqq 1$$

or

$$\overline{(78)}\,\overline{(9 \cdot 10)}\,\overline{(11 \cdot 12)}\,\overline{(13 \cdot 14)} \dots \overline{(8r-3 \cdot 8r-2)}, \qquad r \geqq 2.$$

If $\sigma = (78)(9 \cdot 10)$, then $\bar{\sigma}$ centralizes $\bar{u}$, a_σ centralizes b, but a_σ does not centralize u. However, this leads to the same contradiction as in the preceding case. This completes the proof of our theorem.

We remark here that if $n=4$ in the hypothesis of our theorem, then $G \cong PSL(3,3)$ or M_{11} and if $n=5$, then $G \cong U_3(5)$ by the results in [1].

References

1. Alperin, J., Brauer, R., Gorenstein, D.: Finite groups with quasi-dihedral and wreathed Sylow 2-subgroups (to appear in Trans. Amer. Math. Soc.).
2. Gorenstein, D.: Finite groups. New York: Harper and Row 1968.
3. Lyons, R.: Evidence for the existence of a new simple group (manuscript).
4. Schur, I.: Über die Darstellungen der symmetrischen and alternierenden Gruppen durch gebrochene lineare Substitutionen. J. Math. **139**, 155–250 (1911).
5. Thompson, J. G.: Non-solvable finite groups all of whose local subgroups are solvable. Bull. Amer. Math. Soc. **74**, 383–437 (1968).

Mr. Fredrik L. Smith
Department of Mathematics
Ohio State University
Columbus, Ohio 43210
USA

(Received June 1, 1971)

Math. Z. 124, 67–72 (1972)

Solubility of Groups Admitting a Fixed-point-free Automorphism Group of Type (p, p)*

R. PATRICK MARTINEAU

Introduction

If A is a group of automorphisms of a finite group G, then we say that A acts fixed-point-freely on G if $C_G(A)=1$. An important theorem of Thompson states that in this situation, if A has prime order, then G is nilpotent. In this paper, we prove a result in the case in which A is the direct product of two cyclic groups of the same prime order. Specifically, we prove the following

Theorem. *Let G be a finite group admitting an elementary abelian fixed-point-free group of automorphisms V of order r^2, r a prime. Then G has a normal subgroup F such that both F and G/F are nilpotent.*

Remarks. 1. In a previous paper [2], I proved a similar result under stronger conditions. This paper supersedes that discussion, though several ideas in the proof are the same.

2. The case $r=2$ has already been proved by Glauberman and Bauman ([1], Theorem 10.5.6).

The proof is by induction on the order of G. A theorem of Ward [4] enables us to assume the result in the case that G is soluble. Then if P and Q are V-invariant Sylow p- and q-subgroups of a minimal counter-example G to the theorem with $PQ \neq QP$, we make a detailed study of maximal V-invariant p, q-subgroups of G to find that one of p, q is 2, and that some non-1 element of V centralizes a whole Sylow 2-subgroup of G. Then G has a V-invariant soluble 2-complement L, and we show that L contains a non-1 V-invariant normal subgroup of G, contrary to the minimality of G.

1. Preliminaries

Throughout, all groups are finite. If G is a group, $Z(G)$ denotes the centre of G. If H and X are subsets of G, $C_H(X)$ and $N_H(X)$ denote respectively the centralizer and normalizer of X in H. If the subscript is omitted, G is to be read. $[H, X]$ denotes the group generated by all elements $h^{-1}x^{-1}hx$ for $h \in H$, $x \in X$. If H is a normal subgroup of G, we write $H \lhd G$.

If p is a prime and P a Sylow p-subgroup of G, we write $P \in \mathscr{S}_p(G)$ and $|G|_p = |P|$, and if $P \lhd G$ we say that G is p-closed.

* Research supported by a grant from the National Science Foundation.

If π is a set of primes then G is a π-group if its order is divisible only by primes in π. H is a Hall π-subgroup of G if H is a π-group and the index of H in G is not divisible by any prime in π. $O_\pi(G)$ is the largest normal π-subgroup of G.

$\Phi(G)$ denotes the Frattini subgroup of G, the intersection of all maximal subgroups of G. $F(G)$ denotes the Fitting subgroup of G, the largest normal nilpotent subgroup of G. If A_k is a subset of G for each $k \in K$, then $\langle A_k\colon k \in K\rangle$ denotes the subgroup of G generated by all those A_k.

If α is an automorphism of a group G we denote the image of $g \in G$ under α by g^α, and if A is a set of automorphisms of G, we define $C_G(A)=\{g\colon g^\alpha = g$ for all $\alpha \in A\}$. A is said to act fixed-point-freely on G if $C_G(A)=1$.

Throughout this note we use the following notation: V is the direct product of two (cyclic) groups of order r, r a prime; α_i $(i=0, 1, \dots, r)$ are generators of the $r+1$ subgroups of V of order r; if H is a group admitting V as a group of operators, H_i denotes $C_H(\alpha_i)$.

Let G be a group admitting V as a fixed-point-free group of automorphisms. Then the following facts are well-known, and will be used without reference.

$|G|$ and $|V|$ are coprime.

If H is a V-invariant normal subgroup of G, then V induces a fixed-point-free group of automorphisms of G/H; more precisely $C_{G/H}(\alpha_i) = G_i H/H$ for each i. ([1], Theorem 6.2.2)

$G = \langle G_0, G_1, \dots, G_r\rangle$ ([1], Theorem 6.3.4).

For each prime p, G has a unique V-invariant Sylow p-subgroup P and P contains every V-invariant p-subgroup of G. ([1], Theorem 6.2.2.)

We use frequently the result of Thompson (see, for example, [1], Theorem 10.2.1) that if a group G admits a fixed-point-free automorphism of prime order, then G is nilpotent. We shall also use frequently the above-mentioned theorem of Ward:

If G is a finite soluble group admitting V as a fixed-point-free group of automorphisms, then $G/F(G)$ is nilpotent.

Remark. This is a variant of Theorem 2 of [4]. The condition that G be soluble is omitted in Ward's statement of his Theorem 2, but is necessary in the proof. Also, Ward in fact proves more than we have stated above—he merely requires that G_i be nilpotent for all i, and not that V should act fixed-point-freely on G, and he also gives more information about the factor group $G/F(G)$.

For Lemmas 1 to 5 let G be a group admitting V as a fixed-point-free group of automorphisms.

Lemma 1. *If H is a V-invariant π-subgroup of G with $H_k = 1$ for some k, and K is a V-invariant π'-subgroup of $N(H)$, then $K = C_K(H) \cdot K_k$.*

Proof. If $j \neq k$, HK_j is a group on which α_k acts without fixed points, so $[K_j, H]=1$ by Thompson's theorem. The lemma is proved.

As a corollary we have

Lemma 2. *If H is a V-invariant π-subgroup of G and $H_k=1$ for two values of k, then $N(H)/C(H)$ is a π-group.*

The following corollary of Lemma 2 is particularly useful

Lemma 3. *If p is an odd prime, P is the V-invariant Sylow p-subgroup of G, and $P_k=1$ for two values of k, then G has a normal p-complement.*

Proof. By Lemma 2, if H is a characteristic subgroup of P, then $N(H)/C(H)$ is a p-group, so the result follows from the Thompson normal p-complement theorem [3].

Lemma 4. *If G is a $\{p,q\}$-group, then $|C(O_p(G))|_q=|O_q(G)|$.*

Proof. G is soluble by Burnside's $p^a q^b$ theorem, so $G/F(G)$ is nilpotent by Ward's theorem. Thus if P and Q are the V-invariant Sylow p- and q-subgroups of G then $Q=O_q(G)\cdot N_Q(P)$, so that $C_Q(O_p(G))=O_q(G)\cdot[C_Q(O_p(G))\cap N_Q(P)]$. If we let $K=C(O_p(G))\cap N_Q(P)$, then $[K,P]\leqq O_p(G)$ so $[P,K,K]=1$. Thus $[P,K]=1$ (Theorem 5.3.6 of [1]) so $K\leqq O_{p'}(G)=O_q(G)$.

Lemma 5. *If G is a p-group, then $\langle G_j\colon j\neq k\rangle \lhd G$.*

Proof. Certainly $[G,\alpha_k]\lhd G$. Let $M=[G,\alpha_k]$; then $M=[M,\alpha_k]$, so α_k acts without fixed points on $M/\Phi(M)$. Thus $M=\langle \Phi(M), M_j\colon j\neq k\rangle=\langle M_j\colon j\neq k\rangle$. If $j\neq k$, α_k acts fixed-point-freely on G_j, so $G_j=\Phi(G_j)\cdot[G_j,\alpha_k]=[G_j,\alpha_k]\leqq M$. Thus $M=\langle G_j\colon j\neq k\rangle$ and the lemma is proved.

For the rest of this paper let G be a minimal counter-example to the theorem stated above. By the theorem of Ward mentioned above we need only show that G is soluble to obtain a contradiction. If H is a V-invariant normal subgroup of G, then H and G/H satisfy the conditions of the theorem, so G has no non-1 proper V-invariant normal subgroups. We shall use this remark frequently, in the form that if H is a non-1 V-invariant proper subgroup of G then $N(H)$ is soluble.

2. Maximal $\{p,q\}$-Subgroups

In this section we show that for any pair of primes p,q, any maximal V-invariant $\{p,q\}$-subgroup H of G contains a Sylow subgroup, and that if $|H|_q<|G|_q$ then H is p-closed.

We need Thompson's $X\times Y$ lemma ([1], Theorem 5.3.4):

Lemma 6. *If A and C are p-groups, B is a q-group, p and q distinct primes, A and B normalize C and $[B,A]=[B,C_C(A)]=1$, then $[C,B]=1$.*

An idea of Bender is behind the following useful result.

Lemma 7. *Suppose H is a maximal V-invariant $\{p,q\}$-subgroup of G such that $O_p(H)\neq 1$ and $O_q(H)\neq 1$. If K is a non-1 V-invariant subgroup of $F(H)$, then H contains the V-invariant Hall $\{p,q\}$-subgroup of $N(K)$.*

Proof. We may assume without loss of generality that K is a p-group. Let $U=C_{F(H)}(K)=A\times B$ with $A\in\mathscr{S}_p(U)$ and $B\in\mathscr{S}_q(F(H))$. Since H is a maximal

V-invariant $\{p, q\}$-subgroup of G, H contains the V-invariant Hall $\{p, q\}$-subgroup of the normalizer of any non-1 V-invariant normal subgroup of H. In particular H contains any V-invariant $\{p, q\}$-subgroup of $C(A)$ or $C(B)$ since $Z(O_p(H)) \leqq A$ and $B \lhd H$.

Let $\bar{H}$ be any maximal V-invariant $\{p, q\}$-subgroup of G containing U, and let $F(\bar{H}) = \bar{A} \times \bar{B}$ with $\bar{A}$ a p-group and $\bar{B}$ a q-group. We show that $\bar{B} = B$, whence $\bar{H} = H$. This is enough to prove the lemma.

B normalizes $C_{\bar{A}}(A)$. Also $C_{\bar{A}}(A) \leqq H$ so $C_{\bar{A}}(A)$ normalizes B, whence $[B, C_{\bar{A}}(A)] \leqq \bar{A} \cap B = 1$. But B centralizes A, so by Lemma 6 (with $C = \bar{A}$) $[B, \bar{A}] = 1$. Thus by Lemma 4 applied to $\bar{H}$, $B \leqq \bar{B}$. Similarly $A \leqq \bar{A}$, so $O_p(\bar{H}) \neq 1$ and $O_q(\bar{H}) \neq 1$.

Since $\bar{A}$ centralizes B, $\bar{A} \leqq H$. Similarly $\bar{B} \leqq H$, so we have $F(\bar{H}) \leqq H$. But now we can apply the above arguments with H and $\bar{H}$ interchanged and U and $F(\bar{H})$ interchanged to obtain $\bar{B} \leqq B$. We already have $B \leqq \bar{B}$, so the lemma is proved.

For Lemmas 8 to 10 let P and Q be the V-invariant Sylow p- and q-subgroups of G, and assume $PQ \neq QP$. The major use of Lemma 7 is in the following

Lemma 8. *Suppose H is a maximal V-invariant $\{p, q\}$-subgroup of G. Then either $P \lhd H$ or $Q \lhd H$.*

Proof. If $F(H)$ is a p-group, then since $H/F(H)$ is nilpotent H is p-closed, so by maximality of H, $P \lhd H$. Similarly if $F(H)$ is a q-group we have $Q \lhd H$. Thus we may assume that $O_p(H) \neq 1$ and $O_q(H) \neq 1$.

Let $R = O_p(H)$. If $R_i \neq 1$ then, since G_i is nilpotent by Thompson's theorem, Q_i centralizes R_i so Lemma 7 implies $Q_i \leqq H$. Thus if $R_i \neq 1$ for all i then $Q \leqq H$.

If $R_i = 1$ for two values of i, then by Lemma 2 applied to R and H in place of H and G, $|H|_q = |C(R)|_q$. Then by Lemma 4 applied to H, H is q-closed, whence by maximality of H, $Q \lhd H$.

Assume now that $R_k = 1$ for precisely one k. Suppose $j \neq k$; then we have $Q_j \leqq H$ by the second paragraph of this proof. $[Q_j, R] = 1$ by Lemma 1 applied to R in the role of H, so $Q_j \leqq F(H)$ by Lemma 4. Let $Q^* = \langle Q_j \colon j \neq k \rangle \leqq F(H)$. If $Q^* \neq 1$ we apply Lemma 7 to deduce that the V-invariant Sylow q-subgroup of $N(Q^*)$ is contained in H. But by Lemma 5 with q, Q in place of p, G, we have $Q^* \lhd Q$. Thus $Q \leqq H$. If $Q^* = 1$, then by the argument of the third paragraph of this proof applied to $Q \cap F(H)$ instead of R, we have $P \lhd H$.

We have shown that either $P \lhd H$ or $Q \leqq H$. By symmetry we also have either $Q \lhd H$ or $P \leqq H$. Thus since we are assuming $PQ \neq QP$, so that we cannot have both $P \leqq H$ and $Q \leqq H$, the lemma is proved.

3. Information about P and Q

Lemma 8 will enable us to say that either a large subgroup of P normalizes Q or a large subgroup of Q normalizes P. If both p and q are odd, this will lead to a contradiction as in [2] – the argument is contained essentially in Lemma 9 below – but we are also able to obtain some information if pq is even.

Lemma 9. *Interchanging P and Q if necessary, for some k we have:*

(i) $P_k \leqq N(Q)$ *and for* $j \neq k$ $P_j \nleqq N(Q)$ *and*

(ii) *for* $j \neq k$ $Q_j \leqq N(P)$.

Proof. Suppose false and, without loss of generality, that p is odd. Fix an integer j and let H be a maximal V-invariant $\{p, q\}$-subgroup of G containing $P_j Q_j$ ($P_j Q_j$ is a group since G_j is nilpotent). By Lemma 8 either $P \lhd H$ or $Q \lhd H$. Thus we have that for each j either $P_j \leqq N(Q)$ or $Q_j \leqq N(P)$. If $Q_j \leqq N(P)$ for all j then $QP = PQ$ which is not the case. Thus, since we are assuming the lemma to be false, we have $P_j \leqq N(Q)$ for at least two values of j, say without loss of generality that $P_0 P_1 \leqq N(Q)$.

Let $K = N_P(Q)$ and $S = N_Q(P)$, so that $[K, S] = 1$. $P_0 P_1 \leqq K$ so $K \neq 1$ by Lemma 3 and the fact that p is odd. Let $\mathscr{H}$ be the set of groups R such that:

(i) R is a non-1 V-invariant subgroup of P,

(ii) $P_0 P_1 \leqq R$,

(iii) $[R, S] = 1$,

so that $K \in \mathscr{H}$. Let $R \in \mathscr{H}$. Then (i) and (ii) are satisfied with R replaced by $N_P(R)$.

Let L be the V-invariant Hall $\{p, q\}$-subgroup of $N(R)$, and $H = L \cap C(R)$, so that H is the V-invariant Hall $\{p, q\}$-subgroup of $C(R)$. $P_i \cap C(R) \leqq Z(R)$ for $i = 0, 1$ since $P_0 P_1 \leqq R$, so that by Lemma 3 $H/Z(R)$ has a normal p-complement. The inverse image $\hat{Q}$ in H of this complement contains S, and so $\hat{Q} = \hat{Q} \times Z(R)$ with $S \leqq \hat{Q}$ and $\hat{Q} \lhd L$. Then $[N_P(R), S] \leqq P \cap \hat{Q} = 1$. Thus $N_P(R) \in \mathscr{H}$. $\mathscr{H}$ is non-empty, so $P \in \mathscr{H}$.

Suppose $S \neq 1$. $[P, S] = 1$ by the preceding paragraph. Let H be the V-invariant Hall $\{p, q\}$-subgroup of $N(S)$ so that $P \leqq H$. By Lemma 8 $P \lhd H$, so $N_Q(S)$ normalizes P. Thus $S = Q$ contrary to the assumption $PQ \neq QP$.

Thus $S = 1$. If S does not contain Q_j for any j, then $P_j \leqq K$ for all j, whence P normalizes Q, again a contradiction. If S does not contain Q_j for two values of j, then the lemma is true with P and Q interchanged, so $Q_j = 1$ for two values of j. Then from Lemma 2 with Q in the role of H, K centralizes Q. Now $K \neq 1$, as noted above, so let H be the V-invariant Hall $\{p, q\}$-subgroup of $N(K)$; then $Q \lhd H$ by Lemma 8. But then $N_P(K)$ normalizes Q, contradicting $PQ \neq QP$ again. The lemma is proved.

We suppose without loss of generality that p, q are chosen so that (i) and (ii) of Lemma 9 hold as written, and that $k = 0$.

Lemma 10. $Q = Q_0$.

Proof. First of all, suppose $P_0 = 1$. Then for $j \neq 0$, Q_j centralizes P by Lemmas 1 and 9(ii). But then P normalizes $Q^* = \langle Q_j \colon j \neq 0 \rangle$. If $Q^* \neq 1$, then $PQ = QP$ is exhibited in $N(Q^*)$ by Lemma 5. Thus $Q^* = 1$, so $Q = Q_0$.

We now prove that $P_0 = 1$, so suppose false. Let H be a maximal V-invariant $\{p, q\}$-subgroup of G containing QP_0. P_0 normalizes Q so that if $j \neq 0$ $[P_0, Q_j] \leqq Q \cap P = 1$. $[P_0, Q_0] = 1$ since G_0 is nilpotent, so $[P_0, Q] = 1$, whence by Lemma 4 applied to H $1 \neq P_0 \leqq O_p(H)$. $O_q(H) \neq 1$ since $Q \lhd H$ by Lemma 8. We are in a position to apply Lemma 7 to H.

If $Z(P) \cap O_p(H) \neq 1$ then $P \leqq H$ by Lemma 7, which is not so. Thus $Z(P)_0 = 1$. For $j \neq 0$ Q_j normalizes P, so for $j \neq 0$ $[Q_j, Z(P)] = 1$ by Lemma 1 applied to $Z(P)$ in the role of H.

Again by Lemma 7 $Z(P) \leqq H$, so $Z(P)$ normalizes Q. $Z(P) \cap F(H) = 1$ so by Lemma 4 (with p and q interchanged) $C_{Z(P)}(Q) = 1$. Let $Q^* = \langle Q_j \colon j \neq 0 \rangle$; $Q^* \lhd Q$ by Lemma 5. If $j \neq 0$, α_j acts without fixed points on Q/Q^*. $Z(P)$ normalizes Q and centralizes Q^*, and α_j acts without fixed points on $[Z(P), \alpha_j]$, so $[Z(P), \alpha_j]$ centralizes Q/Q^* by Thompson's theorem. Thus since $[Z(P), \alpha_j]$ centralizes Q^*, $[Z(P), \alpha_j] \leqq C_{Z(P)}(Q) = 1$. But this is true for any $j \neq 0$, whence $Z(P) = 1$ since V acts fixed-point-freely on G. This contradiction completes the proof of the lemma.

4. A 2-Complement

It is now not difficult to obtain a contradiction from Lemma 10.

Lemma 11. *G has a V-invariant soluble 2-complement L. If T is the V-invariant Sylow 2-subgroup of G, then $T = T_i$ for some i and $G = LT_i$.*

Proof. Suppose P and Q are V-invariant Sylow p- and q-subgroups of G and $PQ \neq QP$. By Lemma 10 we may choose the notation so that $Q = Q_0$. Since G clearly does not have a normal q-complement, Lemma 3 implies that $q = 2$. It follows that any two odd-order V-invariant Sylow subgroups of G are permutable, so G has a V-invariant 2-complement L. By Hall's characterization of soluble groups (see, for example, [1], Theorem 6.4.5) L is soluble. Since G is not soluble, there is an odd prime p such that the V-invariant Sylow p-subgroup of G is not permutable with the V-invariant Sylow 2-subgroup T of G. Thus $T = T_i$ for some i.

$$|L| = |G|_{2'}, \quad \text{so} \quad |LT_i| = |L| \cdot |T_i| / |L \cap T_i| = |G|_{2'} \cdot |G|_2 = |G|,$$

and the lemma is proved.

We may now complete the

Proof (of the theorem). $[G, \alpha_i]$ is a V-invariant normal subgroup of G. Clearly $[G, \alpha_i] \neq 1$, for if not, then some α_j would act without fixed points on G, whence G would be nilpotent by Thompson's theorem. But from Lemma 11 $[G, \alpha_i] = [LT_i, \alpha_i] = [L, \alpha_i] \leqq L$ and $L \neq G$ since L is soluble. Thus G has a proper non-1 V-invariant normal subgroup, contrary to the minimality of G.

References

1. Gorenstein, D.: Finite groups. New York: Harper and Row 1968.
2. Martineau, R. P.: On groups admitting a fixed-point-free automorphism group. Proc. London Math. Soc. (3) **22**, 193–202 (1971).
3. Thompson, J. G.: Normal p-complements for finite groups. J. Algebra **1**, 43–46 (1964).
4. Ward, J. N.: Automorphisms of finite groups and their fixed point groups. J. Austral. Math. Soc. **9**, 467–477 (1969).

Dr. R. P. Martineau
University of Oxford
Mathematical Institute
24–29 St Giles
Oxford OX1 3LB
England

(Received May 7, 1971)

Math. Z. 124, 73–78 (1972)

Groups All of Whose Subgroups Are M-Groups

DAVID L. WINTER and PAUL F. MURPHY

1. Introduction

A finite group is called an M-group if and only if each of its irreducible representations over the complex number field is equivalent to a monomial representation. An outstanding problem has been to find a group-theoretical property which characterizes M-groups. It is known that an M-group is solvable and certain sufficient conditions have been given. Some of the difficulty involved in solving the problem is indicated by a result of Dade which states that any finite solvable group can be embedded in an M-group ([6], p. 585).

In this paper a simpler problem is studied: that of finding a group-theoretical characterization of $\tilde{M}$-groups where by an $\tilde{M}$-group we mean a group all of whose subgroups are M-groups. A sufficient condition is obtained for $\tilde{M}$-groups. Furthermore it is shown that, at least within a certain class of solvable groups, this condition is necessary. The condition includes the known group-theoretical sufficient conditions for a group to be an M-group.

All groups considered here are of finite order. All group representations are over the field of complex numbers. A section of a group is a homomorphic image of a subgroup.

2. On E_0-Groups

We shall say that P is an *S-group* if P is a non-abelian p-group for some prime p, the center $Z(P)$ is cyclic and $P/Z(P)$ is elementary abelian. If the p-group P is an S-group, it can be shown that $|P'|=p$, $|P:Z(P)|=p^{2n}$ for some integer n and each non-linear, irreducible, complex character of P is faithful of degree p^n ([6], pp. 353–355 and [8], Theorem 5).

Lemma 1. *Let the p-group P be a normal S-subgroup of the group G such that G acts trivially on $Z(P)$ and irreducibly on $P/Z(P)$. Then*

(1) *if p is odd, P is the central product of $Z(P)$ and an extraspecial p-group of exponent p;*

(2) *if $p=2$ and $|Z(P)|>2$, then P is the central product of $Z(P)$ with a number of dihedral groups of order* 8;

(3) *if $p=2$ and $|Z(P)|=2$, then P is the central product of a number of dihedral groups of order* 8 *or the central product of a number of dihedral groups of order* 8 *and a quaternion group. However, P itself is not a dihedral group of order* 8.

In particular if $p=2$, P has a quaternion section.

Proof. If A is an abelian characteristic subgroup of P, then $A \leqq Z(P)$ and $Z(P)$ is cyclic by assumption. The proofs of ([8], Theorems 3 and 4) may now be used to complete the proof of Lemma 1.

By an *E-group* we shall mean a solvable group G which has a normal S-subgroup P such that G acts trivially on $Z(P)$ and irreducibly on $P/Z(P)$.

An E_0*-group* is a solvable group no section of which is an E-group. Clearly a subgroup of an E_0-group is an E_0-group.

Theorem 1. *The direct product of two E_0-groups is an E_0-group. Consequently, the class of E_0-groups is a formation.*

Proof. Let $W = W_1 \times W_2$ be a counterexample of minimal order where W_i is an E_0-group for $i = 1, 2$. Then W contains a subgroup H which in turn has a normal subgroup V such that H/V is an E-group. Observing $H \leqq HW_1 = W_1 \times ((HW_1) \cap W_2)$, we can deduce

(1) $HW_1 = W = HW_2$.

(2) $W_i \cap V = \langle 1 \rangle$ *for* $i = 1, 2$.

Suppose $W_1 \cap V \neq \langle 1 \rangle$. Since W_2 centralizes W_1, $W_1 \cap V \lhd HW_2 = W$. H/V is a factor group of

$$H/W_1 \cap V \leqq (W_1 \times W_2)/(W_1 \cap V) \cong (W_1/W_1 \cap V) \times W_2$$

and the induction hypothesis yields a contradiction.

By assumption H contains a normal subgroup $P > V$ such that P/V is a p-group, in fact an S-group, and H/V acts irreducibly on $(P/V)/(R/V)$ and trivially on the cyclic group $R/V = Z(P/V)$. We now show

(3) $P \cap W_i \leqq R$ *for* $i = 1, 2$.

It suffices to show $P \cap W_i$ is abelian for then $(P \cap W_i)V \neq P$ and $(P \cap W_i)V \lhd H$. Suppose, for example, $P \cap W_1$ is non-abelian. $P \cap W_1 \lhd HW_2 = W$ and therefore $P \cap W_1 \lhd W_1$. $P \cap W_1 \cong (P \cap W_1)V/V$ by (2) and therefore $P \cap W_1$ is a p-group. H acts trivially on $(R \cap W_1)V/V$. Since $R \cap W_1 \lhd H$, (2) implies H acts trivially on the cyclic group $R \cap W_1$. By the irreducibility of the action of H/V on P/R, $P = (P \cap W_1)R$. $P/R \cong P \cap W_1/P \cap W_1 \cap R = P \cap W_1/R \cap W_1$ and we see that H acts irreducibly on $P \cap W_1/R \cap W_1$. Since $P \cap W_1 \lhd W = W_1 \times W_2$, W_1 acts irreducibly on $P \cap W_1/R \cap W_1$. We now show $R \cap W_1 = Z(P \cap W_1)$. Since W_1 centralizes $R \cap W_1$, $R \cap W_1 \leqq Z(P \cap W_1)$. $Z(P \cap W_1)V \lhd H$ and therefore $Z(P \cap W_1)V \leqq R$ and $Z(P \cap W_1) \leqq R \cap W_1$. $P \cap W_1$ is therefore an S-group and W_1 is an E-group. This proves (3).

If T is a subgroup of $W = W_1 \times W_2$, we shall write

$$T_1 = \{w_1 \in W_1 \mid \text{there exists } w_2 \in W_2 \text{ with } w_1 w_2 \in T\}$$

and

$$T_2 = \{w_2 \in W_2 \mid \text{there exists } w_1 \in W_1 \text{ with } w_1 w_2 \in T\}.$$

It is immediate that T_i is a subgroup of W_i. Since $H \leqq H_1 \times H_2$, it follows that $H_1 = W_1$ and $H_2 = W_2$. Since $V \lhd H$, $V_i \lhd W_i$ for $i = 1, 2$.

Now let $w_1 w_2, w_1' w_2' \in P$ where $w_1, w_1' \in P_1$ and $w_2, w_2' \in P_2$. Then

$$[w_1 w_2, w_1' w_2'] = [w_1, w_1'][w_2, w_2'].$$

Thus P non-abelian implies P_1 or P_2 is non-abelian, say P_1.

Let $w_1 \in P_1$. Then $w_1 w_2 \in P$ for some $w_2 \in W_2$. For some s, $w_1^{p^s} w_2^{p^s} = (w_1 w_2)^{p^s} \in V$ where $w_1^{p^s} \in V_1$. Therefore P_1/V_1 is a p-group.

Let $V w_1 w_2$ be a generator of the cyclic group R/V. Then it is easily seen that R_1/V_1 is a cyclic group with generator $V_1 w_1$. We shall show $R_1/V_1 = Z(P_1/V_1)$. Clearly, $R_1/V_1 \leqq Z(P_1/V_1)$. Let $V_1 y_1 \in Z(P_1/V_1)$. Then there exists $y_2 \in P_2$ such that $y_1 y_2 \in P$. Let $x = x_1 x_2 \in P$ with $x_i \in W_i$. Then $(y_1 y_2)^x = y_1^{x_1} y_2^{x_2} = y_1 v_1 y_2^{x_2}$ for some $v_1 \in V_1$. There exists $v_2 \in V_2$ such that $v_1 v_2 \in V$. Thus

$$(v_1 v_2)^{-1} (y_1 y_2)^{-1} y_1 v_1 y_2^{x_2} = v_2^{-1} y_2^{-1} y_2^{x_2} \in P \cap W_2.$$

By (3) $P \cap W_2 \leqq R$ and therefore $v_2^{-1} y_2^{-1} y_2^{x_2} V \in R/V$. Since $(P/V)'$ is the unique subgroup of R/V of order p, $v_2^{-1} y_2^{-1} y_2^{x_2} V \neq V$ implies $(P/V)' \leqq W_2 V/V$ and a short computation yields P_1/V_1 is abelian. Therefore $v_2^{-1} y_2^{-1} y_2^{x_2} \in V \cap W_2$. By (2) $y_2^{x_2} = y_2 v_2$. Hence $(y_1 y_2)^x = y_1^{x_1} y_2^{x_2} = y_1 v_1 y_2 v_2 = (y_1 y_2)(v_1 v_2)$. It follows that $y_1 y_2 \in R$. Therefore $y_1 \in R_1$ and $V_1 y_1 \in R_1/V_1$ and $Z(P_1/V_1) \leqq R_1/V_1$ as desired.

$(P_1/V_1)/(R_1/V_1) \cong P_1/R_1$ is easily verified to be elementary abelian whence P_1/V_1 is an S-group. The irreducibility of the action of H/V on $(P/V)/(R/V)$ is easily seen to imply that W_1/V_1 acts irreducibly on $(P_1/V_1)/(R_1/V_1)$. Therefore W_1/V_1 is an E-group. This contradiction completes the proof.

3. On $\tilde{M}$-Groups

We now present the main results of this paper.

Theorem 2. *If G is an E_0-group, then G is an $\tilde{M}$-group.*

Proof. Let G be a counterexample of minimal order. Then every section of G except G itself is an M-group. Let χ be an irreducible complex character of G which is not induced from a linear character of a subgroup of G. Suppose χ is not primitive. Then $\chi = \psi^*$ where ψ is an irreducible character of some proper subgroup B of G ([1], (50.2) and (50.6)). Since B is an M-group, transitivity of induction yields that $\chi = \lambda^*$ for some linear character λ of some subgroup of B. Therefore χ is primitive and minimality of $|G|$ implies χ is faithful.

$Z(G)$ consequently is cyclic and $Z(G) \neq \langle 1 \rangle$ because each abelian normal subgroup of G is contained in $Z(G)$. Let $M/Z(G)$ be a minimal normal subgroup of $G/Z(G)$. Then $M/Z(G)$ is elementary abelian of order p^r for some prime p and G acts irreducibly on $M/Z(G)$. Since $Z(G) < M$, M is non-abelian. Let P be a Sylow p-group of M and let Z_1 be the p-complement of $Z(G)$. Clearly $P \cap Z(G) \leqq Z(P)$. Since $M = P \times Z_1 \lhd G$, $P \lhd G$ and $Z(P) \lhd G$. Therefore $Z(P) \leqq Z(G)$, $Z(P)$ being an abelian normal subgroup of G. Thus we have $Z(P) = P \cap Z(G)$. $P/Z(P) = P/P \cap Z(G) \cong PZ(G)/Z(G) = M/Z(G)$ is elementary abelian. Therefore P is a normal S-subgroup of G. It is clear that G acts trivially on $Z(P)$ and irreducibly on $P/Z(P)$ and this proves Theorem 2.

Corollary (Dornhoff, [2]). *Let G be a solvable group and let $N \lhd G$. Suppose G/N is supersolvable and the Sylow p-subgroups of N are modular for p odd and quaternion-free for $p=2$. Then G is an $\tilde{M}$-group.*

Proof. It suffices to show that G cannot be an E-group as the same proof may be applied to each section of G. Suppose G is an E-group and P a normal p-subgroup satisfying the appropriate conditions. Suppose first that $P \cap N$ is non-abelian. Then $P=(P \cap N)\,Z(P)$. Clearly $Z(P) \cap N \leqq Z(P \cap N)$. Since $Z(P \cap N) \lhd G$, $Z(P \cap N) \leqq Z(P)$. Therefore $Z(P) \cap N = Z(P \cap N)$ whence $Z(P \cap N)$ is cyclic and $P/Z(P)=(P \cap N)\,Z(P)/Z(P) \cong P \cap N/Z(P) \cap N = P \cap N/Z(P \cap N)$. Thus $P \cap N$ is an S-group and G must act irreducibly on $P \cap N/Z(P \cap N)$. By Lemma 1, $P \cap N$ has a non-abelian section of order p^3 which is of exponent p if p is odd or is a quaternion group if $p=2$. This is in contradiction to the assumption on the p-Sylow subgroup of N ([2], Lemma 1.11). If $P \cap N$ is abelian, then $P \cap N \leqq Z(P)$. Therefore N acts trivially on $P/Z(P)$ and G/N acts irreducibly on $P/Z(P)$. Hence $(PN/N)/(Z(P)\,N/N) \cong PN/Z(P)\,N \cong P/Z(P)$ is a non-cyclic chief factor of G/N, contradicting the supersolvability of G/N.

Lemma 2. *Let G be an E-group with normal p-subgroup P as in the definition of E-group. If P is a p-Sylow subgroup of G, then G is not an M-group.*

Proof. Let $|P:Z(P)|=p^{2n}$ and let φ be a complex irreducible non-linear character of P. Then φ has degree p^n. Since $\sum_P |\varphi(x)|^2=|P|$, it is easily seen that $\varphi(x)=0$ for $x \in P-Z(P)$. It follows that φ is invariant under G. Hence φ has an extension to an irreducible character χ of G([6], p. 572). Suppose G is an M-group. Then G has a subgroup H of index p^n which has a linear character λ such that $\lambda^*=\chi$. Since $Z(P) \leqq Z(G)$, $Z(P) \leqq H$. Therefore $P \cap H \lhd P$ and since $P \cap H \lhd H$, $P \cap H \lhd G=PH$. Because $|P:P \cap H|=|PH:H|=p^n$, this contradicts the irreducibility of the action of G on $P/Z(P)$. Therefore G is not an M-group.

Theorem 3. *Let G be a solvable group which has a normal subgroup A all of whose Sylow subgroups are abelian and such that G/A has Fitting height at most 2. Then G is an $\tilde{M}$-group if and only if G is an E_0-group.*

Proof. In view of Theorem 2 we need only show that if G is an $\tilde{M}$-group, then G is an E_0-group. Since the conditions are inherited by all sections of G, it suffices to show that if G is an E-group, then G is not an $\tilde{M}$-group.

To this end suppose G has a normal non-abelian p-subgroup P satisfying the conditions necessary to make G an E-group. Then $P \cap A$ is a normal abelian subgroup of G and so $P \cap A \leqq Z(P)$. Therefore A acts trivially, and G/A irreducibly on $P/Z(P)$. Let F/A be the Fitting subgroup of G/A. By assumption G/F is nilpotent. Let R and T be such that R/A is the p-Sylow subgroup of F/A and T/A is the normal p-complement. Then $F/A=(R/A)\times(T/A)$ and $P \leqq R$. Hence T acts trivially on $P/Z(P)$. Since R/A is a p-group the subspace V of $P/Z(P)$ of fixed points for R/A is not $\{0\}$. Since $R/A \lhd G/A$, V is invariant under G/A. Hence $V=P/Z(P)$ and R acts trivially on $P/Z(P)$. Therefore F acts trivially on $P/Z(P)$ and G/F irreducibly. Since G/F is nilpotent it follows that its p-Sylow subgroup is trivial on $P/Z(P)$ while its normal p-complement acts irreducibly. Let H be a Hall p'-subgroup of G. It now follows that HP is an E-group. By Lemma 2 HP is not an M-group. Therefore G is not an $\tilde{M}$-group.

4. Applications and Remarks

(a) The proof of Theorem 2 shows that if G is a solvable complex primitive linear group of degree d, then G is an E-group. If P is the corresponding normal p-subgroup, then $G/C(P/Z(P))$ acts trivially on $Z(P)$ and faithfully and irreducibly on $P/Z(P)$. It is known then that $G/C(P/Z(P))$ is isomorphic to an irreducible subgroup of 1) the symplectic group $\mathrm{Sp}(2n, p)$ if p is odd or 2) one of the orthogonal groups $O_\varepsilon(2n, 2)$, $\varepsilon = \pm 1$, if $p=2$. Here n is the positive integer such that $p^{2n} = |P:Z(P)|$. By Clifford's theorem d is a multiple of p^n. This provides a new way of approaching certain questions on primitive solvable linear groups.

(b) We know of no E-group which is an $\tilde{M}$-group but it is possible for an E-group to be an M-group. Let $Q \cdot Q_1$ be the central product of two quaternion groups Q and Q_1. Let α be an automorphism of $Q \cdot Q_1$ of order 3 such that $Q^\alpha = Q$, $Q_1^\alpha = Q_1$ and α is non-trivial on both Q and Q_1. Then $G = \langle\alpha\rangle\, Q \cdot Q_1$ is an M-group ([2], p. 249). The presence of $Q \lhd G$ with α acting irreducibly on $Q/Z(Q)$ implies G is an E-group.

(c) The following terminology is convenient for some applications.

Definition. A *D-group* is a group $\langle y\rangle\, P$ which is a semidirect product of the normal p-subgroup P which is an S-group and a cyclic group $\langle y\rangle$ of order a power of a prime $q \neq p$ where y acts trivially on $Z(P)$ and non-trivially on $P/Z(P)$. A *D_0-group* is a solvable group no section of which is a D-group.

If G is a D_0-group, then G is an $\tilde{M}$-group.

Proof. If not, G has a section K which is an E-group. Thus for some prime p, K contains a normal p-subgroup P which is an S-group. $K/C_K(P/Z(P))$ cannot be a p-group because a non-trivial p-group can have no faithful irreducible representation over a field of characteristic p ([3], p. 62). Therefore there is a q-Sylow subgroup Q of K with $Q \nsubseteq C_K(P/Z(P))$ for some prime $q \neq p$. Let $y \in Q - C_K(P/Z(P))$. Then $\langle y\rangle\, P$ is a D-group and is a section of G. This contradiction completes the proof.

(d) We apply (c) to a generalization of an example of Seitz ([9], Example 2.8). Let N be a Suzuki 2-group [4] of order 2^{2n} where $2^n - 1$ is a Mersenne prime larger than 3. N has an automorphism α of order $2^n - 1$ which permutes its $2^n - 1$ involutions cyclically. $Z(N) = N' = \Omega_1(N) = \Phi(N)$, $|Z(N)| = 2^n$ and N has exponent 4. It follows that α acts fixed-point-free on N. Let $G = \langle\alpha\rangle\, N$ be the relative holomorph. Suppose G has a section H/K which is a D-group. Replacing H by a suitable conjugate we can assume $\alpha \in H$. Then α acts fixed-point-free on $H \cap N$ and on $H \cap N/K$. Hence α cannot act trivially on the center of $H \cap N/K$, a contradiction. Therefore G is a D_0-group and therefore an $\tilde{M}$-group. On the other hand the corollary cannot be applied to G. For, N is the minimal normal subgroup of G such that G/N is supersolvable and by a theorem of Dornhoff ([2], Theorem 2.7) N has a quaternion section.

(e) The work of Seitz and Wright [10] shows that under the assumptions of the corollary of the preceding section the supersolvable residual of G is an A-group, i.e., a group whose p-Sylow subgroups are abelian for all primes p. Huppert [5] proved that if N is a normal subgroup of the solvable group G such that N is an A-group and G/N is supersolvable, then G is an M-group. Thus the corollary is also a consequence of [10] and [5]. The proof in this paper would be shorter if N is assumed to be an A-group.

(f) It has been shown that the wreath product of an M-group with an $\tilde{M}$-group is an M-group ([9], Theorem 2.2 and [7]). Consequently, the wreath product of an M-group with an E_0-group is an M-group and the wreath product of an E_0-group with an E_0-group is an M-group.

References

1. Curtis, C.W., Reiner, I.: Representation theory of finite groups and associative algebras. New York: Interscience 1962.
2. Dornhoff, L.: M-groups and 2-groups. Math. Z. **100**, 226–256 (1967).
3. Gorenstein, D.: Finite groups. New York: Harper and Row 1968.
4. Higman, G.: Suzuki 2-groups. Illinois J. Math. **7**, 79–96 (1963).
5. Huppert, B.: Monomiale Darstellungen endlicher Gruppen. Nagoya Math. J. **6**, 93–94 (1953).
6. — Endliche Gruppen I. Berlin-Heidelberg-New York: Springer 1967.
7. Kerber, A.: Zur Theorie der M-Gruppen. Math. Z. **115**, 4–7 (1970).
8. Rigby, J.F.: Primitive linear groups containing a normal nilpotent subgroup larger than the center of the group. J. London Math. Soc. **35**, 389–400 (1960).
9. Seitz, G.M.: M-groups and the supersolvable residual. Math. Z. **110**, 101–122 (1969).
10. — Wright, C.R.B.: On finite groups whose Sylow subgroups are modular or quaternion-free. J. Algebra **13**, 374–381 (1969).

Dr. David L. Winter
Department of Mathematics
Michigan State University
East Lansing, Michigan 48823
USA

Dr. Paul F. Murphy
Department of Mathematics
California State Polytechnic College
San Luis Obispo, California 93401
USA

(Received June 1, 1971)

Math. Z. 124, 79–82 (1972)

Über die Länge der Knotenlinien schwingender Membranen

JOCHEN BRÜNING und DIETER GROMES

Es wird eine asymptotische (untere) Abschätzung für die Länge der Knotenlinien schwingender Membranen angegeben. Das Ergebnis ist ein Analogon zu bekannten Eigenschaften der Nullstellen von Eigenfunktionen Sturm-Liouvillescher Probleme.

Wir gehen aus von

Satz 1 ([1], S. 395). *Es bezeichne y_n die n-te Eigenfunktion eines Sturm-Liouvilleschen Problems im Intervall $I=[a,b]$. Dann hat y_n in (a,b) genau $n-1$ (notwendig einfache) Nullstellen.*

Es stellt sich die Frage, ob diese Eigenschaft ein Analogon bei Eigenfunktionen elliptischer Randwertprobleme in höheren Dimensionen hat. Beschränken wir uns dabei auf den zweidimensionalen Fall. Es sei also G ein $s+1$-fach zusammenhängendes Gebiet im R^2 mit hinreichend glattem Rand S; es bezeichne ferner F die Fläche und U den Umfang von G. In G betrachten wir das Problem

$$\begin{aligned} \Delta u+\lambda u&=0 \quad \text{in } G,\\ u&=0 \quad \text{auf } S. \end{aligned} \tag{1}$$

Die Aussage von Satz 1, die sich ja nicht unmittelbar übertragen läßt, ist offenbar äquivalent zu der Behauptung, daß die Nullstellen von y_n I in genau n Teile zerlegen. In Analogie dazu können wir nach der Anzahl $k(n)$ der Gebiete fragen, in die G durch die Knotenlinien der n-ten Eigenfunktion u_n von (1) zerlegt wird. Courant bewies ([1], S. 393), daß $k(n)\leqq n$, und Pleijel [4] verschärfte dies zu

$$\limsup_{n\to\infty}\frac{k(n)}{n}\leqq\frac{4}{j^2}<0{,}7, \tag{2}$$

wo j die kleinste Nullstelle der nullten Besselfunktion bezeichnet. Aus (2) ersieht man, daß $k(n)=n$ nur endlich oft auftreten kann; es gibt sogar Beispiele, wo unendlich oft der Fall $k(n)=2$ auftritt [5], so daß $k(n)$ in keiner befriedigenden Beziehung zu n steht. Bekannte Beispiele von Knotenlinien schwingender Membranen [1], [5] legen nun die Vermutung nahe, daß zwar nicht die Anzahl der Teilgebiete, wohl aber die Länge l_n der Knotenlinien mit n gegen unendlich strebt. Setzt man

$$L_n = l_n + \frac{U}{2}, \tag{3}$$

und erklärt man den Inhalt $N-1$-dimensionaler Punktmengen im R^N nach Minkowski ([2], S. 994), so hat (3) auch für $N=1$ einen Sinn, und Satz 1 liefert $L_n = n$. Die Größe L_n kann also als legitime Verallgemeinerung der Anzahl der Nullstellen in höheren Dimensionen angesehen werden. Wir wollen beweisen:

Satz 2. *Es gilt die Abschätzung*

$$L_n \geqq \frac{F\sqrt{\lambda_n}}{2j} - \pi(s-1)\frac{j}{2\sqrt{\lambda_n}}. \tag{4}$$

Mit der Weylschen Beziehung $\lambda_n \sim \frac{4\pi n}{F}$ ergibt sich daraus die

Folgerung. *Es ist*

$$\liminf_{n\to\infty} \frac{L_n^2}{Fn} \geqq \frac{\pi}{j^2}. \tag{5}$$

Die Länge der Knotenlinien wächst also mindestens wie $\sqrt{n}$. Zum Beweis von Satz 2 benötigen wir zwei Hilfssätze.

Hilfssatz 1. *G_i sei ein s_i+1-fach zusammenhängendes Teilgebiet von G, das nur von den Knotenlinien der Eigenfunktion u_n zum Eigenwert λ_n von (1) und eventuell von S berandet wird, aber keine Knotenlinien im Innern enthält. Bezeichnet F_i die Fläche, U_i den Umfang und r_i den Inkreisradius von G_i, so gilt:*

$$r_i \leqq \frac{j}{\sqrt{\lambda_n}} \leqq \frac{U_i}{2\pi}. \tag{6}$$

Beweis. Die linke Seite der Ungleichung folgt aus der Beziehung $k(1)=1$ und der Monotonieeigenschaft der Eigenwerte ([1], S. 392), die rechte aus der Tatsache, daß unter allen eingespannten Membranen gleichen Umfangs die kreisförmige den tiefsten Grundton hat ([1], S. 402).

Hilfssatz 2. *Für ein $s+1$-fach zusammenhängendes Polygon mit Fläche F, Umfang U und Inkreisradius r gilt die Beziehung*

$$F \leqq Ur + (s-1)\pi r^2. \tag{7}$$

Beweis. Für $s=0$ hat Fejes-Tóth in [3] einen Beweis angegeben, der sich beinahe wörtlich auf den Fall $s>0$ übertragen läßt. Sein Hilfssatz 2 ändert sich nur insofern, als sich die Formel

$$f_1 + 2f_2 + 3f_3 + \cdots = F + L\rho - (s-1)\pi\rho^2$$

anstelle der ursprünglichen ergibt ($s=0$). Beachtet man nämlich, daß man zur Zerlegung eines $s+1$-fach zusammenhängenden Polygons in $k+1$ konvexe

Teilpolygone genau $k+s$ Diagonalen benötigt und bildet man ein Papiermodell wie in [3], so ergibt sich (mit den dortigen Bezeichnungen) für den Gesamtinhalt H der Papierblätter

$$\begin{aligned} H &= \sum_{i=1}^{k+1} (F_i + L_i \rho + \pi \rho^2) - \sum_{i=1}^{k+s} (2 d_i \rho + \pi \rho^2) \\ &= F + L\rho - (s-1)\pi \rho^2. \end{aligned}$$

Von hier ab verläuft der Beweis von Hilfssatz 2 wie bei Fejes-Tóth.

Beweis von Satz 2. Wir wenden Hilfssatz 2 auf jedes der durch die Knotenlinien erzeugten Teilgebiete G_i an, $1 \leqq i \leqq k(n)$. Das ist möglich, da die Knotenlinien und S stückweise glatt sind, also G_i durch Polygone approximiert werden kann. Da die rechte Seite von (7) für $s=0$ und $r \leqq U/2\pi$ in r monoton ist, können wir wegen (6) r_i durch $j/\sqrt{\lambda_n}$ ersetzen und erhalten:

$$F_i \leqq U_i \frac{j}{\sqrt{\lambda_n}} + (s_i - 1) \frac{\pi j^2}{\lambda_n}.$$

Summation über i ergibt

$$\begin{aligned} F &\leqq \frac{j}{\sqrt{\lambda_n}} \sum_{i=1}^{k(n)} U_i + \frac{\pi j^2}{\lambda_n} \sum_{i=1}^{k(n)} (s_i - 1) \\ &\leqq \frac{2j}{\sqrt{\lambda_n}} L_n + (s-1)\pi \frac{j^2}{\lambda_n}, \end{aligned}$$

da offenbar $\sum_{i=1}^{k(n)} U_i = 2L_n$ und $\sum_{i=1}^{k(n)} (s_i - 1) \leqq s-1$.

Tatsächlich ist die in der zweiten Ungleichung links stehende Summe nichts anderes als die Gesamtzahl aller inneren Ränder der G_i minus die Gesamtzahl aller äußeren Ränder der G_i. Nun bringt ein innerer Rand aber offenbar genau dann einen positiven Beitrag zur Summe, wenn er schon innerer Rand von G selbst ist; ein äußerer Rand bringt genau dann einen negativen Beitrag zur Summe, wenn er nicht zugleich auch innerer Rand eines der G_i ist, was zumindest für S zutrifft. Also muß die obige Ungleichung gelten. Das bedeutet aber

$$L_n \geqq \frac{F\sqrt{\lambda_n}}{2j} - (s-1)\pi \frac{j}{2\sqrt{\lambda_n}},$$

wie behauptet.

Offen bleibt die Frage, ob auch eine obere Abschätzung für L_n gilt. Für ein Rechteck mit den Seiten a, b und a^2/b^2 irrational (d.h. es gibt nur einfache Eigenwerte), gilt

$$\liminf_{n\to\infty} \frac{L_n^2}{Fn} = \frac{4}{\pi}, \qquad \limsup_{n\to\infty} \frac{L_n^2}{Fn} = \frac{8}{\pi}, \tag{8}$$

was eine solche Vermutung nahelegt. Allerdings ist uns darüber bis jetzt nichts bekannt.

Literatur

1. Courant, R., Hilbert, D.: Methoden der mathematischen Physik I, 2. Auflage. Berlin: Springer 1930.
2. Enzyklopädie der mathematischen Wissenschaften II, 3.2. Leipzig 1923 – 1927.
3. Fejes-Tóth, L.: Elementarer Beweis einer isoperimetrischen Ungleichung. Acta Math. Acad. Sci. Hungar. 1, 273 – 275 (1950).
4. Pleijel, A.: Remarks on Courant's nodal line theorem. Commun. Pure Appl. Math. 9, 543 – 550 (1956).
5. Stern, A.: Bemerkungen über asymptotisches Verhalten von Eigenwerten und Eigenfunktionen. Dissertation, Göttingen 1925.

J. Brüning
Mathematisches Institut
D-3550 Marburg
Universitätsstraße 24
Deutschland

Dr. D. Gromes
Institut für theoretische Physik
D-6900 Heidelberg
Philosophenweg 16
Deutschland

(Eingegangen am 16. Juli 1971)

Math. Z. 124, 83–88 (1972)

Extending a Norm from a Vector Lattice to its Dedekind Completion

ROSALIND REICHARD

Introduction

Let $(E, \leqq, \|\cdot\|)$ be a normed vector lattice. In [5], Nishiura questioned whether or not there exists more than one extension of $\|\cdot\|$ to $\hat{E}$, the Dedekind completion of E. Solovev answered this question in the negative in [6]. In fact, as pointed out by Vulich in [10], several Soviet mathematicians have results relating to this question and others posed by Nishiura.

The present paper is concerned with other problems related to extending a monotone norm from a vector lattice E to its Dedekind completion $\hat{E}$. It is known [6; Lemma 2] that if $(E, \leqq, \|\cdot\|)$ is a normed vector lattice and $\|\cdot\|$ is order-continuous then $\|\cdot\|$ must extend uniquely to $\hat{E}$. It will be shown that this is not true when $\|\cdot\|$ is σ-continuous or, surprisingly, even when $\|\cdot\|$ is semicontinuous. In fact an assertion of Solovev's [7; Theorem 3] along these lines is shown to be false. Although semicontinuous norms do not, in general, extend uniquely, it is shown that a large class of semicontinuous norms do have unique extensions. Finally, turning from norm properties to lattice properties, we note that whenever E has the projection property then every monotone norm on E must extend uniquely to $\hat{E}$.

1. Preliminaries

A *normed vector lattice* $(E, \leqq, \|\cdot\|)$ is a normed vector space E over the real number field with a partial order "$\leqq$" which makes E into a vector lattice and the norm $\|\cdot\|$ on E is such that $\|x\| \leqq \|y\|$ whenever $|x| \leqq |y|$ and x and y are in E. ($\|\cdot\|$ is said to be *monotone* on E.) An Archimedean vector lattice E always has a Dedekind completion $\hat{E}$ [9, p. 109]. An *extension* of $\|\cdot\|$ to $\hat{E}$ is a norm $\|\cdot\|'$ defined on $\hat{E}$ such that $\|x\| = \|x\|'$ for all x in E and $(\hat{E}, \leqq, \|\cdot\|')$ is a normed vector lattice. Given a normed vector lattice $(E, \leqq, \|\cdot\|)$ we can always extend $\|\cdot\|$ to $\hat{E}$ by defining $\|\hat{x}\|_M = \inf(\|w\|: w \geqq |\hat{x}|, w \in E)$ for all $\hat{x}$ in $\hat{E}$ [9, p. 179]. Also, $\|\cdot\|_M$ is the maximal extension of $\|\cdot\|$ to $\hat{E}$ in the sense that if $\|\cdot\|'$ is any other extension of $\|\cdot\|$ to $\hat{E}$ we must have $\|\hat{x}\|' \leqq \|\hat{x}\|_M$ for all $\hat{x}$ in $\hat{E}$. There are examples of normed vector lattices with two different norm extensions which are not even norm equivalent [6, p. 1362]. However, all examples of this, which were given previous to the present paper, are examples in which the original norm is not semicontinuous.

Let $(E, \leqq, \|\cdot\|)$ be a normed vector lattice. $\|\cdot\|$ is said to be *continuous* if $\inf(\|u_\alpha\|: \alpha\in\mathscr{A})=0$ whenever $0\leqq u_\alpha\downarrow_\alpha 0$ (i.e., $\inf(u_\alpha: \alpha\in\mathscr{A})=0$ in E and $\{u_\alpha\}$ is a net in E which is directed downward). $\|\cdot\|$ is said to be *semicontinuous* if $\sup(\|u_\alpha\|: \alpha\in\mathscr{A})=\|u\|$ whenever $0\leqq u_\alpha\uparrow u$ in E. The definitions for σ-continuous and σ-semicontinuous are the same with arbitrary nets replaced by sequences. Note that a continuous norm is always semicontinuous but $\|x\|=\sup(|x(n)|: n\in N)$ is a monotone semicontinuous norm on $E=l_\infty$ which is not continuous. (Let $x_n=\{x_n(i): i=1,2\ldots\}$ where $x_n(i)=0$ if $i\neq n$ and $x_n(n)=1$ for $n=1,2,\ldots$. Then $0\leqq x_n\downarrow_n 0$ in l_∞ but $\|x_n\|=1\to 1$ as $n\to\infty$.)

2. Continuous and σ-Continuous Norms

As has been mentioned, it is known that if $\|\cdot\|$ is a continuous norm on a normed vector lattice $(E, \leqq, \|\cdot\|)$ then $\|\cdot\|$ extends uniquely to $\hat{E}$. However, since this result depends on one lemma and two theorems in [6] and one theorem in [9] the following direct proof is presented here.

Theorem 2.1. *If $(E, \leqq, \|\cdot\|)$ is a normed vector lattice and $\|\cdot\|$ is continuous, then $\|\cdot\|$ extends uniquely to $\hat{E}$.*

Proof. Let $\phi(\hat{x})=\sup(\|y\|: 0\leqq y\leqq|\hat{x}|, y\in E)$ for all $\hat{x}$ in $\hat{E}$. If $\|\cdot\|_1$ is any extension of $\|\cdot\|$ to $\hat{E}$ we must have $\phi(\hat{x})\leqq\|\hat{x}\|_1\leqq\|\hat{x}\|_M$ for all $\hat{x}$ in $\hat{E}$. Hence, to show that $\|\cdot\|$ extends uniquely to $\hat{E}$ we need only prove that $\phi(\hat{x})=\|\hat{x}\|_M$ for all $\hat{x}$ in $\hat{E}$.

Choose $\hat{x}$ in $\hat{E}$ and let $A=\{w-y: w\geqq|\hat{x}|, 0\leqq y\leqq|\hat{x}|, w, y\in E\}$. A is a directed set in E so we can let $A=\{u_\alpha: \alpha\in\mathscr{A}\}$ where $\mathscr{A}$ is a directed set. Then $0\leqq u_\alpha\downarrow_\alpha 0$ in E since $\sup(y: 0\leqq y\leqq|\hat{x}|, y\in E)=\inf(w: w\geqq|\hat{x}|, w\in E)=|\hat{x}|$.

Using the hypothesis that $\|\cdot\|$ is continuous we must have $\inf(\|u_\alpha\|: \alpha\in\mathscr{A})=0$. Hence, given any $\varepsilon>0$ there exists an element $w\geqq|\hat{x}|$ and an element $0\leqq y\leqq|\hat{x}|$ with $w, y\in E$ such that $\|w-y\|<\varepsilon$. However, $\|w\|-\|y\|\leqq\|w-y\|$. Therefore, $\|w\|-\|y\|<\varepsilon$ and it follows that $\phi(\hat{x})=\|\hat{x}\|_M$. This proves that $\|\cdot\|$ extends uniquely to $\hat{E}$.

Since the property of σ-continuity is considerably weaker than the property of continuity of a norm, one might suspect that there exists a normed vector lattice $(E, \leqq, \|\cdot\|)$ with $\|\cdot\|$ σ-continuous and $\|\cdot\|$ having more than one extension to $\hat{E}$. The following example supports this suspicion.

Example. Let X be an uncountable set of points and let $E=\{f: f$ is a real-valued, bounded function on X and f is constant except possibly on a finite set of points in $X\}$. Then $\hat{E}=\{f: f$ is a real-valued, bounded function on $X\}$. For every f in E let $c(f)=$ that real number such that $\{t: f(t)=c(f)\}$ is uncountable. Also, let

$$\|f\|_1=\sup(|f(t)|: t\in X) \quad \text{and define} \quad \|f\|=\|f\|_1+2|c(f)|.$$

$(E, \leqq, \|\cdot\|)$ is a normed vector lattice and $\|\cdot\|$ is σ-continuous. The σ-continuity of $\|\cdot\|$ follows because $0\leqq f_n\downarrow 0$ in E implies the uniform convergence of the sequence $\{f_n\}$ to 0, and hence, $\|f_n\|\to 0$ as $n\to\infty$.

We now exhibit an extension of $\|\cdot\|$ which is distinct from $\|\cdot\|_M$. Define $\|\hat{x}\|_{ML}=\inf(\lim_{n\to\infty}\|\hat{x}_n\|_M\colon 0\leqq\hat{x}_n\uparrow|\hat{x}| \text{ in } \hat{E})$. Given any $\hat{x}$ in $\hat{E}$ $\|\hat{x}\|_{ML}$ is the infimum of the $\|\cdot\|_M$ norm limit of all sequences $\{\hat{x}_n\}$ in $\hat{E}$ such that $0\leqq\hat{x}_n\uparrow|\hat{x}|$. $\|\cdot\|_{ML}$ is called the Lorentz seminorm associated with $\|\cdot\|_M$ (see [2]). It is not difficult to show that $\|\cdot\|_{ML}$ agrees with $\|\cdot\|$ on E. It then follows that $\|\cdot\|_{ML}$ is a monotone norm extension of $\|\cdot\|$ to $\hat{E}$. Let $\{t_n\colon n\in N\}$ be any countable set of points in X. Define $f(t)=1$ if $t=t_n$ for $n=1,2,\ldots$, and $f(t)=0$ otherwise. Then $f\in\hat{E}$, $\|f\|_M=3$ but $\|f\|_{ML}\leqq 1$. This last fact follows because if $f_n(t)=1$ for $t=t_1,t_2,\ldots,t_n$ and $f_n(t)=0$ for all other t in X $(n=1,2,\ldots)$ then $0\leqq f_n\uparrow f$ and $\|f_n\|_M=1$.

3. Semicontinuous Norms

Although a semicontinuous norm need not be continuous, the two properties are similar enough so that one might conjecture that semicontinuous norms extend uniquely as do continuous ones. In the following we shall show that this conjecture is false.

If E is a vector lattice, and element e in E^+ ($E^+=$the positive elements in E) is said to be a *strong unit* for E if for every element x in E there exists a real number $\lambda>0$ such that $|x|\leqq\lambda e$. If an Archimedean vector lattice E has a strong unit e then e is also a strong unit for $\hat{E}$. If e is a strong unit for E then $\|x\|=\inf(\lambda\colon \lambda e\geqq|x|)$ defines a norm on E such that $(E,\leqq,\|\cdot\|)$ is a normed vector lattice [9, p. 180]. Also, the fact that $\|\cdot\|$ is semicontinuous follows directly from its definition.

It is known ([1, Theorem 1.5] and [7, Lemma 1]) that if $(E,\leqq,\|\cdot\|)$ is a normed vector lattice and $\|\cdot\|$ is semicontinuous then $\|\hat{x}\|_s=\sup(\|y\|\colon 0\leqq y\leqq|\hat{x}|, y\in E)$ is a semicontinuous norm extension of $\|\cdot\|$ to $\hat{E}$. As the example $E=c$ with $\|x\|=\sup(|x(n)|\colon n\in N)+\lim_{n\to\infty}|x(n)|$ demonstrates, this is not necessarily true in the case when $\|\cdot\|$ is not semicontinuous. In this case $\|\hat{x}\|_s=\sup(|\hat{x}(n)|\colon n\in N)+\lim_{n\to\infty}\inf|\hat{x}(n)|$ which is not a norm because it is not subadditive. It is also easy to see that if $(E,\leqq,\|\cdot\|)$ is a normed vector lattice and $\|\cdot\|_1$ and $\|\cdot\|_2$ are 2 semicontinuous norm extensions of $\|\cdot\|$ to $\hat{E}$ then $\|\hat{x}\|_1=\|\hat{x}\|_2$ for all $\hat{x}$ in $\hat{E}$.

Using the information presented above we shall now give an example of a normed vector lattice $(E,\leqq,\|\cdot\|)$ where $\|\cdot\|$ is semicontinuous but $\|\cdot\|$ has two different extensions to $\hat{E}$.

Example. Let $E=c$ and choose $e=(1,2,1,2\ldots)$ in l_∞. e is a strong unit for l_∞. Hence, for any $\hat{x}$ in $\hat{E}=l_\infty$ $\|\hat{x}\|_1=\inf(\lambda\colon \lambda e\geqq|\hat{x}|)$ defines a semicontinuous norm on $\hat{E}$. It follows that $\|\cdot\|=\|\cdot\|_1$ restricted to E is a semicontinuous norm on E. Also, $\|\hat{x}\|_s=\sup(\|y\|\colon 0\leqq y\leqq|\hat{x}|, y\in E)=\|\hat{x}\|_1$ since both $\|\cdot\|_s$ and $\|\cdot\|_1$ are semicontinuous extensions of $\|\cdot\|$ to $\hat{E}$. We shall now show that $\|\hat{x}\|_M=\inf(\|w\|\colon w\geqq|\hat{x}|, w\in E)\neq\|\hat{x}\|_s$ for some $\hat{x}$ in $\hat{E}$. Consider $e=(1,2,1,2\ldots)$. $\|e\|_s=1$. However, given any $w\in E=c$ such that $w\geqq e$ we must have $\lim_{n\to\infty}w(n)\geqq 2$. It follows that $\|w\|\geqq 2$ since $\|w\|=\inf(\lambda\colon \lambda e\geqq w)$. Therefore, $\|e\|_M=2\neq 1=\|e\|_s$. This proves that $\|\cdot\|_s$ and $\|\cdot\|_M$ are two distinct extensions of $\|\cdot\|$ to $\hat{E}$.

Remarks. 1. Solovev has asserted that if $(E, \leqq, \|\cdot\|)$ is a normed vector lattice satisfying:

$$\sup(\|u\|: u\in A)=\|x\| \quad \text{whenever} \quad \sup(u: u\in A)=x \text{ where } A\subset E^+ \qquad \text{(P)}$$

then $\|\cdot\|$ extends uniquely to $\hat{E}$ [7, Theorem 3]. Note that (P) is stronger than semicontinuity and if $(E, \leqq, \|\cdot\|)$ is a Banach lattice satisfying (P) then E is an AM-space. However, the previous example gives a norm satisfying (P) as well as being semicontinuous and yet $\|\cdot\|$ has two different extensions. Hence, the above example proves that Solovev's assertion was false.

2. In the previous example both $\|\cdot\|_s$ and $\|\cdot\|_M$ are equivalent. This is true because $\|\cdot\|_s$ is complete (any norm defined by means of a strong unit on a Dedekind complete vector lattice is complete [9, p. 181]) so that we can find a positive number k such that $\|\hat{x}\|_M\leqq k\|\hat{x}\|_s$ for all $\hat{x}\in\hat{E}$ [4, Theorem 30.28]. Now since $\|\hat{x}\|_s\leqq\|\hat{x}\|_M$ for all $\hat{x}$ in $\hat{E}$ it follows that $\|\cdot\|_s$ and $\|\cdot\|_M$ are equivalent.

If we let $E=c$ and $\hat{E}=l_\infty$ but we define $\|x\|=\inf(\lambda: \lambda e\geqq|x|)$ where $e=(1,2,1,3,1,4,\ldots,1,n,\ldots)$ for all x in E we have an example of a normed vector lattice $(E, \leqq, \|\cdot\|)$ such that $\|\cdot\|$ is semicontinuous and $\|\cdot\|_s$ and $\|\cdot\|_M$ are two extensions of $\|\cdot\|$ to $\hat{E}$ which are not equivalent. Let

$$e_1=(1,1,1,\ldots), \quad e_2=(1,2,1,2,1,2,\ldots),$$
$$e_3=(1,1,1,3,1,3,\ldots),\ldots, e_n=(1,1,\ldots,1,n,1,n,\ldots),\ldots.$$

Then $\|e_n\|_M=n$ and $\|e_n\|_s=1$ for $n=1,2,\ldots$. Hence, $\|e_n\|_M\to+\infty$ as $n\to+\infty$ but $\|e_n\|_s\to 1$ as $n\to+\infty$. This proves that $\|\cdot\|_M$ and $\|\cdot\|_s$ are not equivalent.

The following theorem will show that some classes of semicontinuous norms do extend uniquely.

A vector lattice E is *universally complete* if it is Dedekind complete and $\sup(x_\alpha: \alpha\in\mathscr{A})=x$ exists in E whenever $\{x_\alpha: \alpha\in\mathscr{A}\}\subseteq E$ and $\inf(x_{\alpha_1}, x_{\alpha_2})=0$ for all $\alpha_1,\alpha_2\in\mathscr{A}$ such that $\alpha_1\neq\alpha_2$. The vector lattice $E^\#$ is said to be a *universal completion* of E whenever $E^\#$ is universally complete and E can be isomorphically embedded as an order dense vector sublattice of $E^\#$. A vector lattice E has a universal completion if and only if E is Archimedean, and any two universal completions are isomorphic [3, Theorem 34.4].

In Theorem 39.1 of *Modern Spectral Theory* [3], Nakano characterizes a normed vector lattice $(E, \leqq, \|\cdot\|)$ having the property that whenever $\{x_\alpha: \alpha\in\mathscr{A}\}$ is a set of elements in E which is bounded above in E then $\sup(\|x_\alpha\|: \alpha\in\mathscr{A})=\inf(\|y\|: y$ is an upper bound for $\{x_\alpha\})$. Such a normed vector lattice must be a norm dense vector sublattice of a space $C(K)$ of continuous, bounded functions, vanishing at infinity, and defined on a locally compact, Hausdorff space K. The first part of the proof of this theorem can be used to prove:

If $(E, \leqq, \|\cdot\|)$ is a normed vector lattice satisfying (P) and $\{x_\alpha: \alpha\in\mathscr{A}\}\subseteq E^+$ such that $\|x_\alpha\|\leqq M$ for all $\alpha\in\mathscr{A}$ then $\sup(x_\alpha: \alpha\in\mathscr{A})=\bar{x}$ exists in $E^\#$.

As noted in the remark following the preceding example. (P) by itself is not enough to insure uniqueness of norm extensions. However, in the following theorem we shall prove that if $(E, \leqq, \|\cdot\|)$ is a normed vector lattice satisfying (P) and also E has the property that $\bar{e} \wedge w \in E$ for all $w \in E$ where $\bar{e} = \sup(x\colon x \in E^+, \|x\| \leqq 1)$ (which exists in $E^\#$ by the above discussion) then $\|\cdot\|$ extends uniquely to $\hat{E}$. This added condition on E is not at all unusual as any Archimedean vector lattice with a strong unit satisfies it.

Theorem 3.1. *Let* $(E, \leqq, \|\cdot\|)$ *be a normed vector lattice satisfying* (P). *In addition, letting* $\bar{e} = \sup(x \in E^+\colon \|x\| \leqq 1)$ *in* $E^\#$ *we shall assume that* $\bar{e} \wedge w \in E$ *whenever* $w \in E$. *Then* $\|\cdot\|$ *extends uniquely to* $\hat{E}$.

Proof. Extend $\|\cdot\|$ to $E^\#$ (allowing the extension to be infinite on some elements of $E^\#$) by defining $\|\bar{x}\|^* = \sup(\|y\|\colon 0 \leqq y \leqq |\bar{x}|, y \in E)$ for all $\bar{x}$ in $E^\#$. It is not difficult to show that $\|\cdot\|^*$ satisfies property (P) on $E^\#$. Let $\bar{E} = \{\bar{x}\colon \bar{x} \in E^\#$ and $\|\bar{x}\|^* < +\infty\}$. $(\bar{E}, \leqq, \|\cdot\|^*)$ is a normed vector lattice, $\hat{E} \subseteq \bar{E}$ and $\bar{e} \in \bar{E}$. This latter fact follows because if $A = \{x \in E^+\colon \|x\| \leqq 1\}$ then $\sup(x\colon x \in A) = \bar{e}$ which implies that $\|\bar{e}\|^* = \sup(\|x\|\colon x \in A) = 1$. Note also that $\bar{e} = \sup(\bar{x} \in \bar{E}^+\colon \|\bar{x}\|^* \leqq 1)$ implying that $|\bar{y}| \leqq \bar{e}$ if and only if $\|\bar{y}\|^* \leqq 1$ for all $\bar{y}$ in $E^\#$.

$\bar{E}$ has $\bar{e}$ as a strong unit and $\|\bar{x}\|^* = \inf(\lambda\colon \lambda\bar{e} \geqq |\bar{x}|)$ for all $\bar{x}$ in $\bar{E}$. To prove this statement we choose $\bar{x}$ in $\bar{E}$ and let $a = \|\bar{x}\|^*$. Then $\left\|\frac{1}{a}|\bar{x}|\right\|^* = 1$ which implies that $\frac{1}{a}|\bar{x}| \leqq \bar{e}$. Hence, $|\bar{x}| \leqq a\bar{e}$. Also, $\|\bar{x}\|^* = \inf(\lambda\colon \lambda\bar{e} \geqq |\bar{x}|)$ since $|\bar{x}| \leqq \|\bar{x}\|^* e$.

In order to show that $\|\cdot\|$ extends uniquely to $\hat{E}$ we need only show that for every $\hat{x}$ in $\hat{E}$ $\|\hat{x}\|^* = \inf(\lambda\colon \lambda\bar{e} \geqq |\hat{x}|) \geqq \inf(\|w\|\colon w \geqq |\hat{x}|, w \in E) = \|\hat{x}\|_M$. Choose any $\lambda > 0$ such that $\lambda\bar{e} \geqq |\hat{x}|$. By hypothesis, $w \wedge \bar{e} \in E$ whenever $w \in E$. Hence, if $w \geqq |\hat{x}|$ and $w \in E$ then $w' = w \wedge \lambda\bar{e} \in E$ and $w' \geqq |\hat{x}|$. Since $\|w'\| \leqq \|\lambda\bar{e}\|^* = \lambda$ it follows that for all λ such that $\lambda\bar{e} \geqq |\hat{x}|$ there exists an element $w' \geqq |\hat{x}|$ with $w' \in E$ such that $\|w'\| \leqq \lambda$. Hence,

$$\inf(\|w'\|\colon w' \geqq |\hat{x}|, w' \in E) \leqq \inf(\lambda\colon \lambda\bar{e} > |\hat{x}|).$$

This completes the proof of the theorem.

We remark here that in the example preceding this theorem $\bar{e} = (1, 2, 1, 2, \ldots)$ and although $\|\cdot\|$ satisfies property (P) we do not have $\bar{e} \wedge w \in E$ for all $w \in E$. In fact, $\bar{e} \wedge (2, 2, 2, \ldots) = \bar{e} \notin E$.

4. The Projection Property

Up to this point we have considered norm conditions which affect the uniqueness of norm extensions. We shall now consider a lattice condition on E.

A vector sublattice M in E is an *ideal* in E if $x \in M$ whenever $|x| \leqq y$ and $y \in M$. An ideal M in E is a *band* in E if $x \in M$ whenever $x = \sup(x_\alpha\colon \alpha \in \mathscr{A})$ and $\{x_\alpha\colon \alpha \in \mathscr{A}\} \subseteq M$. A band M in E is a *projection band* whenever given any x in E^+, $\sup(y\colon y \in M$ and $y \leqq x)$ exists in E. $\sup(y\colon y \in M$ and $y \leqq x)$ is an element of M and is called the *projection of x onto M*. A vector lattice E is said to have the *projection property* whenever every band in E is a projection band. The vector lattice of all bounded real sequences with finite range has the projection property but is not Dedekind complete.

Theorem 4.1. *If* $(E, \leqq, \|\cdot\|)$ *is a normed vector lattice and* E *has the projection property then* $\|\cdot\|$ *extends uniquely to* $\hat{E}$.

Proof. Veksler has shown [8, Lemma 3] that whenever E has the projection property and $\hat{x} \in \hat{E}^+$ there exists a monotone sequence $\{x_n\} \subseteq E^+$ such that $0 \leqq x_n \uparrow \hat{x}$ in $\hat{E}$ and there exists an element $\hat{z}$ in $\hat{E}^+$ and a sequence of positive real numbers $\lambda_n \downarrow 0$ such that $|x_n - \hat{x}| \leqq \lambda_n \hat{z}$ for $n = 1, 2, \ldots$.

Let $\|\cdot\|_1$ be any extension of $\|\cdot\|$ to $\hat{E}$ and choose $\hat{x}$ in $\hat{E}$. Using Veksler's result we know there exists a monotone sequence $\{x_n\} \subseteq E^+$ as described above. Hence, there exists an element $\hat{z} \in \hat{E}^+$ and a sequence $\lambda_n \downarrow 0$ such that $||\hat{x}| - x_n| \leqq \lambda_n \hat{z}$, $n = 1, 2, \ldots$. Then $\||\hat{x}| - x_n\|_1 \to 0$ as $n \to \infty$. Hence, $\|\hat{x}\|_1 - \|x_n\| \to 0$ as $n \to \infty$. But then $\sup(\|x_n\| : n \in N) = \|\hat{x}\|_1$ where $\|\cdot\|_1$ is any extension of $\|\cdot\|$ to $\hat{E}$. This implies that $\|\hat{x}\|_1 = \|\hat{x}\|_M$ for all $\hat{x}$ in $\hat{E}$ and $\|\cdot\|_M$ is the unique extension of $\|\cdot\|$ to $\hat{E}$.

References

1. Kawaii, I.: Locally convex lattices. J. Math.Soc. Japan **9**, 281–314 (1957).
2. Luxemburg, W.A.J., Zaanen, A.C.: Notes on Banach function spaces III. Kon. Nederl. Akad. Wetensch. Proc. Ser. A **66**, 239–250 (1963).
3. Nakano, H.: Modern spectral theory. Tokyo: Maruzen Co. 1950.
4. — Modulared semi-ordered linear spaces. Tokyo: Maruzen Co. 1950.
5. Nishiura, T.: Completions of normed linear lattices. Colloquium Math. **19**, 271–275 (1968).
6. Solovev, V.A.: On the extension of a monotone norm from a normed vector lattice to its Dedekind completion. Sibir. Mat. Žurn. **7**, 1360–1369 (1966).
7. — The extension of a semicontinuous norm and a monotonically complete norm from a KN-lineal to its K-completion. Vestnik Leningrad. Univ. **23**, 52–57 (1968).
8. Veksler, A.I.: Concerning a new construction of the Dedekind completion of some lattices and quotient spaces of l-groups. Sibir. Mat. Žurn. **10**, 1206–1213 (1969).
9. Vulich, B.Z.: Introduction to the theory of partially ordered spaces. Groningen: Wolters-Noordhoff 1967.
10. — Remarks on the completion of normed linear lattices. Colloquium Math. **21**, 101–102 (1970).

Prof. R. Reichard
Mathematics Department
Morehouse College
Atlanta, Georgia 30034
USA

(Received June 28, 1971)

Math. Z. 124, 89–99 (1972)

Almost Diagonal Matrices over Dedekind Domains

LAWRENCE S. LEVY*

Call a pair of matrices A and B over a ring R *equivalent* (notation: $A \sim B$) if $B = PAQ$ for square matrices P and Q which are invertible over R. Clearly every matrix A is equivalent to a matrix of the form

$$\begin{bmatrix} B_1 & 0 & & 0 & 0 \\ 0 & B_2 & & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & B_d & 0 \\ 0 & 0 & & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \quad \text{or} \quad [B_1 \quad 0]$$

where the B_i and zeros are matrices (some of the blocks of zeros possibly missing) and the B_i are *indecomposable* in the sense that they cannot be further reduced in this manner. (The diagonal placement of the B_i should be interpreted to mean that if the entry in the lower right-hand corner of B_i is in row j and column k, then the entry in the upper left-hand corner of B_{i+1} is in row $j+1$ and column $k+1$. We do *not* suppose that each B_i is square.)

How large can indecomposable matrices get? Can they be unboundedly large?

The two main results of this paper are that, for a Dedekind domain R: (1) The (ideal) class number, when it is finite, is a bound to the number of rows and columns in every indecomposable matrix over R; (2) A necessary and sufficient condition for the existence of a bound to the size of all indecomposable matrices over R is the existence of a positive integer e such that, for every ideal J of R, the ideal J^e is principal. (The bound is then $2e-1$.) These theorems are supplemented by examples of "large" indecomposable matrices.

Each of the results (1) and (2) generalizes the well-known fact that, over a principal ideal domain, every matrix is equivalent to a diagonal matrix.

Fixed Notation. The letter R will always denote a Dedekind domain.

1. Interpretation by Modules

The purpose of this section is to put some results of Steinitz and Krull into a form applicable to the problems of this paper.

* This research was partially supported by NSF Grant GP-7073.

For an $m \times n$ matrix A over R, let M_A denote the R-submodule of $R^{(n)}$ (= the n-th cartesian power of R) generated by the rows of A; and let

$$S_A = R^{(n)}/M_A.$$

(1.1) **Theorem.** *For $m \times n$ matrices A and B over R, $A \sim B \Leftrightarrow S_A \cong S_B$ (isomorphism of R-modules).*

In view of (1.1) (to be proved below), we will need a description of precisely which modules can occur as an S_A for some A which is $m \times n$ of rank r.

Let L be an *integral* (that is, nonzero) ideal of R, and let the factorization of L into powers of maximal ideals be

$$L = P_1^{e(1)} \ldots P_t^{e(t)} \quad (P_i \neq P_j \text{ when } i \neq j).$$

We will call the ideals $P_i^{e(i)}$ the *separated divisors* of L. For a finite sequence $L_1, \ldots, L_r$ of integral ideals, we define its *separated divisors*, notation

$$\operatorname{div}\{L_i\}_{i=1,r}$$

to be the collection, counting multiplicity, of all separated divisors of the individual ideals L_i, where "counting multiplicity" means that if P^e is a separated divisor of b of the ideals L_i, then P^e occurs b times in $\operatorname{div}\{L_i\}_{i=i,r}$.

For *fractional* ideals (= finitely generated nonzero submodules of the field of quotients Q of R) H and K, *class* H = *class* K will mean that $qH = K$ for some $0 \neq q \in Q$.

(1.2) **Separated Divisor Theorem.** (i) *The modules S_A, where A is $m \times n$ of rank r, are precisely those of the form*

$$S(L_1, \ldots, L_r; H)_{m \times n} = \begin{cases} \left(\bigoplus_{i=1}^{r} R/L_i\right) \oplus H \oplus R^{n-r-1} & \text{if } r<n \\ \bigoplus_{i=1}^{r} R/L_i & \text{if } r=n \end{cases}$$

where each L_i is an integral ideal, H a fractional ideal; and where

$$\textit{class } H = \textit{class} \prod_{i=1}^{r} L_i \quad \text{if} \quad r=m$$

$$\textit{class } H = \textit{class } R \quad \text{if}^1 \quad r = n \text{ or } 0.$$

(ii) *If A and A' are both $m \times n$ of rank r, then $A \sim A'$ if and only if*

$$\operatorname{div}\{L_i\}_{i=1,r} = \operatorname{div}\{L'_i\}_{i=1,r} \quad \textit{and class } H = \textit{class } H'.$$

In stating the above theorem we have adopted the convention that *the direct sum of the empty family of modules is the zero module*, and *the product of the empty family of ideals is R*. Thus $R^{(0)} = 0$ and $S(\phi, R)_{m \times n} = R^{(n)}$.

[1] If $r=n$ the choice of H is, of course, irrelevant. However, the choice indicated will enable us to avoid stating many special cases later.

Proof of (1.1). Krull [6, p. 18] calls two matrices A and B—*not necessarily* of the same size—over R "equivalent" if there exist invertible matrices P and Q, integers u and v, and blocks of zeros of appropriate size so that

$$\begin{bmatrix} P & 0 \\ 0 & I_u \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Q & 0 \\ 0 & I_v \end{bmatrix}=\begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix}.$$

(He writes this merely as $PAQ=B$.) We will call this equivalence *Krull equivalence* and use the notation $A\sim_K B$. Krull showed that, *for matrices A and B over R,*

$$(1.3)\qquad \begin{aligned} A\sim_K B &\Leftrightarrow S_A\cong S_B \\ &\Leftrightarrow \beta(M_A)=M_B \quad \text{for some } R\text{-automorphism } \beta \text{ of } R^{(n)} \end{aligned}$$

where the second equivalence holds provided A and B have the same number n of columns. The equivalence of the extreme left and right sides of (1.3) is a restatement of [6, Theorems 11 and 18], while the equivalence

$$S_A\cong S_B\Leftrightarrow\beta(M_A)=M_B$$

is a combination of [6, Theorems 19 and 17].

Another result we will need is [5, Theorem 2]: *for fractional ideals H_i and K_i of R,*

$$(1.4)\qquad H_1\oplus\cdots\oplus H_r\cong K_1\oplus\cdots\oplus K_r\Leftrightarrow \text{class}\prod_{i=1}^{r}H_i=\text{class}\prod_{i=1}^{r}K_i.$$

To prove Theorem 1.1 it will suffice, in view of (1.3), to prove that for matrices A and B *of the same size* $m\times n$,

$$(1.5)\qquad A\sim B\Leftrightarrow A\sim_K B.$$

For a proof of the nontrivial implication "$\Leftarrow$" let σ_A: $R^{(m)}\to R^{(n)}$ be right multiplication by A. Since every submodule of $R^{(n)}$ is projective [10, (5.3) p. 13 and (3.2) p. 132] the homomorphism σ_A maps $R^{(m)}$ *onto* a projective module, and hence its kernel is a direct summand of $R^{(m)}$. Thus we obtain the top row of the following diagram (for some U_A).

$$\begin{tikzcd} R^{(m)}=(\ker\sigma_A)\oplus U_A \arrow[r, "\sigma_A"] \arrow[d, dashed, "\alpha=\text{isomorphism}"'] & R^{(n)} \arrow[d, "\beta=\text{isomorphism}"] \\ R^{(m)}=(\ker\sigma_B)\oplus U_B \arrow[r, "\sigma_B"] & R^{(n)} \end{tikzcd}\qquad (\sigma_B \text{ is 1-1 on } U_B).$$

The bottom row is obtained similarly. Let β be the isomorphism given by (1.3), and note that $\sigma_A(R^{(m)})=M_A$. We now define the isomorphism (=isomorphism onto) α. On U_A define α to be the composition: σ_A then β then $(\sigma_B|U_B)^{-1}$. This shows that $U_A\cong U_B$ and hence, by (1.4), that $\ker\sigma_A\cong\ker\sigma_B$. Define α, on $\ker\sigma_A$, to be *any* isomorphism of $\ker\sigma_A$ onto $\ker\sigma_B$.

Since α and β are isomorphisms, they are given by right multiplication by invertible matrices (over R), which we will call P^{-1} and Q respectively. Then commutativity of the diagram above shows

$$XAQ = XP^{-1}B \qquad \text{for all } X \in R^{(m)}.$$

Taking X to be the element of $R^{(m)}$ which has 1 in coordinate i and zero elsewhere, we see that row i of AQ equals row i of $P^{-1}B$. Since this is true for every i, we get $AQ = P^{-1}B$, that is, $PAQ = B$, as desired.

Proof of Theorem (1.2). Let M be any submodule of $R^{(n)}$. Then [6, Theorems 13 and 15] state that there exist simultaneous decompositions

$$R^{(n)} = R y_1 \oplus \cdots \oplus R y_{r-1} \oplus H^{-1} \oplus H \oplus R^{(n-r-1)}, \tag{1}$$

$$M = L_1 y_1 \oplus \cdots \oplus L_{r-1} y_{r-1} \oplus L_r H^{-1} \tag{2}$$

where H^{-1} and H are isomorphic to fractional ideals and, by (1.4), H^{-1} is isomorphic to the ideal inverse to H; and where each L_i is an integral ideal. Since $H^{-1}/L_r H^{-1} \cong R/L_r$ [3, 18.24; or see Krull's proof, scattered through 6, §10], it follows from (1) and (2) that

$$R^{(n)}/M \cong R/L_1 \oplus \cdots \oplus R/L_r \oplus H \oplus R^{(n-r-1)} \tag{3}$$

where neither of the last two terms occur if $n = r$. This will prove that $S_A = R^{(n)}/M_A$ has the form $S(L_1, \ldots, L_r; H)_{m \times n}$ provided we know that the integer r in (3) equals the rank of A, and this follows from (2).

To obtain the "exceptional cases" stated in (i) of the theorem, suppose first that $r = m$. Then the right multiplication map $\sigma_A: R^{(m)}$ onto M_A is 1-1 and hence $R^{(m)} \cong M_A$. Thus, by (2) and (1.4) we see that class $(L_1 \ldots L_r H^{-1}) =$ class R, that is, class $H = \text{class} \prod_{i=1}^{r} L_i$.

When $r = 0$, so that no L_i's actually occur in (2), we see from (1) and (1.4) that class H = class R. Finally, we recall that the exceptional case $r = n$ is just a notational convention.

It is not difficult to see, in view of the above computations, that all the possibilities described in (i) can actually occur.

The proof of (ii) can be found by combining the formula

$$R/L = \oplus \{R/P^e \mid P^e \in \operatorname{div}\{L\}\}$$

[3, 18.18] with the Krull-Schmidt theorem [3, 14.5] and (1.4) above.

(1.6) *Remarks.* For the case that R is the ring of all algebraic integers in an algebraic number field, the content of Theorems 1.1 and 1.2 can be found scattered through [9]. However, I am not sure whether all of Steinitz's proofs are valid in an arbitrary Dedekind domain; and, in addition, I was not able to supply the steps omitted from his proof (p. 343, the paragraph before §67) that his definition of matrix equivalence ($\neq$ Krull equivalence) is the same as the "unimodular" equivalence we are using.

The strictly module-theoretic portions of Steinitz's and Krull's theorems about *pairs* of modules have been generalized in [7].

The following lemma will be used repeatedly in the next section. Its proof is an immediate consequence of the fact that $S_{\operatorname{diag}(B,C)} = S_B \oplus S_C$, where

$$\operatorname{diag}(B, C) = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}.$$

(1.8) **Diagonalization Lemma.** *Let*

$$S_B = S(L_1, \dots, L_r; H)_{m \times n} \quad \textit{and} \quad S_C = S(L'_1, \dots, L'_s; H)_{p \times q}.$$

Then

$$S_{\operatorname{diag}(B,C)} = S(L_1, \dots, L_r, L'_1, \dots, L'_s; HH')_{(m+p) \times (n+q)}.$$

Note. The Diagonalization Lemma remains true when $p=0$ (in which case $\operatorname{diag}(B, C)$ should be interpreted to mean B followed by q columns of zeros) and when $q=0$ (here $\operatorname{diag}(B, C)$ means B followed by p rows of zeros). In both of these cases we interpret S_C to mean $S(\phi; R)_{p \times q} = R^{(q)}$.

2. Diagonal Decomposition

In this section $\mathscr{C}(R)$ will denote the group of ideal classes of R: $\{\text{class } H \mid H = \text{fractional ideal}\}$ with multiplication $(\text{class } H)(\text{class } K) = \text{class}(HK)$. The order of $\mathscr{C}(R)$ will be written $|\mathscr{C}(R)|$.

(2.1) **Lemma.** *Let $A_{m \times n}$ be an $m \times n$ matrix of rank r over R. Then*

$$A \sim \begin{bmatrix} D_{p \times q} & 0 \\ 0 & 0 \end{bmatrix}_{m \times n} \quad \textit{where } p=r \textit{ or } r+1 \textit{ and } q=r \textit{ or } r+1.$$

Proof. Let $S_A = S(L_1, \dots, L_r; H)_{m \times n}$. First suppose that m and n are both $\geqq r+1$, and put $p=q=r+1$. Then

$$(*) \qquad S_A \cong S(L_1, \dots, L_r; H)_{p \times q} \oplus S(\phi; R)_{(m-p) \times (n-q)}.$$

None of the restrictions on the ideal class of H enumerated in the Separated Divisor Theorem (§1) arise in this case. The matrix associated with

$$S(\phi; R)_{(m-p) \times (n-q)} \cong R^{(n-q)}/0$$

is the zero matrix $O_{(m-p) \times (n-q)}$. If we choose for D a $p \times q$ matrix such that $S_D = S(L_1, \dots, L_r; H)_{p \times q}$ then $(*)$ together with the Diagonalization Lemma establish the desired result for the case being considered.

If $r=m$ the Separated Divisor Theorem shows that class $H = \text{class}(L_1 \dots L_r)$; and hence we can take $p=r$. Similarly if $r=n$ we can take $q=r$.

Note that if A is, say, 2×4 of rank 2 the lemma asserts that

$$A \sim \begin{bmatrix} a & b & c & 0 \\ d & e & f & 0 \end{bmatrix}.$$

(2.2) **Theorem.** *Suppose $\mathscr{C}(R)$ is finite and let $A_{m\times n}$ be indecomposable over R. Then m and n are both $\leqq|\mathscr{C}(R)|$.*

Proof. We will suppose that either m or n is $>|\mathscr{C}(R)|$ and show that A is "strictly decomposable", that is,

$$A\sim\begin{bmatrix}B_1 & 0\\ 0 & B_2\end{bmatrix} \quad\text{or}\quad \begin{bmatrix}B_1\\ 0\end{bmatrix} \quad\text{or}\quad [B_1 \quad 0]$$

with B_1, B_2, and the zero blocks actually present.

If m or $n\geqq r+2$, then A is decomposable by the previous lemma. Hence we can suppose

(1) $$m=r \text{ or } r+1 \quad\text{and}\quad n=r \text{ or } r+1 \quad\text{hence}\quad r\geqq|\mathscr{C}(R)|.$$

Let $S_A=S(L_1,\ldots,L_r;H)_{m\times n}$, and consider the ideals

(2) $$(L_1),(L_1L_2),(L_1L_2L_3),\ldots,(L_1L_2\ldots L_r).$$

If some two of these ideals are in the same class, say $\alpha\prod_{i=1}^{u}L_i=\prod_{i=1}^{v}L_i$ with $u<v$, then we can use invertibility of the ideals L_i to cancel the first u of them, getting $\alpha R=\prod_{i=u+1}^{v}L_i$. Thus, *after a suitable renumbering of the L_i*, we obtain

(3) $$\operatorname{class}\prod_{i=1}^{p}L_i=\operatorname{class}R \qquad (\text{for some } p\leqq r-1).$$

This will yield the decomposition (by the Diagonalization Lemma of § 1)

(4) $$S_A\cong\underbrace{S(L_1,\ldots,L_p;R)_{p\times p}}_{Term\,1}\oplus\underbrace{S(L_{p+1},\ldots,L_r;H)_{(m-p)\times(n-p)}}_{Term\,2}$$

provided that the restrictions in the Separated Divisor Theorem (when they occur) are met.

Note that *Term* 1 represents a nonsingular matrix. Eq. (3) gives the condition which must be satisfied because its rank p equals its number of rows; and since R occurs after the semicolon in *Term* 1, the condition due to $p=$ the number of columns is also satisfied.

The restrictions in the Separated Divisor Theorem occur in *Term* 2 when its rank $r-p$ equals $m-p$ or $n-p$. (Note that $r-p\neq 0$ since $p\leqq r-1$); that is when $r=m$ or n. When $r=m$, the Separated Divisor Theorem, applied to A, shows that

$$\operatorname{class}H=\operatorname{class}(L_1L_2\ldots L_r).$$

Since the product of the first p of these ideals is principal (by (3)), we see that $L_{p+1}L_{p+2}\ldots L_r$ is in the class of H, as required for *Term* 2 of (4). Similarly we dispose of the case $r=n$.

Thus we have established the decomposition (4) whenever some two of the ideals in (2) belong to the same class.

Suppose, now, that no two of the ideals in (2) belong to the same class. We conclude from the inequality in (1) that, in (2), each element of $\mathscr{C}(R)$ appears exactly once. In particular, the principal class occurs, say as class $(L_1 L_2 \dots L_p)$. If $p \leqq r-1$, as in (3), then we again obtain the decomposition (4). If $p=r$, then *Term* 2 of (4) becomes $S(\phi; H)$, which is admissible only if H is principal (in which case (4) works again).

Thus we are left with the case that $H \neq$ principal, $L_1 L_2 \dots L_r$ *is* principal, and where (2) enumerates every class exactly once. Since H is not principal, $r \neq n$ (by the Separated Divisor Theorem, applied to A); so $r+1=n$.

Since (2) enumerates all the elements of $\mathscr{C}(R)$, there is a q such that

(5) $$\text{class } H = \text{class} \prod_{i=1}^{q} L_i \qquad (q \leqq r-1)$$

where the inequality results from the present hypothesis that $L_1 L_2 \dots L_r$ is principal. It follows that $q<m$, that is, A has enough rows to admit the following "strict decomposition"

(6) $$S_A \cong S(L_1, \dots, L_q; H)_{q \times (q+1)} \oplus S(L_{q+1}, \dots, L_r; R)_{(m-q) \times (n-q-1)}.$$

To see that A has enough columns, merely observe that $q+1 \leqq r < r+1 = n$. That (6) satisfies the remaining conditions imposed by the Separated Divisor Theorem now follows from (5) and the corresponding condition (if any) satisfied by A.

(2.3) **Proposition**. *Suppose $\mathscr{C}(R)$ contains an element of order $e>1$. Then there exist indecomposable matrices over R of the following types:*

(i) *$e \times e$ nonsingular.*

(ii) *$e \times e$ of rank $e-1$.*

(iii) *$e \times (e-1)$ and $(e-1) \times e$ of rank $e-1$.*

Proof. By hypothesis there is a fractional ideal L such that L^e is principal, but L^d is not principal for $1 \leqq d < e$. In fact there is an integral such L (replace L by rL where r is a common denominator for some finite set of generators of L). By the Separated Divisor Theorem there is a nonsingular $e \times e$ matrix A such that

(1) $$S_A = S(\underbrace{L, L, \dots, L}_{e \text{ terms}}; R)_{e \times e}.$$

In any decomposition of the form

(2) $$A \sim \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix}$$

both B and D must be nonsingular, say $S_B = S(L_1, \dots, L_p; R)_{p \times p}$ and $S_D = (L_{p+1}, \dots, L_e; R)_{(e-p) \times (e-p)}$ where $1 \leqq p < e$ and

(3) $$\prod_{i=1}^{p} L_i \quad \text{is principal}.$$

Then the Diagonalization Lemma shows that $S_A \cong S(L_1, \ldots, L_e; R)_{e\times e}$ and therefore

(4) $$\operatorname{div}\{L_1, L_2, \ldots, L_e\} = \operatorname{div}\{\underbrace{L, L, \ldots, L}_{e \text{ terms}}\}.$$

Let P be a maximal ideal and suppose P^d is the highest power of P which divides L. Then (4) shows that the highest power of P which divides each L_i must be P^d. In other words, every $L_i = L$. But then (3) becomes "L^e is principal" contradicting our choice of e. Thus A is indecomposable and we have established (i).

For (ii) and (iii) take the same ideal L and choose a matrix A of one of the following types

(5) $$S_A = S(\underbrace{L, L, \ldots, L}_{e-1 \text{ terms}}; R)_{e\times e},$$

(6) $$S_A = S(\underbrace{L, L, \ldots, L}_{e-1 \text{ terms}}; R)_{e\times(e-1)}.$$

In either case suppose we have a decomposition of the form (2) with B of rank p and D of rank q $(p+q=e-1)$. If A is of the form (6) then B must have p columns and D must have q columns; and hence one of them must be nonsingular. We can choose the equivalence (2) in such a way that B is nonsingular. The argument given for (i) now establishes the desired contradiction.

The same reasoning, applied to (5) shows that there is no decomposition of the form (2) with B nonsingular. The remaining possibility is that B is $p\times(p+1)$ of rank p and D is $(q+1)\times q$ of rank q. Then the two exceptional cases in the Separated Divisor Theorem show

$$S_B = S(\underbrace{L, L, \ldots, L}_{p \text{ terms}}; L^p) \quad \text{and} \quad S_D = S(\underbrace{L, L, \ldots, L}_{q \text{ terms}}; R).$$

But then $S_B \oplus S_D$ is not isomorphic to S_A (by the second part of the Separated Divisor Theorem and the Diagonalization Lemma) because $L^p \neq$ principal.

The proof is now complete.

Combining this with the theorem which preceeded it we get:

(2.4) **Corollary.** *If $\mathscr{C}(R)$ is finite and cyclic, the maximum size for an indecomposable matrix A is $|\mathscr{C}(R)|$ rows $\times\, |\mathscr{C}(R)|$ columns. Such an A must either be nonsingular or have rank $|\mathscr{C}(R)|-1$.*

(2.5) **Proposition.** *If $\mathscr{C}(R)$ has an element of infinite order, then there exist unboundedly large indecomposable nonsingular matrices over R.*

Proof. By hypothesis R has an ideal H no nonzero power of which is principal. Hence the same is true of some prime factor P of H. For every nonzero element x in P there is a factorization (e and each $e(i)>0$)

(1) $$P^e \underbrace{P_1^{e(1)} P_2^{e(2)} \ldots P_t^{e(t)}}_{\text{call this ideal } Q} = Rx \qquad (P, P_1, \ldots, P_t \text{ distinct primes}).$$

Choose an x which makes t as small as possible (note that, by hypothesis, $t \neq 0$).

Let n be an arbitrary positive integer. By the Separated Divisor Theorem there is a nonsingular $ne \times ne$ matrix A such that

(2) $$S_A = S(PQ^n, \underbrace{P, P, \dots, P}_{ne-1 \text{ terms}}; R)_{ne \times ne}.$$

We prove that A is indecomposable. If not, then in any "strict decomposition" $A = \operatorname{diag}(B, C)$, both B and C must be nonsingular. Hence any such decomposition of A would yield a decomposition

(3) $$\begin{aligned} S_{\operatorname{diag}(B,C)} &= S(L_1, \dots, L_r; R)_{r \times r} \oplus S(L_{r+1}, \dots, L_{ne}; R)_{(ne-r) \times (ne-r)} \\ &\cong S(L_1, \dots, L_{ne}; R)_{ne \times ne} \qquad (ne > r). \end{aligned}$$

Then the second assertion of the Separated Divisor Theorem shows:

(4) $$\operatorname{div}\{L_i\}_{i=1,ne} = \{\underbrace{P, P, \dots, P}_{ne \text{ terms}}, P_1^{ne(1)}, \dots, P_t^{ne(t)}\}.$$

This forces the highest power of P which divides each L_i to be P itself. Since the term $S(L_1, \dots, L_r; R)_{r \times r}$ in (3) comes from the nonsingular matrix B, we must have $\prod_{i=1}^{r} L_i$ principal, say Ry. Factoring this into a product of primes, with the help of (4), gives an expression (after a suitable renumbering of the P_i)

(5) $$P^r P_1^{ne(1)} P_2^{ne(2)} \dots P_u^{ne(u)} = Ry \qquad \text{for some } u \leqq ne.$$

Minimality of t in (1) shows that $u = t$. Hence combining (1) and (5) shows

$$P^{ne-r} = R(x^n/y) = \text{principal}$$

so that, by our choice of P, $ne - r = 0$, contrary to (3); and this completes the proof.

(2.6) **Proposition.** *Let $\mathscr{C}(R)$ have finite exponent e, and let A be an indecomposable $m \times n$ matrix over R. Then m and n are both $\leqq 2e-1$.*

Proof. Recall that the rank r of an indecomposable matrix cannot differ from m or n by more than 1 (2.1). Therefore it will suffice to show that if $r \geqq 2e-1$, then A must be strictly decomposable. Since the case $e=1$, that is, $|\mathscr{C}(R)| = 1$ was covered in (2.2), we can suppose $e > 1$.

Let $S_A = S(L_1, \dots, L_r; H)_{m \times n}$ and factor each L_i into a product of powers of distinct prime ideals. Let P be one of the primes which arise in this way, and let the power of P which appears in the factorization of $L_1, \dots, L_r$ be, respectively,

(1) $$P^{d(1)}, P^{d(2)}, \dots, P^{d(r)} \quad (\text{each } d(i) \geqq 0).$$

A theorem of Erdős, Ginzburg, and Ziv [2] states that if $a_1, a_2, \dots, a_{2v-1}$ is a sequence of elements in a solvable group of order v, then the a_i can be renumbered in such a way that $a_1 \cdot a_2 \dots a_v = 1$. (A shorter proof of this theorem is given in [8].) It is easily seen, by repeated use of the theorem, that the theorem remains true if the order of the group merely *divides* v.

Since the cyclic subgroup of $\mathscr{C}(R)$ generated by class P has order dividing e, we can apply the above theorem to the first $2e-1$ exponents in (1), remembering that $r \geqq 2e-1$, to obtain

$$(2) \qquad (\text{class } P^{d(1)})(\text{class } P^{d(2)}) \dots (\text{class } P^{d(e)}) = \text{class } R;$$

in other words, the product of the first e terms in (1) is principal. Note that the necessary rearrangement of (1) can be accomplished by redistributing the appropriate powers of P among the L_i without altering either $\operatorname{div}\{L_i\}$ or the distribution of the remaining primes which divide the L_i.

After doing this with each prime P which appears in the factorization of some L_i we obtain that $\prod_{i=1}^{e} L_i$ is principal. This permits the decomposition

$$(3) \qquad S_A = S(L_1, \dots, L_e; R)_{e \times e} \oplus S(L_{e+1}, \dots, L_r; H)_{(m-e) \times (n-e)}$$

provided that the exceptional cases in the elementary divisor theorem do not cause difficulty with the summand on the extreme right of (3). The situations in which these exceptional cases arise are

$$(4) \qquad \begin{aligned} & r-e=m-e && (\Leftrightarrow r=m) \\ \text{or } & r-e=n-e && (\Leftrightarrow r=n) \\ \text{or } & r-e=0 && (\Leftrightarrow r=e). \end{aligned}$$

There is one more possible exceptional situation: If $m-e=0=n-e$ then the extreme right-hand term of (3) is zero.

The first two possibilities cannot cause difficulty because

$$\text{class}\left(\prod_{i=e+1}^{r} L_i\right) = \text{class}\left(\prod_{i=1}^{r} L_i\right) \quad \left(\text{since } \prod_{i=1}^{e} L_i \text{ is principal}\right)$$

so the condition in question is guaranteed by the corresponding condition for S_A.

If $r=e$, then from $r \geqq 2e-1$ we get the case $e=1$ which has already been disposed of.

Finally, if $m-e=0=n-e$, that is, $m=n=e$, then A is $e \times e$; and since $e \leqq 2e-1$ when $e \geqq 1$, there is nothing to be proved about A.

Thus matrices corresponding to the right-hand side of (3) decompose A as desired.

Remark. I am indebted to Sylvia Wiegand and Henry Mann for calling my attention to the above theorem of Erdős, Ginzburg, and Ziv, thereby considerably improving my original bound in Proposition 2.6. However, since the papers referred to are not all easy to obtain, I would like to add that only the *existence* of a bound to the size of indecomposable matrices (not its actual value) will be needed in the proof of Theorem 2.7 below.

It is easy to see that e^e will do, in place of $2(e-1)$, in Proposition 2.6: If $r \geqq e^e$ then some exponent must occur in (1) at least e times (modulo e), and hence we can obtain Eq. (2) with $d(1) \equiv d(2) \equiv \dots \equiv d(e)$ (modulo e).

Establishing the second main theorem is now merely a matter of putting together the preceding results.

(2.7) **Theorem.** *There is a bound to the sizes of all indecomposable matrices over R if and only if $\mathscr{C}(R)$ has finite exponent.*

Proof. If $\mathscr{C}(R)$ does not have finite exponent then either: (i) It contains elements of unboundedly large finite order, in which case (2.3) shows that R has unboundedly large indecomposable matrices; or else (ii) It contains an element of infinite order, in which case (2.5) shows R has unboundedly large indecomposable matrices.

Conversely, if $\mathscr{C}(R)$ has finite exponent e, then (2.6) gives a bound for all indecomposable matrices over R.

Remark. In this paper we have found an interpretation of two of the invariants of $\mathscr{C}(R)$ in terms of the matrix theory over R, namely the order and exponent of $\mathscr{C}(R)$. In view of the theorem of Claborn [1] that every abelian group can appear as $\mathscr{C}(R)$ for *some* Dedekind domain R, it would be interesting to know what other invariants of $\mathscr{C}(R)$ can appear in the matrix theory over R.

References

1. Claborn, L.: Every Abelian group is a class group. Pacific J. Math. **18**, 219–222 (1966).
2. Erdős, P., Ginzburg, A., Ziv, A.: Bull. Research Council Israel **10** (1961).
3. Curtis, C.W., Reiner, I.: Representation theory of finite groups and associative algebras. New York: Wiley 1962.
4. Kaplansky, I.: Elementary divisors and modules. Trans. Amer. Math. Soc. **66**, 464–491 (1949).
5. – Modules over Dedekind rings and valuation rings. Trans. Amer. Math. Soc. **72**, 327–340 (1952).
6. Krull, W.: Matrizen, Moduln und verallgemeinerte Abelsche Gruppen im Bereich der ganzen algebraischen Zahlen. Heidelberger Akademie der Wissenschaften **2**, 13–38 (1932).
7. Levy, L.S.: Decomposing pairs of modules. Trans. Amer. Math. Soc. **122**, 64–80 (1966).
8. Mann, H.B.: Two addition theorems. J. Combinat. Theory **3**, 233–235 (1967).
9. Steinitz, E.: Rechteckige Systeme und Moduln in algebraischen Zahlkörpern. I and II. Math. Ann. **71**, 328–354 (1912) and **72**, 297–345 (1912).
10. Cartan, H., Eilenberg, S.: Homological algebra. Princeton, N.J.: Princeton University Press 1956.

Prof. Lawrence S. Levy
Mathematics Department
University of Wisconsin
Madison, Wis.
USA

(Received March 1, 1971)

Math. Z. 124, 100–118 (1972)

Eine Kennzeichnung der symplektischen und orthogonalen Gruppen über $GF(4)$

KLAUS-JÜRGEN FLEISCHER

In dieser Arbeit wird der folgende Satz bewiesen:

Hauptsatz. *Sei G eine endliche Gruppe, die von einer Konjugiertenklasse D von Involutionen erzeugt wird. Außerdem gelte*

(i) *Sind $d, e \in D$, so gilt $(d \cdot e)^n = 1$ für ein $n \in \{2, 3, 5\}$.*

(ii) *Es existieren drei paarweise verschiedene Elemente $d, e, f \in D$ mit $(d \cdot e)^3 = 1 = (d \cdot f)^5$.*

(iii) $\mathbb{O}_2(G) = 1 = \mathbb{Z}(G)$.

(iv) *Sind $d, f \in D$ zwei verschiedene Elemente mit $(d \cdot f)^5 = 1$, so ist $\mathbb{C}_D(e) \cap \langle d, f \rangle \neq \emptyset$ für alle $e \in D$.*

Dann ist

(a) *$G \simeq \mathrm{Sp}(2n, 4)$ für ein $n \geqq 1$, falls zwei Elemente $c, d \in D$ existieren mit $c d \in D$,*

(b) *$G \simeq O^{\pm}(2n, 4)$ für ein $n \geqq 2$, falls $c d \notin D$ für alle $c, d \in D$.*

Alle unter (a) und (b) angegebenen Gruppen erfüllen die Voraussetzungen. Involutionen, die der Bedingung (i) genügen, heißen $\{3, 5\}$-Transpositionen in Analogie zu den von Fischer untersuchten $\{3\}$-Transpositionen [4]. Bedingung (ii) schließt Entartungsfälle aus; gibt es keine zwei verschiedenen Elemente aus D, deren Produkt die Ordnung 5 hat, so wird G von $\{3\}$-Transpositionen erzeugt, und diese Gruppen wurden in [4] bestimmt. Gibt es keine zwei Elemente aus D, deren Produkt die Ordnung 3 hat, so wird G von $\{5\}$-Transpositionen erzeugt. Es wird vermutet, daß alle derartigen Gruppen auflösbar sind.

Bedingung (iii) schließt Gruppen mit auflösbaren Normalteilern aus. Andere im wesentlichen einfache Gruppen, die ebenfalls von $\{3, 5\}$-Transpositionen erzeugt werden, werden durch Bedingung (iv) ausgeschlossen. Dies sind die Gruppen $SU(n, 4^2)$ für $n \geqq 3$ und $O(n, 5)$ für $n \geqq 3$. Es ist nicht bekannt, ob es weitere Gruppen mit dieser Eigenschaft gibt.

Durch (iv) ist es möglich, alle Untergruppen zu bestimmen, die von 3 Involutionen aus D erzeugt werden, was für den Beweis des Satzes wichtig ist. Dies geschieht im 2. Kapitel.

Der eigentliche Beweis des Hauptsatzes erstreckt sich auf die Kapitel 3 bis 6. Wir bestimmen dabei zunächst die Struktur von $\langle \mathbb{C}_D(d) \rangle$, des „$D$-Zentralisators" einer Involution $d \in D$. Kapitel 3 bestimmt die Anzahl der Involutionen aus D im Zentrum; im 4. Kapitel zeigen wir mit Induktion, daß $\langle \mathbb{C}_D(d)\rangle$ modulo seinem maximalen auflösbaren Normalteiler zu einer symplektischen Gruppe $\mathrm{Sp}(2n-2, 4)$ isomorph ist. Diese Kenntnisse, zusammen mit Bedingung (iv), gestatten uns, die Mächtigkeit von D zu berechnen.

Im 5. Kapitel konstruieren wir dann für den Fall (a) aus D einen symplektischen Raum, auf dem G als Automorphismengruppe operiert. Für Fall (b) zeigen wir im 6. Kapitel, daß der von D gebildete Graph zu einem Graph der Transvektionen einer orthogonalen Gruppe isomorph ist und daß G wie eine orthogonale Gruppe auf D operiert. (Ein Graph besteht hierbei aus den Involutionen, zusammen mit den Ordnungen der Produkte von je zwei Involutionen.)

Ich möchte Herrn B. Fischer für Ratschläge und hilfreiche Diskussionen ganz besonders danken.

1. Bezeichnungen und Grundlagen

Alle betrachteten Gruppen sind endlich. Zur Bezeichnung von speziellen Gruppen verwenden wir folgende Symbole

C_p	die zyklische Gruppe der Ordnung p;
D_{2p}	die Diedergruppe der Ordnung $2p$;
E_{p^f}	die elementar abelsche Gruppe der Ordnung p^f;
A_5	die alternierende Gruppe vom Grad 5;
Σ_n	die symmetrische Gruppe vom Grad n;
$W(D_4)$	die Weyl-Gruppe zur Lie-Algebra D_4.

Die Bezeichnung der klassischen Gruppen richtet sich nach Huppert [5].

Ist g ein Gruppenelement, so bezeichnet $o(g)$ die Ordnung von g. Werden im folgenden Involutionen graphisch dargestellt, so bedeutet eine einfache Verbindungslinie zweier Elemente d, e, daß $o(d \cdot e)=3$ ist; eine dreifache Verbindung besagt, daß $o(d \cdot e)=5$ ist. Sind zwei Involutionen nicht verbunden, so sind sie vertauschbar. Falls nicht ausdrücklich anders vorausgesetzt, so sind zwei verschieden dargestellte Elemente stets voneinander verschieden. (Es gilt also $o(d \cdot e)=$Anzahl der Verbindungsstriche$+2$.) Eine derartige Darstellung einer Menge D_0 von Involutionen nennen wir den *Graph von* D_0.

(1.1) Definition. Sei G eine Gruppe und D eine Menge von Involutionen von G mit $\langle D\rangle = G$. Sei N eine (endliche) Menge natürlicher Zahlen ≥ 3. Dann heißt D eine Menge von *N-Transpositionen* von G, wenn folgende Bedingungen gelten

(i) $D^G = D$.

(ii) Sind $d, e \in D$, so ist $o(d \cdot e) \in \{1, 2, N\}$.

D heißt eine Menge von *N^+-Transpositionen*, falls zusätzlich gilt

(iii) Zu jedem $n \in N$ gibt es Elemente $d, e \in D$ mit $o(d \cdot e)=n$.

Im folgenden bezeichnet D stets eine Konjugiertenklasse von $\{3, 5\}^+$-Transpositionen und $G=\langle D\rangle$ das Erzeugnis von D. Außerdem gelte für D stets die Bedingung

(Z) *Sind* $d, x \in D$ *mit* $o(d \cdot x) = 5$, *so ist* $\mathbb{C}_D(e) \cap \langle d, x\rangle \neq \emptyset$ *für alle* $e \in D$.

Wir bezeichnen mit (S) und (O) folgende zwei Bedingungen (symplektischer und orthogonaler Fall):

(S) *Es existieren zwei Elemente* $c, d \in D$ *mit* $cd \in D$.

(O) *Für zwei beliebige Elemente* $c, d \in D$ *gilt stets* $cd \notin D$.

Mit diesen Bezeichnungen formulieren wir noch einmal den

Hauptsatz. *Ist* $\mathbb{Z}(G) = 1 = \mathbb{O}_2(G)$, *so ist*

(a) $G \simeq \mathrm{Sp}(2n, 4)$ *für ein* $n \geqq 1$, *falls* (S) *gilt;*

(b) $G \simeq O^{\pm}(2n, 4)$ *für ein* $n \geqq 2$, *falls* (O) *gilt.*

Wir werden in dieser Arbeit einige Eigenschaften von symplektischen und orthogonalen Räumen über $GF(4)$ und deren zugehörigen Gruppen benötigen, die wir im folgenden kurz zusammenstellen wollen. Beweise finden sich teilweise in [3]; einige spezielle Eigenschaften sind leicht verifizierbar.

Sei V ein $2n$-dimensionaler Vektorraum über $GF(4)$, dem Körper mit 4 Elementen, wobei $n \geqq 1$. Ist auf V eine nicht ausgeartete symmetrische Bilinearform $(\ ,\ )$ definiert, für die $(v, v) = 0$ für alle $v \in V$ gilt, so ist V ein symplektischer Raum. Sind $a, b \in V$ zwei Vektoren mit $(a, b) \neq 0$, so heißt das Erzeugnis $\langle a, b\rangle$ eine hyperbolische Ebene von V. Das orthogonale Komplement einer hyperbolischen Ebene $\langle a, b\rangle$ ist ein $(2n-2)$-dimensionaler Unterraum $U = \langle a, b\rangle^{\perp}$, und V ist die direkte Summe von $\langle a, b\rangle$ und U. Wir schreiben $V = \langle a, b\rangle \perp U$. Offenbar läßt sich V in eine direkte Summe von n hyperbolischen Ebenen zerlegen, die paarweise orthogonal zueinander sind.

Sei $a \in V$ mit $a \neq 0$. Dann ist die Abbildung

$$t(a):\quad v \to v + (v, a)\, a \qquad \text{für } v \in V$$

eine Transvektion von V. Umgekehrt gibt es zu jeder Transvektion t von V ein $a \in V \setminus \{0\}$, so daß $t = t(a)$. Die Abbildung $a \to t(a)$ für $a \in V \setminus \{0\}$ ist bijektiv. Alle Transvektionen von V sind Involutionen, und es gilt für $a, b \in V \setminus \{0\}$

$$\begin{aligned} &t(a) \cdot t(b) = t(b) \cdot t(a) && \text{falls } (a, b) = 0;\\ &o\big(t(a) \cdot t(b)\big) = 3 && \text{falls } (a, b) = 1;\\ &o\big(t(a) \cdot t(b)\big) = 5 && \text{falls } (a, b) \notin \{0, 1\}. \end{aligned}$$

Da alle symplektischen Gruppen von ihren Transvektionen erzeugt werden, ist $T = \{t(a) \mid a \in V \setminus \{0\}\}$ eine Konjugiertenklasse von $\{3, 5\}^+$-Transpositionen von $\mathrm{Sp}(2n, 4)$.

Durch zusätzliche Definition einer quadratischen Form Q auf V mit

$$Q(k \cdot a + h \cdot b) = k^2 \cdot Q(a) + h^2 \cdot Q(b) + k \cdot h \cdot (a, b)$$

für $k, h \in GF(4)$ und $a, b \in V$ wird V zu einem orthogonalen Raum über $GF(4)$. Die symplektische Transvektion $t(a)$ ist genau dann auch orthogonale Transvektion, wenn $Q(a)=1$. Für Q gibt es bis auf Isomorphie zwei Möglichkeiten; in einem Fall ist die Dimension des maximalen singulären Unterraums n, im andern Fall $n-1$. (Ein Unterraum W heißt singulär, wenn für jedes $w \in W$ gilt, daß $Q(w)=0$.) Wir bezeichnen die zugehörigen Gruppen mit $O^{+}(2n, 4)$ und $O^{-}(2n, 4)$.

(1.2) Lemma. *Sei G eine der Gruppen* $\mathrm{Sp}(2n, 4)$, $O^{+}(2n, 4)$ *oder* $O^{-}(2n, 4)$. *Sei T die Menge der Transvektionen aus G.*

(i) *Ist* $G \simeq \mathrm{Sp}(2n, 4)$, *so ist* $|T| = 4^{2n}-1$.

(ii) *Ist* $G \simeq O^{+}(2n, 4)$, *so ist* $|T| = 4^{n-1}(4^{n}-1)$.

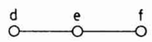
(iii) *Ist* $G \simeq O^{-}(2n, 4)$, *so ist* $|T| = 4^{n-1}(4^{n}+1)$.

Diese Zahlen ergeben sich durch Abzählen der Vektoren $\neq 0$ im symplektischen Fall und durch Abzählen der Vektoren a mit $Q(a)=1$ im orthogonalen Fall.

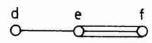
2. Gruppen, die von 3 Involutionen erzeugt werden

In diesem Kapitel bestimmen wir alle Untergruppen von G, die von drei Involutionen aus D erzeugt werden. Nach Coxeter und Moser [2], gelten folgende Aussagen

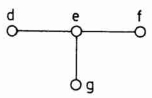
(2.1) Lemma. *Seien d, e, f, g Involutionen.*

(i) *Gilt* 

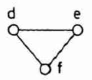
, so ist $\langle d, e, f\rangle \simeq \Sigma_4$.

(ii) *Gilt* 

, so ist $\langle d, e, f\rangle = A$, *wobei* $A \simeq A_5 \times \mathbb{Z}(A)$ *und* $|\mathbb{Z}(A)| \leqq 2$.

(iii) *Gilt* 

, so ist $\langle d, e, f, g\rangle = W$, *wobei* $W \simeq W(D_4)$ *oder* $W \simeq W(D_4)/\mathbb{Z}(W(D_4))$.

(iv) *Gilt* 

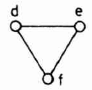
, und ist $o(d^{e} \cdot f)=3$, *so ist* $\langle d, e, f\rangle = H$, *und* $|H|=54$.

Die Involutionen in (ii) und (iii) können jeweils zwei verschiedene Gruppen erzeugen, die sich durch die Ordnung des Zentrums unterscheiden. Der Bequemlichkeit halber werden wir in Zukunft für beide Gruppen die Bezeichnungen A_5^{*} und $W^{*}(D_4)$ gebrauchen. $A \simeq A_5^{*}$ bedeutet also $A \simeq A_5$ oder $A \simeq A_5 \times C_2$; Entsprechendes gilt für $W^{*}(D_4)$.

(2.2) Lemma. *Seien* $d, e, f \in D$ *mit* $o(d \cdot e)=3$ *und* $o(d \cdot f)=5$. *Dann ist* $\langle d, e, f\rangle \simeq A_5^{*}$.

Beweis. Wegen (Z) ist e mit einem Element $g \in \langle d, f\rangle \cap D$ vertauschbar. Nach (2.1.ii) folgt $\langle d, e, f\rangle = \langle d, e, g\rangle \simeq A_5^{*}$.

(2.3) Lemma. *Seien $d, e, f \in D$ mit*



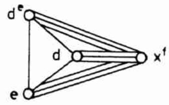
Dann gilt $d^e \cdot f = f \cdot d^e$.

Beweis. Angenommen, es ist $o(d^e \cdot f) = 5$. Dann ist $\langle d, e, f\rangle = \langle d, d^e, f\rangle \simeq A_5^*$ nach (2.2); in A_5^* existieren jedoch keine drei Elemente, welche die vorausgesetzten Relationen erfüllen.

Angenommen, es ist $o(d^e \cdot f) = 3$. Dann ist $\langle d, e, f\rangle = H$ eine Gruppe der Ordnung 54 nach (2.1.iv). Es folgt $|d^H| = 9$. Sei $x \in D$ mit $o(x \cdot f) = 5$. Wegen (Z) ist jede der neun Involutionen aus d^H mit einer Involution aus $\langle x, f\rangle$ vertauschbar. O.B.d.A. kann deshalb x mit d und e vertauschbar gewählt werden. Dann gilt für jedes $y \in \{d, e, d^e\}$

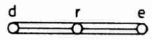
$$o(y \cdot x^f) = o(y^f \cdot x) = o(f^y \cdot x) = o(f \cdot x) = 5;$$

also



Nach (2.2) ist $\langle d, e, x^f\rangle \simeq A_5^*$; jedoch gibt es diese Relation in A_5^* nicht. Demnach ist $d^e \cdot f = f \cdot d^e$.

(2.4) Lemma. *Seien $d, r, e \in D$ mit*



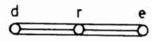
Dann ist $o(r^{de} \cdot r) \neq 5$.

Beweis. Angenommen, es ist $o(r^{de} \cdot r) = 5$. Es ist $o(r^{de} \cdot d) = o(r \cdot d) = 5$; $o(r^{de} \cdot r^d) = o(r^e \cdot r) = 5$ und $o(r^{de} \cdot d^{rd}) = o(r^e \cdot d^r) = o(r^{er} \cdot d) \neq 2$. Da r^{de} mit einem Element aus $\{d, d^r, d^{rd}, r^d, r\}$ vertauschbar ist, gilt $o(r^{de} \cdot d^r) = 2$. Deshalb ist $o(r^{de \cdot r} \cdot d) = 2 = o(r^{der \cdot de} \cdot d) = o(r^{(r^{de})} \cdot d)$. Wegen $o(r^{de} \cdot r) = 5$ ist $r \in \langle r^{de \cdot r}, r^{(r^{de})}\rangle$, und somit auch $o(r \cdot d) = 2$, entgegen der Voraussetzung.

Sei V ein symplektischer Raum der Dimension $2n \geqq 4$, und seien $a, b \in V$ mit $(a, b) = 1$. Sei $c \in \langle a, b\rangle^\perp$ mit $c \neq 0$. Sei $B = \langle t(a), t(\alpha \cdot b), t(a + c)\rangle$, wobei $\alpha \in GF(4)$ und $0 \neq \alpha \neq 1$. Dann ist $|B| = 160$, und B ist das semidirekte Produkt eines elementar abelschen Normalteilers der Ordnung 16 mit einer D_{10}. In B gibt es 20 konjugierte Transvektionen, die in 5 Klassen zu je 4 paarweise vertauschbare Involutionen zerfallen. (Wir übergehen hier die Verifikation dieses Ergebnisses.)

Bezeichnung. *Wir werden im folgenden den Isomorphietyp dieser Gruppe B mit B_5 bezeichnen.*

(2.5) Lemma. *Seien $d, r, e \in D$ mit*



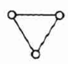
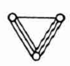
Ist $o(r^{de} \cdot r) = 3$, so ist $\langle d, r, e\rangle \simeq A_5^*$. Ist $o(r^{de} \cdot r) = 2$, so ist $\langle d, r, e\rangle \simeq B_5$.

Beweis. Ist $o(r^{de} \cdot r) = 3$, so folgt nach (2.2), daß $\langle d, r, e\rangle = \langle r^{de}, r, e\rangle \simeq A_5^*$.

Sei $o(r^{de} \cdot r) = 2$. Dann erfüllen die Transvektionen $t(a)$, $t(\alpha \cdot b)$ und $t(a+c)$ die gleichen Relationen wie d, r und e. Demnach besitzt $\langle d, r, e\rangle$ als homomorphes Bild eine B_5. (Wegen $d \neq r \neq e \neq d$ gibt es keine Relation zwischen d, r und e, die nicht auch von den Transvektionen erfüllt wird.) Das Nebenklassenabzählverfahren von Todd und Coxeter [6] ergibt $|\langle d, r, e\rangle| \leqq 160$. Demnach ist $\langle d, r, e\rangle \simeq B_5$.

Damit sind alle Untergruppen von G, die von 3 Involutionen aus D erzeugt werden, bestimmt. Es gilt folgender

(2.6) Satz. *Seien $d, e, f \in D$ mit $d \neq e \neq f \neq d$. Dann ist $\langle d, e, f\rangle$ isomorph zu einer der folgenden Gruppen*

d, e, f		$\langle d, e, f\rangle$	*Beweis*
(1)		E_8 oder E_4	
(2)		$C_2 \times D_6$	
(3)		$C_2 \times D_{10}$	
(4)		Σ_4	(2.1.i)
(5)		A_5^*	(2.1.ii)
(6)		A_5^* oder B_5	(2.4) und (2.5)
(7)		Σ_4 oder Σ_3	(2.3)
(8)		A_5^*	(2.2)
(9)		A_5^*	(2.2)

Die Relation

ist wegen (Z) zu einer Relation (6) äquivalent.

3. TI-Mengen

Ist C der Zentralisator einer Involution $d \in D$, so ist es möglich, daß $\mathbb{Z}(C)$ noch weitere Involutionen aus D enthält. Diese bilden zusammen mit d eine TI-Menge. In diesem Kapitel bestimmen wir die Mächtigkeit solcher TI-Mengen.

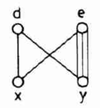
(3.1) Bezeichnungen. *Sei $d \in D$. Dann ist*

$$\begin{aligned} E_d &= \{e \in D \mid \mathbb{C}_D(e) = \mathbb{C}_D(d)\}; \\ D_d &= \mathbb{C}_D(d) \setminus E_d; \\ A_d &= D \setminus \mathbb{C}_D(d); \\ A_{d,3} &= \{x \in D \mid o(x \cdot d) = 3\}; \\ A_{d,5} &= \{x \in D \mid o(x \cdot d) = 5\}. \end{aligned}$$

(3.2) Lemma. *Sei $d \in D$ und $e \in E_d$. Ist $A_{d,3} \cap A_{e,3} \neq \emptyset$, so gilt $A_{d,3} = A_{e,3}$.*

Beweis. Sei $x \in A_{d,3} \cap A_{e,3}$. Angenommen, es gibt ein $y \in A_{d,3} \cap A_{e,5}$. Dann tritt einer der folgenden drei Fälle ein:

1)

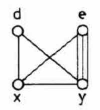


Dann ist

$$2 = o(x \cdot y) = o(x \cdot d^{yd}) = o(x \cdot e^{yd}) = o(d^x \cdot e^y) = o(e^x \cdot e^y) = o(x \cdot e^{ye}),$$

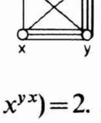
da $\mathbb{C}_D(d^{yd}) = \mathbb{C}_D(e^{yd})$ und $\mathbb{C}_D(d^x) = \mathbb{C}_D(e^x)$. Wegen $o(x \cdot y) = 2 = o(x \cdot e^{ye})$ folgt $o(x \cdot e) = 2$, entgegen der Voraussetzung.

2)



Nach (2.3) ist $d \cdot x^y = x^y \cdot d$. Es ist aber $\langle e, x, y \rangle \simeq A_5^*$ und $o(e \cdot x^y) \neq 2$. Demnach ist $e \notin E_d$; ein Widerspruch.

3)



Dann ist $\langle d, x, y \rangle \simeq A_5^*$ und $o(d \cdot x^{yx}) = 2$. Außerdem folgt $\langle e, x, y \rangle \simeq A_5^*$ und $o(e \cdot x^{yx}) \neq 2$. Dies besagt, daß $e \notin E_d$; ein Widerspruch.

Somit ist $A_{d,3} \cap A_{e,5} = \emptyset$, und es folgt $A_{d,3} = A_{e,3}$.

(3.3) Lemma. *Seien d, e vertauschbare Elemente von D mit $o(d \cdot f) = o(e \cdot f)$ für alle $f \in D \setminus \{d, e\}$. Dann ist $d \cdot e \in \mathbb{O}_2(G)$.*

Beweis. Offenbar ist $e \in E_d$ und $A_{d,3} = A_{e,3}$. Demnach gehören d, e zu $V_d = \{v \in E_d | A_{v,3} = A_{d,3}\}$. Sei $Q(V_d) = \{v_1 \cdot v_2 | v_1, v_2 \in V_d\}$ und $E = E_d$. Wir werden zeigen, daß die Elemente von $Q(V_d)$ auf $\{E^g | g \in G\}$ trivial operieren; also daß $E^{gq} = E^g$ für $g \in G$ und $q \in Q(V_d)$.

O.B.d.A. können wir annehmen, daß $q = d \cdot e$. Es genügt zu zeigen, daß $\mathbb{C}_D(f^{de}) = \mathbb{C}_D(f)$ ist für $f \in E^g$. Ist $f \in \mathbb{C}_D(d) = \mathbb{C}_D(e)$, so ist $f^{de} = f$. Ist $f \in A_{d,3} = A_{e,3}$, so folgt, daß $\mathbb{C}_D(d^f) = \mathbb{C}_D(e^f)$ und somit $\mathbb{C}_D(f^d) = \mathbb{C}_D(f^e)$. Sei daher $f \in A_{d,5} = A_{e,5}$. Nach (2.6) ist $\langle d, e, f \rangle \simeq A_5^*$ oder $\simeq B_5$. Wäre $\langle d, e, f \rangle \simeq A_5^*$, so gäbe es ein Element in $A_{d,3} \cap A_{e,5}$. Deshalb ist $\langle d, e, f \rangle = B \simeq B_5$.

Angenommen, es gäbe ein $x \in \mathbb{C}_D(f^{de}) \cap A_f$. Dann ist x wegen (Z) mit genau einem Element $a \in \mathbb{C}_D(d) \cap B$ vertauschbar. Wiederum wegen (Z) folgt aber aus $e \in E_d$, daß auch $a \in E_d$. Demnach ist x mit jedem Element von $\mathbb{C}_D(d) \cap B$ vertauschbar. Dieser Widerspruch beweist, daß $\mathbb{C}_D(f^{de}) = \mathbb{C}_D(f)$ und daß somit $Q(V_d)$ auf $\{E^g | g \in G\}$ trivial operiert.

Sei $M = \langle Q(V_d)^G \rangle$. Es folgt nun $M \leqq \mathbb{O}_2(G)$ und demnach $d \cdot e \in \mathbb{O}_2(G)$.

(3.4) Korollar. *Sei $d \in D$ und $e \in E_d$ mit $d \neq e$. Sei $\mathbb{O}_2(G) = 1$. Dann ist $A_{d,3} \cap A_{e,3} = \emptyset$.*

Beweis. Angenommen, es ist $A_{d,3} \cap A_{e,3} \neq \emptyset$. Nach (3.2) ist dann $A_{d,3} = A_{e,3}$, und aus (3.3) folgt $de \in \mathbb{O}_2(G) = 1$. Es ist also $d = e$ im Widerspruch zur Voraussetzung.

(3.5) Lemma. *Sei $d \in D$ und $x \in A_d$. Dann ist $\langle \mathbb{C}_D(d), x \rangle = G$.*

Beweis. Sei $\langle \mathbb{C}_D(d), x \rangle = H$. Ist $x \in A_{d,5}$, so folgt wegen (Z), daß

$$H = \langle \mathbb{C}_D(d) \cup \mathbb{C}_D(d^x) \cup \mathbb{C}_D(d^{xd}) \cup \mathbb{C}_D(x^d) \cup \mathbb{C}_D(x) \rangle = \langle D \rangle = G.$$

Ist $x \in A_{d,3}$, so existiert wegen $A_{x,5} \neq \emptyset$ und (Z) ein $e \in A_{x,5} \cap \mathbb{C}_D(d)$. Da d und e unter $\langle d, x, e \rangle \simeq A_5^*$ konjugiert sind, ist $\mathbb{C}_D(e) \subseteq H$, und es folgt

$$H \geqq \langle \mathbb{C}_D(e), x \rangle = G.$$

(3.6) Satz. *Sei $d \in D$, und sei $\mathbb{O}_2(G) = 1 = \mathbb{Z}(G)$. Dann ist*

(a) $|E_d| = 3$, *falls (S) gilt:*

(b) $|E_d| = 1$, *falls (O) gilt.*

Beweis. Es gelte (S). Sei $c \in D$ mit $cd \in D$, und sei $f \in \mathbb{C}_D(d)$. Angenommen, es ist $o(c \cdot f) = p \in \{3, 5\}$. Dann ist $(cd \cdot f)^2 = (cf)^2 \neq 1$ und $(cd \cdot f)^p = d$; also $o(cd \cdot f) = 2p$, was nicht sein kann. Demnach ist $\mathbb{C}_D(d) = \mathbb{C}_D(c) = \mathbb{C}_D(cd)$, und $|E_d| \geqq 3$.

Angenommen, es ist $|E_d| > 3$, und seien $\{c, d, cd, b\} \subseteq E_d$. Sei $x \in A_{b,3}$. Dann folgt $x \in A_{c,5} \cap A_{d,5} \cap A_{cd,5}$ nach (3.4), und $\langle x, c, d \rangle$ ist zu keiner der Gruppen aus (2.6) isomorph; ein Widerspruch. Demnach gilt (a).

Angenommen, es gilt (O) und $|E_d| \geqq 2$. Sei $e \in E_d$ mit $e \neq d$, und sei $x \in A_{d,3}$. Dann ist $x \in A_{e,5}$ nach (3.4), und $\langle d, x, e \rangle = A \simeq A_5^*$. Es gibt genau ein $a \in \mathbb{C}_D(d) \cap A$ mit $d \neq a \neq e$. Ist $y \in \mathbb{C}_D(d) = \mathbb{C}_D(e)$, so folgt wegen (Z), daß $y \in \mathbb{C}_D(a)$.

Demnach ist $\mathbb{C}_D(e) \subseteq \mathbb{C}_G(ade)$ und außerdem $x \in \mathbb{C}_G(ade)$, da $ade \in \mathbb{Z}(A)$. Nach (3.5) ist aber $\langle \mathbb{C}_D(e), x\rangle = G$; deshalb ist $ade \in \mathbb{Z}(G) = 1$; im Widerspruch zu (O). Deshalb gilt auch (b).

4. Der Zentralisator einer Involution

In diesem Kapitel bestimmen wir die Struktur von $\langle D_d \rangle$.

(4.1) Lemma. *Sei $\mathbb{Z}(G) = 1$, und sei $d \in D$. Ist $D_d = \emptyset$, so ist $G \simeq \mathrm{Sp}(2, 4)$.*

Beweis. Sei $D_d = \emptyset$. Seien $d, e, x \in D$ mit $o(d \cdot e) = 3$ und $o(d \cdot x) = 5$. Wegen (Z) gilt $D = \bigcup_{y \in Y} E_y$, wobei $Y = \langle x, d \rangle \cap D$. Demnach ist $|D| \leqq 15$. Andererseits ist $\langle d, e, x \rangle \simeq A_5^*$, und es folgt $D \subseteq \langle d, e, x \rangle = G$. Wegen $\mathbb{Z}(G) = 1$ ist $G \simeq A_5$, und $A_5 \simeq \mathrm{Sp}(2, 4)$ nach [1].

Demnach ist der Hauptsatz für $D_d = \emptyset$ richtig, und wir können im folgenden Induktion verwenden. Wir setzen für den Rest des Kapitels folgende

Voraussetzung. *Es sei $\mathbb{Z}(G) = 1 = \mathbb{O}_2(G)$ und $D_d \neq \emptyset$.*

(4.2) Lemma. *D_d ist eine Konjugiertenklasse in $\langle D_d \rangle$.*

Beweis. Angenommen, D_d ist keine Konjugiertenklasse in $\langle D_d \rangle$. Dann gibt es eine Zerlegung $D_d = \bigcup_{i=1}^{m} W_i$ mit $m > 1$, wobei die W_i Konjugiertenklassen in $\langle W_i \rangle$ sind, und $W_j \subseteq \mathbb{C}_D(W_i)$ für $i \neq j$. Es ist o.B.d.A. $|W_1| \geqq |W_i|$ für $1 \leqq i \leqq m$. Sei $W = W_1$.

Wäre $|W| = 1$, so wäre $\langle E_d \cup D_d \rangle$ abelsch und $D_d = \emptyset$, im Widerspruch zur Voraussetzung. Deshalb ist $|W| > 1$.

Sei $K = \mathbb{N}_D(W) \setminus W$. Dann ist K ein normaler Komplex in $\langle \mathbb{N}_D(W) \rangle$. Es folgt $K \subseteq \mathbb{C}_D(W)$, und $\{E_d, W_2\} \subseteq K$. Sei $x \in K$. Dann ist $W \subseteq D_x$. Da W eine Konjugiertenklasse in $\langle W \rangle$ ist, folgt wegen der Maximalität von $|W|$, daß W ein normaler Komplex von $\langle D_x \rangle$ ist. Demnach ist $\mathbb{C}_D(x) \subseteq \mathbb{N}_D(W)$ für alle $x \in K$.

Es ist sogar $\mathbb{N}_D(W) = \mathbb{C}_D(x)$ für $x \in K$. Sonst wäre nämlich $G \leqq \mathbb{N}_G(W)$ nach (3.5), und $W^G = W$, was nicht sein kann. Deshalb ist $\mathbb{C}_D(d) = \mathbb{N}_D(W) = \mathbb{C}_D(x)$ für $x \in K$, und $K \subseteq E_d$; ein Widerspruch zu $W_2 \subseteq K$.

Wir konstruieren nun folgendermaßen eine Kette von Normalteilern von $\langle D_d \rangle$:

$$N_1 = M = \mathbb{O}_2(\langle D_d \rangle)$$
$$N_{2i}/N_{2i-1} = \mathbb{Z}(\langle D_d \rangle / N_{2i-1}) \quad \text{für } i = 1, 2, \ldots$$
$$N_{2i+1}/N_{2i} = \mathbb{O}_2(\langle D_d \rangle / N_{2i}) \quad \text{für } i = 1, 2, \ldots.$$

Offenbar gibt es eine natürliche Zahl k, so daß $N_k = N_{k+1} = N_{k+2}$, und es ist dann $\mathbb{O}_2(\langle D_d \rangle / N) = 1$ und $\mathbb{Z}(\langle D_d \rangle / N) = 1$ für $N = N_k$.

(4.3) Lemma. *Seien $e, f \in D_d$. Dann gilt $e \cdot f \in N$ genau dann, wenn*

$$e \cdot f \in M = \mathbb{O}_2(\langle D_d \rangle).$$

Beweis. Sei $e \cdot f \in N$. Sei k die kleinste natürliche Zahl, so daß $N_k = N$. Ist $k=1$, so ist nichts zu zeigen. Sei also $k>1$, und sei $g \in D_d \cap A_e$. Dann ist $o(g \cdot e) = p \in \{3, 5\}$.

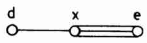
Angenommen, es ist $o(gN \cdot eN) = 1$. (Dabei bezeichnen gN und eN Elemente der Faktorgruppe $\langle D_d \rangle / N$.) Dann ist $eg \in N$; also $eg N_{k-1} \in \mathbb{Z}(\langle D_d \rangle / N_{k-1})$ oder $\in \mathbb{O}_2(\langle D_d \rangle / N_{k-1})$. Wegen $o(g \cdot e) = p$ impliziert dies $eg \in N_{k-1}$, und nach Induktion folgt $eg \in M$, was nicht sein kann.

Demnach ist $o(gN \cdot eN) \neq 1 \neq o(gN \cdot fN)$. Da $eN = fN$, gilt

$$p = o(g \cdot e) = o(gN \cdot eN) = o(gN \cdot fN) = o(g \cdot f).$$

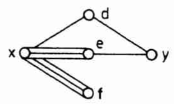
Nach (3.3), angewandt auf die Menge D_d, folgt nun $ef \in \mathbb{O}_2(\langle D_d \rangle) = M$.

(4.4) Lemma. *Seien $d, x, e \in D$ mit*



und sei $f \in \mathbb{C}_D(d) \cap \langle d, x, e \rangle$ mit $d \neq f \neq e$. Ist $y \in A_{d,3} \cap A_{e,3}$, so ist $y \in A_{f,3}$.

Beweis. Angenommen, es sind y und f vertauschbar.



$o(x\,y) = ?$

Wegen (Z) ist y außerdem mit einem Element von $\{x, x^e, e^{xe}, e^x\}$ vertauschbar. O.B.d.A. genügt es, zwei Fälle zu betrachten.

1. Fall. $o(x \cdot y) = 2$. Dann ist

$$o(e \cdot y) = o(x^{dfxdfx} \cdot y) = o(x^{fdxd} \cdot y) = o(x^{fxdx} \cdot y) = o(x^{fx} \cdot d^y) = o(x^{fx} \cdot d) = 5,$$

ein Widerspruch zur Voraussetzung.

2. Fall. $o(e^x \cdot y) = 2$. Dann ist $o(x^{de} \cdot d) = o(x \cdot d) = 3$; $o(y \cdot d) = 3$ und

$$o(x^{de \cdot y} \cdot d) = o(x^{de} \cdot y^d) = o(x^e \cdot y) = o(x^{exex} \cdot y) = o(e \cdot y) = 3.$$

Nun ist aber $o(x^{de} \cdot y) = o(x^f \cdot y) = o(x \cdot y) = 5$, und $\langle x^{de}, y, d \rangle \simeq A_5^*$. Dann folgt aber aus $o(x^{de} \cdot d) = 3$ und $o(y \cdot d) = 3$, daß $o(x^{de \cdot y} \cdot d) = 5$; ein Widerspruch.

Angenommen, es ist $o(f \cdot y) = 5$. Wegen (Z) ist y mit einem Element von $\{x, x^f, f^{xf}, f^x\}$ vertauschbar, und nach geeigneter Wahl von x (evtl. Vertauschung von d und e) gilt $y \in \mathbb{C}_D(f^{xf})$. Dann ist $o(x \cdot y) = 5$. Es folgt, daß entweder $e^x \in \mathbb{C}_D(y)$ oder $x^e \in \mathbb{C}_D(y)$. Wäre $e^x \in \mathbb{C}_D(y)$, so wäre auch $\langle e^x, f^{xf} \rangle \cap D \subseteq \mathbb{C}_D(y)$, und $d \in \mathbb{C}_D(y)$. Deshalb ist $x^e \in \mathbb{C}_D(y)$. Nun sind $\{x^e, f^{xf}, e^{xf}\}$ drei paarweise vertauschbare Involutionen aus $\langle x, e, f \rangle \simeq A_5^*$, und wegen (Z) ist auch $e^{xf} \in \mathbb{C}_D(y)$. Es ergibt sich der folgende Widerspruch

$$3 = o(e \cdot y) = o(e^{fxexf} \cdot y) = o(x^{ef} \cdot y) = o(x^d \cdot y) = o(d \cdot y^x) = 5.$$

Deshalb ist $o(f \cdot y) = 3$.

(4.5) Satz. *Es ist* $\langle D_d\rangle/N \simeq \mathrm{Sp}(2n-2, 4)$ *für geeignetes* $n \geqq 2$.

Beweis. Die Aussage folgt nach Induktion, wenn wir beweisen, daß in $\langle D_d\rangle/N$ die Bedingung (S) erfüllt ist. Wegen (3.6) genügt es zu zeigen, daß zwei verschiedene Elemente $e, f \in D_d$ existieren, so daß $\mathbb{C}_D(e) \cap D_d = \mathbb{C}_D(f) \cap D_d$ und $ef \notin N$.

Sei $x \in A_{d,3}$. Angenommen, es ist $D_d \cap A_{x,5} = \emptyset$. Dann gilt wegen (Z), daß $A_{x,5} \subseteq \langle x, E_d\rangle$, und $\langle x, E_d\rangle \cap D$ ist eine TI-Menge von G. Sei $y \in D_d \cap A_{x,3}$. Dann ist wegen $d^{yx} \in \langle x, E_d\rangle$ auch $x^{yx} = y \in \langle x, E_d\rangle$. Dies kann jedoch nicht sein; es ist also $D_d \cap A_{x,3} = \emptyset$, und $D_d = D_x$. Nun folgt $D_d^{\langle x, \mathbb{C}_D(d)\rangle} = D_d^G = D_d$, ein Widerspruch.

Demnach existiert also ein $e \in D_d \cap A_{x,5}$, und $\langle d, x, e\rangle = A \simeq A_5^*$. Somit gibt es ein $f \in D_d \cap A$ und $f \neq e$. Wegen (Z) folgt $\mathbb{C}_D(e) \cap D_d = \mathbb{C}_D(f) \cap D_d$.

Angenommen, es ist $D_d \cap A_{e,3} = \emptyset$. Dann ist auch $D_d \cap D_x = \emptyset$; denn D und D_d sind Konjugiertenklassen. Sicherlich gibt es jedoch ein $z \in D_d \cap A_{e,5}$, und x ist mit einer Involution aus $\langle e, z\rangle$ vertauschbar; ein Widerspruch zu $D_d \cap D_x = \emptyset$.

Sei $y \in D_d \cap A_{e,3}$. Wegen (4.4) ist $y \notin A_{f,3}$, und es folgt $ef \notin N$.

Bezeichnung. Sei $e \in D_d$ und $M = \mathbb{O}_2(\langle D_d\rangle)$. Wir setzen $q = |e^M|$.

(4.6) Lemma. *Es ist* $|D_d| = q \cdot |D_d N/N|$.

Beweis. Da $D_d \cap N = \emptyset$, genügt es zu zeigen, daß $eM \cap D_d = e^M$. Klar ist, daß $e^M \subseteq eM \cap D_d$. Sei also $f \in eM \cap D_d$. Dann gibt es ein $y \in D_d \cap A_{e,3} = D_d \cap A_{f,3}$. Es gilt dann $f = e^{yefy} \in e^M$.

(4.7) Lemma. *Sei* $e \in D_d$. *Dann ist* $q = |e^M| \leqq 4$ *und* $|e^M \cap D_x| \leqq 1$ *für* $x \in A_d$. *Ist* $q = 4$, *so gilt* $|e^M \cap D_x| = 1$ *für alle* $x \in A_d$.

Beweis. Sei $f \in e^M$ und $x \in A_d$. Offenbar gilt

$$\mathbb{C}_D(d) \cap \mathbb{C}_D(e) = \mathbb{C}_D(d) \cap \mathbb{C}_D(f) = \mathbb{C}_D(e) \cap \mathbb{C}_D(f).$$

Deshalb kann x nicht sowohl mit e als auch mit f vertauschbar sein, und $|e^M \cap D_x| \leqq 1$.

Sei $y \in D_d \cap A_e \cap A_{x,5}$. (Bei geeigneter Wahl von e existiert solch ein y sicherlich.) Dann ist jedes Element von e^M mit einem Element aus $\{x, x^y, x^{yx}, y^x\}$ vertauschbar. Wegen $|e^M \cap D_x| \leqq 1$ für alle $x \in A_d$ folgt $q = |e^M| \leqq 4$. Ist $q = 4$, so folgt außerdem $|e^M \cap D_x| = 1$ für alle $x \in A_d$.

(4.8) Bezeichnung. Sei $d \in D$ und $e \in D_d$. Dann ist

$$F(d, e) = \{f \in D_d \mid \mathbb{C}_D(e) \cap D_d = \mathbb{C}_D(f) \cap D_d\}.$$

Nach (4.5) gilt offenbar $|F(d, e)| = 3 \cdot q$.

(4.9) Lemma. *Es ist*

(a) $q = 4$, *falls* (S) *gilt;*

(b) $q = 1$, *falls* (O) *gilt.*

Beweis. Angenommen, es ist $q>1$. Sei $e\in D_d$ und $f\in e^M$ mit $f\neq e$. Sei $x\in A_{e,3}\cap A_{f,5}$. Dann ist $\langle e,x,f\rangle=A\simeq A_5^*$. Sei $g\in A\cap \mathbb{C}_D(e)$ und $e\neq g\neq f$. Wegen (*Z*) gilt $\mathbb{C}_D(e)\cap D_d=\mathbb{C}_D(f)\cap D_d\subseteq \mathbb{C}_D(g)\cap D_d$, und $g\in F(d,e)\cup E_d$. Sei $y\in D_d\cap A_{e,3}$. Dann ist auch $y\in A_{f,3}$, und nach (4.4) folgt $y\in A_{g,3}$. Deshalb ist $g\in e^M$, und $q=4$.

Es gelte (*S*). Dann ist $E_e\subseteq F(d,e)$, und außerdem liegt das in (4.5) konstruierte f in $F(d,e)$. Dabei ist $f\notin E_e$; demnach ist $q>1$ und nach dem ersten Abschnitt $q=4$.

Es gelte nun (*O*). Angenommen, es ist $q>1$; also $q=4$. Dann ist $|F(d,e)\cap D_x|=3$ für jedes $x\in A_d$ nach (4.7). Sei $h\in F(d,e)\cap D_x$. Es ist $F(h,d)=(F(d,e)\cup\{d\})\smallsetminus\{h\}$. Dann ist aber $|F(h,d)\cap D_x|=2$, was nicht sein kann. Deshalb ist $q=1$.

(4.10) Korollar. *Es ist*

(a) $|D_d|=4^{2n-1}-4$, *falls* (*S*) *gilt;*

(b) $|D_d|=4^{2n-2}-1$, *falls* (*O*) *gilt.*

Beweis. Es ist D_dN/N die Menge der symplektischen Transvektionen von $\langle D_dN/N\rangle\simeq \mathrm{Sp}(2n-2,4)$. Die Aussage folgt deshalb nach (1.2.i) und (4.6).

5. Die symplektischen Gruppen

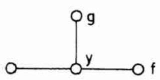
Wir werden einen symplektischen Vektorraum konstruieren, auf dem G als Automorphismengruppe operiert. Für dieses Kapitel gelte folgende

Voraussetzung. *Es sei* $\mathbb{Z}(G)=1=\mathbb{O}_2(G)$; $D_d\neq\emptyset$, *und es gelte* (*S*).

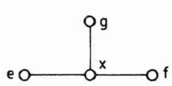
(5.1) Lemma. *Sei* $d\in D$, $e\in D_d$, $x\in A_d$ *und* $M=\mathbb{O}_2(\langle D_d\rangle)$. *Dann ist* $|e^M\cap A_{x,3}|=1$.

Beweis. Sei $e^M=\{e,f,g,h\}$. Dann ist $F(d,e)=E_e\cup E_f\cup E_g\cup E_h$. Nach (4.7) ist x mit genau drei Elementen aus $F(d,e)$ vertauschbar; o.B.d.A. sei $E_h\subseteq D_x$. Da $\langle E_a,x\rangle\simeq A_5$ für jedes $a\in A_x$, ist $|F(d,e)\cap A_{x,3}|=3$. Es genügt demnach zu zeigen, daß $|e^M\cap A_{x,3}|\leqq 1$.

Angenommen, es ist $|e^M\cap A_{x,3}|>1$. Seien $e,g\in A_{x,3}$. Ist $f\in A_{x,5}$, so folgt wie im Beweis von (4.9), daß $\langle e,f,x\rangle=A\simeq A_5^*$ und $\mathbb{C}_D(e)\cap A\subseteq e^M$, was nicht sein kann. Deshalb ist auch $f\in A_{x,3}$. Sei $y\in D_d\cap A_{e,3}$. Dann gilt



und $\langle e,f,g,y\rangle\simeq W^*(D_4)$. Da $(ef)^y\in M$, ist $g^{yefy}\in g^M=e^M$; jedoch keines der Elemente $\{e,f,g\}$. Also ist $g^{yefy}=h$ und $(efgh)^y=efgh$. Sei $e_1\in E_e$ mit $e_1\neq e$, und $e_1^M=\{e_1,f_1,g_1,h_1\}$. In $\langle e,y,e_1\rangle\simeq A_5$ gibt es einen inneren Automorphismus i, der e auf e_1 abbildet. Demnach gilt $efgh=(efgh)^i=e_1f_1g_1h_1$. Ebenso folgt $efgh=ee_1ff_1gg_1hh_1=(efgh)^2$; also $efgh=1$. Nun gilt



und $\langle e, f, g, x\rangle = W \simeq W^*(D_4)$. Wegen $h=efg$ ist $h\in W$. Es folgt $h\in \mathbb{Z}(W)$, was wiederum wegen $h=efg$ nicht sein kann.

(5.2) Lemma. *Sei $d\in D$. Es ist $|D|=4^{2n}-1$ und $|A_d|=3\cdot 4^{2n-1}$.*

Beweis. Wegen (Z) ist $|D|=5\cdot|\mathfrak{C}_D(d)|-4\cdot|\mathfrak{C}_D(d)\cap\mathfrak{C}_D(x)|$ für $x\in A_d$, und $|A_d|=|D|-|\mathfrak{C}_D(d)|$. Nach (4.10) und (4.7) ist $|\mathfrak{C}_D(d)|=4^{2n-1}-1$ und

$$|\mathfrak{C}_D(d)\cap\mathfrak{C}_D(x)|=4^{2n-2}-1.$$

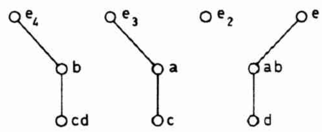
Durch Rechnung folgt die Behauptung.

(5.3) Lemma. *$\mathfrak{C}_G(d)$ hat auf $A_{d,5}$ genau zwei Bahnen.*

Beweis. Sei $x\in A_{d,5}$. Für jedes $z\in A_{d,5}$ gilt, daß $\langle x, d, z\rangle = X$ zu einer der Gruppen D_{10}, A_5^* oder B_5 isomorph ist. In jedem Fall ist z unter $\mathfrak{C}_X(d)$ zu x oder d^{xd} konjugiert, so daß $\mathfrak{C}_G(d)$ auf $A_{d,5}$ höchstens zwei Bahnen hat.

Angenommen, $\mathfrak{C}_G(d)$ hat auf $A_{d,5}$ nur eine Bahn. Seien $a, b\in A_{d,5}$, wobei $b\in E_a\smallsetminus\{a\}$. Dann gibt es ein $g\in\mathfrak{C}_G(d)$ mit $a^g=b$, und o.B.d.A. ist g ein 2-Element. Es folgt $c^g=cd$ für $c\in E_d\smallsetminus\{d\}$. Wir betrachten nun G als Permutationsgruppe auf $E=\{E_d|d\in D\}$. (Wenn im folgenden von Permutationen die Rede ist, so bezieht sich das stets auf diese Darstellung.) Da $|A_d|=3\cdot 4^{2n-1}$, induziert jedes $d\in D$ eine gerade Permutation von E, und wegen $g\in\langle D\rangle$ bewirkt auch g eine gerade Permutation von E. Wir werden jedoch zeigen, daß g als ungerade Permutation auf E operieren muß.

Sei $e\in D_d$ und $F(d,e)=F$. Sei $E_e=\{e, f, ef\}$ und $e^M=\{e, e_2, e_3, e_4\}$. Ist $F^g=F$, so normalisiert g als 2-Element mindestens eine der drei Mengen e^M, f^M und $(ef)^M$. Sei o.B.d.A. $e^{Mg}=e^M$. Dann gelten bei geeigneter Bezeichnung folgende Relationen (es ist nur die Ordnung 3 gekennzeichnet).



Da $a^g=b$; $b^g=a$, folgt $e^g=e$, $e_2^g=e_2$ und $e_3^g=e_4$. Somit induziert g auf F eine Transposition.

Ist $F^g\neq F$, so induziert g auf $\{E_y|y\in F\cup F^g\cup F^{g^2}\cup\ldots\}$ eine gerade Permutation, da alle vorkommenden Zyklen in gerader Anzahl auftreten. Wegen $|D_d|=4^{2n-1}-4$ ist die Anzahl der $F(d,e)$, die von g festgelassen werden, ungerade, und g induziert insgesamt auf $\{E_e|e\in D_d\}$ eine ungerade Permutation.

Sei $x\in A_d$. Dann ist $\langle x, E_d\rangle\simeq A_5$. Sei $A=(\langle x, E_d\rangle\cap D)\smallsetminus\{E_d\}$. Ist $A^g=A$, so induziert g auf A eine ungerade Permutation. Wegen $|A_d|=3\cdot 4^{2n-1}$ geschieht dies in einer geraden Anzahl von Fällen. Ist $A^g\neq A$, so bewirkt g auf $\{E_x|x\in A\cup A^g\cup A^{g^2}\cup\ldots\}$ eine gerade Permutation, da alle vorkommenden Zyklen in gerader Anzahl auftreten. Demnach induziert g auf $\{E_x|x\in A_d\}$ eine gerade Permutation, und somit auf E eine ungerade. Das liefert uns den gesuchten Widerspruch.

Es sei V ein Vektorraum über $GF(4)$, dessen Basiselemente die Involutionen aus D sind. Zwischen den Basiselementen sei folgende Form definiert.

Sei $d \in D$; $x \in A_{d,5}$ und $GF(4)=\{0, 1, \alpha, \alpha^2\}$.

$$\begin{array}{ll} (d, e)=0 & \text{für } e \in \mathfrak{C}_D(d); \\ (d, y)=1 & \text{für } y \in A_{d,3}; \\ (d, z)=\alpha & \text{für } z \in x^{\mathfrak{C}_G(d)}; \\ (d, z)=\alpha^2 & \text{für } z \in A_{d,5} \setminus x^{\mathfrak{C}_G(d)}; \\ (a, b)=(d, b^g) & \text{für } a, b \in D; \text{ wobei } g \text{ so gewählt ist, daß } a^g=d. \end{array}$$

Diese Form sei auf V als Bilinearform fortgesetzt. Sei V_0 der Nullraum von V bezüglich dieser offenbar symplektischen Form. (Es ist klar, daß das Element d durch die Definition nicht vor den anderen ausgezeichnet wurde.)

(5.4) Lemma. *Zu jedem $k \in \{\alpha, \alpha^2\}$ und $d \in D$ existiert ein $c \in E_d$ mit $(k \cdot d, v)=(c, v)$ für alle $v \in V$. Umgekehrt existiert zu jedem $c \in E_d \setminus \{d\}$ ein $k \in \{\alpha, \alpha^2\}$ mit $(k \cdot d, v)=(c, v)$ für alle $v \in V$.*

Beweis. Es genügt zu zeigen, daß $(k \cdot d, v)=(c, v)$ für alle $v \in D$. Ist $v \in \mathfrak{C}_D(d)$, so folgt für jedes $c \in E_d$

$$(k \cdot d, v)=k \cdot (d, v)=0=(c, v).$$

Sei $k \in \{\alpha, \alpha^2\}$ gegeben, und sei $a \in A_{d,3}$. Nach (5.3) ist $c \in E_d$ durch $(c, a)=k$ eindeutig bestimmt. Sei $v \in A_d$ mit $(c, v)=k$. Nach (5.3) ist $\mathfrak{C}_G(c)=\mathfrak{C}_G(d)$, und a und v sind unter $\mathfrak{C}_G(c)$ konjugiert. Demnach ist auch $v \in A_{d,3}$ und $(d, v)=1$. Ist $w \in E_v$ mit $(c, w)=k^2$, so folgt entsprechend $w \in A_{cd,3}$ und $(d, w)=k$. Es ist dann $(c, vw)=1$ und $(d, vw)=k^2$. Demnach gilt für alle $v \in A_d$

$$(k \cdot d, v)=k \cdot (d, v)=(c, v).$$

Die zweite Aussage ist eine triviale Folgerung aus der ersten.

(5.5) Lemma. *Sei $d \in D$ und $c \in E_d$. Dann ist $d+c+cd \in V_0$, d.h. $(d+c+cd, v)=0$ für alle $v \in V$.*

Beweis. Wieder genügt es zu zeigen, daß $(d+c+cd, v)=0$ für alle $v \in D$. Ist $v \in \mathfrak{C}_D(d)$, so ist $(d, v)=0=(c, v)=(cd, v)$. Ist $v \in A_d$, so ist $\langle E_d, v\rangle \simeq A_5$, und o.B.d.A. gilt $v \in A_{d,3}$. Da c und cd unter $\mathfrak{C}_G(v)$ nicht konjugiert sind, folgt $(c, v)=k$ und $(cd, v)=k^2$ für ein $k \in \{\alpha, \alpha^2\}$, und es ist

$$(d+c+cd, v)=(d, v)+(c, v)+(cd, v)=1+k+k^2=0.$$

(5.6) Lemma. *Sei $d \in D$; $e \in D_d$ und $x \in A_{d,3} \cap A_{e,3}$. Sei $n(d, e)=e^M \cap D_x$. Dann ist $n(d, e)$ von der Wahl von x unabhängig, und es gilt $n(d, n(d, e))=e$.*

Beweis. Sei $x \in A_{d,3} \cap A_{e,3}$ und $f=e^M \cap D_x$. Sei $M_f=\mathfrak{O}_2(\langle \mathfrak{C}_D(f)\rangle)$. Dann sind d und e unter M_f konjugiert. Für jedes $y \in A_{d,3} \cap A_{e,3}$ folgt deshalb nach (5.1), daß $y \in D_f$. Somit ist $n(d, e)$ von der Wahl von x unabhängig. Sei $m \in M$ mit $e^m=f$ und $f^m=e$. Da $d^m=d$, folgt $n(d, f)=e=n(d, n(d, e))$.

(5.7) Lemma. *Sei $d \in D$; $e \in D_d$ und $f = n(d, e)$. Dann ist $d+e+f \in V_0$.*

Beweis. Es genügt zu zeigen, daß $(d+e+f, v)=0$ für alle $v \in D$. Ist $v \in \mathfrak{C}_D(d)$, so ist $(e, v)=(f, v)$; denn wenn $o(e \cdot v)=5=o(f \cdot v)$, so ist $\langle e, v, f\rangle \simeq B_5$, und e und f sind unter $\mathfrak{C}_G(v)$ konjugiert. In diesem Fall gilt demnach

$$(d+e+f, v)=(d, v)+(e, v)+(f, v)=0+(e, v)+(e, v)=0.$$

Sei $v \in A_{d,3}$. Nach (5.6) ist $v \in A_{e,3}$ genau dann, wenn $v \in D_f$, und $v \in A_{f,3}$ genau dann, wenn $v \in D_e$. In beiden Fällen folgt $(d+e+f, v)=0$. Sei $v \in A_{e,5} \cap A_{f,5}$. Sei $w \in E_v \cap A_{e,3}$. Da $w \in A_d$, ist $w \notin A_{f,3}$, und $(w, e) \neq (w, f)$. Nach (5.4) gibt es ein $k \in \{\alpha, \alpha^2\}$ mit $(w, e)=(k \cdot v, e)$ und $(w, f)=(k \cdot v, f)$. Demnach ist $(v, e) \neq (v, f)$, und es folgt $(v, e)+(v, f)=1$. Somit ist $(d+e+f, v)=0$.

Sei $v \in A_{d,5}$. Dann gibt es ein $w \in E_v \cap A_{d,3}$. Es ist $(d+e+f, w)=0$, und es gibt nach (5.4) ein $k \in \{\alpha, \alpha^2\}$ mit $(d+e+f, k \cdot v)=0$. Demnach ist auch $(d+e+f, v)=0$.

(5.8) Lemma. *Sei $d \in D$ und $e \in A_{d,3}$. Dann ist $d+e+d^e \in V_0$.*

Beweis. Wieder genügt es zu zeigen, daß $(d+e+d^e, v)=0$ für alle $v \in D$. Ist $v \in \mathfrak{C}_D(d)$, so ist $o(v \cdot e)=o(v \cdot d^e)$, und $(v, e)=(v, d^e)$, da e und d^e unter $\mathfrak{C}_G(v)$ konjugiert sind. Somit ist $(d+e+d^e, v)=0+(e, v)+(e, v)=0$.

Wir können daher annehmen, daß $v \in A_d \cap A_e \cap A_{d^e}$. Sei $v \in A_{d,3}$. Dann ist $v \in A_{e,5} \cap A_{d^e,5}$ nach (2.3). Es folgt $\langle d, e, v\rangle \simeq A_5^*$, und e und d^e sind unter $\mathfrak{C}_G(v)$ nicht konjugiert. Deshalb ist

$$(d, v)+(e, v)+(d^e, v)=1+(e, v)+((e, v)+1)=0.$$

Ein $v \in A_{d,5} \cap A_{e,5} \cap A_{d^e,5}$ gibt es nicht; denn es wäre $\langle d, e, v\rangle \simeq A_5^*$, und darin gibt es keine solche Relation.

(5.9) Lemma. *Sei $d \in D$ und $y \in A_{d,5}$. Dann gibt es ein $a \in D$, so daß $d+y+a \in V_0$.*

Beweis. Es ist $(d, y)=k$ für $k \in \{\alpha, \alpha^2\}$. Sei $c \in (E_d \setminus \{d\}) \cap A_{y,5}$ und $z \in (E_y \setminus \{y\}) \cap A_{d,5}$. Dann gilt $(c, y)=k^2=(d, z)$ und $(c, z)=1$. Wegen $(d, y)=(k^2 \cdot c, y)=(d, k^2 \cdot z)$ folgt aus (5.4), daß $(d, v)=(k^2 \cdot c, v)$ und $(y, v)=(k^2 \cdot z, v)$ für alle $v \in V$ ist. Wiederum nach (5.4) existiert ein $a \in E_{c^z}$, so daß $(a, v)=(k^2 \cdot c^z, v)$ für alle $v \in V$. Es gilt nun für alle $v \in V$

$$\begin{aligned}(d+y+a, v)&=(d, v)+(y, v)+(a, v)=(k^2 \cdot c, v)+(k^2 \cdot z, v)+(k^2 \cdot c^z, v)\\ &=k^2 \cdot ((c, v)+(z, v)+(c^z, v))=k^2 \cdot (c+z+c^z, v).\end{aligned}$$

Nach (5.8) ist $(c+z+c^z, v)=0$; demnach ist auch $(d+y+a, v)=0$ für alle $v \in V$.

Es sei $W=V/V_0$. Dann ist W offenbar ein nicht ausgearteter symplektischer Raum über $GF(4)$. Nach (5.4), (5.5), (5.7), (5.8) und (5.9) läßt sich jedes Element von W, mit Ausnahme der Null, in der Form $d+V_0$ darstellen, wobei $d \in D$. Andererseits gilt $d+e \in V_0$ für $d, e \in D$ nur dann, wenn $d=e$. Deshalb ist $|W|=|D|+1=4^{2n}$ und $\dim W=2n$. Wir können nun die Elemente von $W \setminus \{0\}$ mit den Involutionen aus D identifizieren.

Definition. Sei $g \in G$. Sei g^+ die folgende Abbildung von W auf sich

$$0\,g^+ = 0; \qquad w\,g^+ = w^g \qquad \text{für } w \in D.$$

Es ist g^+ die von g *induzierte* Abbildung von W auf sich.

(5.10) Satz. *Es ist* $G \simeq \mathrm{Sp}(2n, 4)$.

Beweis. Wir werden zeigen, daß die Elemente von D die symplektischen Transvektionen auf W induzieren. Dann ist klar, daß G die volle symplektische Gruppe auf W induziert; denn diese wird ja durch die Transvektionen erzeugt. Da wegen $\mathbb{Z}(G)=1$ ein nichttriviales Element von G niemals die Identität auf W induziert, folgt dann $G \simeq \mathrm{Sp}(2n, 4)$.

Sei $d \in D$ und $w \in W$.

Ist $w = 0$, so gilt $w\,d^+ = w = w + 0 = w + (w, d) \cdot d$.

Ist $w \in \mathbb{C}_D(d)$, so gilt $w\,d^+ = w^d = w = w + 0 = w + (w, d)\,d$.

Ist $w \in A_{d,3}$, so gilt $w\,d^+ = w^d = w + d = w + (w, d)\,d$.

Sei nun $w \in A_{d,5}$ und $(w, d) = k$, wobei $k \in \{\alpha, \alpha^2\}$. Aus dem Beweis von (5.9) folgt, daß $w + k \cdot d = k \cdot (k^2 \cdot w)^d$. Wegen $(k^2 \cdot w)^d \in E_{w^d}$ und $(k \cdot (k^2 \cdot w)^d, d) = k \cdot ((k^2 \cdot w)^d, d) = k = (w^d, d)$ folgt $k \cdot (k^2 \cdot w)^d = w^d$. Es ist somit

$$w\,d^+ = w^d = k \cdot (k^2 \cdot w)^d = w + k \cdot d = w + (w, d)\,d,$$

und jedes $d \in D$ induziert auf W eine symplektische Transvektion.

6. Die orthogonalen Gruppen

In diesem Kapitel gelte stets die

Voraussetzung. *Es sei* $\mathbb{Z}(G) = 1 = \mathbb{O}_2(G)$, *und es gelte* (*O*).

(6.1) Lemma. *Sei* $d \in D$, $e \in D_d$ *und* $x \in A_d$. *Dann ist* $|F(d, e) \cap D_x| \leqq 1$.

Beweis. Es ist $F(e, d) = (F(d, e) \cup \{d\}) \setminus \{e\}$. Angenommen, es ist

$$|F(d, e) \cap D_x| > 1.$$

Dann ist o.B.d.A. $e \in D_x$, und es folgt $1 \leqq |F(e, d) \cap D_x| \leqq 2$, was nicht sein kann.

(6.2) Lemma. *Sei* $d \in D$ *und* $x \in A_d$. *Sei* N *der maximale auflösbare Normalteiler von* $\langle D_d \rangle$. *Dann ist* $\langle D_d \cap D_x \rangle N/N \simeq O^{\pm}(2n-2, 4)$.

Beweis. Sei zunächst $n = 2$. Dann ist $\langle D_d \rangle / N \simeq A_5$, und $\langle D_d \cap D_x \rangle \simeq D_6$ oder D_{10} wegen (6.1). Da $D_6 \simeq O^+(2, 4)$ und $D_{10} \simeq O^-(2, 4)$, gilt die Behauptung in diesem Fall.

Sei nun $n \geqq 3$. Seien $e, f \in D_d \cap D_x$ mit $e \in D_f$, und sei $y \in D_d \cap A_e \cap A_f$. Dann ist $\langle F(d, e), y \rangle = A \simeq A_5^*$. Da $|F(d, e) \cap D_x| = 1$ nach (6.1), existiert ein $z \in (A \cap D) \setminus F(d, e)$, das in D_x liegt. Da f mit ganz $F(d, e)$ vertauschbar ist, gilt $z \in A_f$, und e und f sind in $\langle e, z, f \rangle$ konjugiert. Demnach ist $D_d \cap D_x$ eine Konjugiertenklasse von $\langle D_d \cap D_x \rangle$.

Angenommen, es gilt $D_d \cap D_x \cap \mathbb{C}_D(e) = D_d \cap D_x \cap \mathbb{C}_D(f) = X$. Dann ist $\langle X \rangle N_e/N_e \simeq \mathrm{Sp}(2n-4, 4)$, wobei N_e der maximale auflösbare Normalteiler von $\langle D_e \rangle$ ist, und $\mathbb{C}_{D_e}(X) = \langle F(e, d), x \rangle \simeq A_5^*$. Da $f \in \mathbb{C}_{D_e}(X)$, folgt $f \in F(e, d)$ und somit $f \in F(d, e)$, was wegen (6.1) nicht sein kann. Deshalb gilt (O) auch in $D_d \cap D_x$. Es folgt daher nach Induktion, daß $\langle D_d \cap D_x \rangle / N_1 \simeq O^{\pm}(2m, 4)$ für einen geeigneten Normalteiler N_1. Es ist $m = n-1$, da $\langle D_d \cap D_x \cap \mathbb{C}_D(e) \rangle N_e/N_e \simeq \mathrm{Sp}(2n-4, 4)$. Eine von Transvektionen erzeugte orthogonale Untergruppe einer symplektischen Gruppe über $GF(4)$ ist aber stets natürlich eingebettet; daher folgt $\langle D_d \cap D_x \rangle N/N \simeq O^{\pm}(2n-2, 4)$.

(6.3) Korollar. *Es ist* $|D| = 4^{n-1}(4^n \pm 1)$.

Beweis. Sei $d \in D$ und $x \in A_{d,5}$. Dann ist $|D| = 5 \cdot |\mathbb{C}_D(d)| - 4 \cdot |\mathbb{C}_D(d) \cap \mathbb{C}_D(x)|$. Nach (6.2) und (1.2) ist $|\mathbb{C}_D(d) \cap \mathbb{C}_D(x)| = 4^{n-2}(4^{n-1} \mp 1)$. Es folgt

$$|D| = 5 \cdot 4^{2n-2} - 4^{n-1}(4^{n-1} \mp 1) = 4^{n-1}(4^n \pm 1).$$

Wir werden nun den Graph von D mit dem Graph der Transvektionen einer orthogonalen Gruppe vergleichen. Dazu brauchen wir den Begriff der Isomorphie zweier Graphen.

(6.4) Definition. Seien D_1 und D_2 zwei Mengen von $\{3, 5\}$-Transpositionen zweier Gruppen G_1 und G_2. Der Graph von D_1 ist zum Graph von D_2 genau dann *isomorph*, wenn es eine bijektive Abbildung φ von D_1 auf D_2 gibt, so daß gilt

$$o(d \cdot e) = o(d\varphi \cdot e\varphi)$$

für alle $d, e \in D_1$.

(6.5) Satz. *Der Graph von D ist zu einem Graph der Transvektionen einer $O^{\pm}(2n, 4)$ isomorph.*

Beweis. Sei $d \in D$ und $x \in A_{d,5}$. Da $\langle D_d \rangle / N \simeq \mathrm{Sp}(2n-2, 4)$ ist, entsprechen die Elemente von D_d eineindeutig den Vektoren $\neq 0$ eines symplektischen Raumes $\tilde{V}_{\mathrm{Sp}}(2n-2, 4)$. Sei $\tilde{V} = \{\bar{e} \mid e \in D_d\} \cup \{0\}$, wobei $^-$ die eineindeutige Zuordnung bedeutet. Es sei $V = \tilde{V} \perp \langle \bar{d}, \bar{x} \rangle$ ein $2n$-dimensionaler symplektischer Raum, wobei $\bar{d}, \bar{x}$ zwei weitere Basiselemente sind mit $(\bar{d}, \bar{x}) = \alpha \in GF(4)$. Sei $Q(\bar{e}) = 1$, falls $e \in D_d \cap D_x$, und sei $Q(\bar{d}) = Q(\bar{x}) = 1$. Da $\langle D_d \cap D_x \rangle N/N \simeq O^{\pm}(2n, 4)$, läßt sich Q auf V als quadratische Form fortsetzen. Damit wird V zu einem orthogonalen Raum. Sei T die Menge der orthogonalen Transvektionen von V.

Es sei $Y = \langle d, x \rangle \cap D$. Sei φ folgende Abbildung von D auf T

$$\begin{array}{lll}
d & \to t(\bar{d}), & \\
d^x & \to t(\bar{d} + \alpha \cdot \bar{x}), & \\
d^{xd} & \to t(\alpha \cdot \bar{d} + \alpha \cdot \bar{x}), & \\
x^d & \to t(\alpha \cdot \bar{d} + \bar{x}), & \\
x & \to t(\bar{x}), & \\
e & \to t(\bar{e} + k \cdot \bar{d}), & \text{wobei } k = Q(\bar{e})^2 + 1 \text{ für } e \in D_d, \\
e^y & \to (e\varphi)^{y\varphi} & \text{für } e \in D_d \text{ und } y \in Y.
\end{array}$$

Wegen $e\varphi = t(\bar{e}) = (e\varphi)^{y\varphi}$ für $e \in D_d \cap D_x$ und $y \in Y$ ist die Definition eindeutig; offenbar ist φ bijektiv. Es ist zu zeigen, daß $o(r \cdot s) = o(r\varphi \cdot s\varphi)$ für alle $r, s \in D$ ist.

Nach Definition ist φ ein Graphenisomorphismus von Y auf $\langle t(\bar{d}), t(\bar{x})\rangle \cap T$. Außerdem ist $o(d \cdot e) = 2 = o(d\varphi \cdot e\varphi)$ für alle $e \in D_d$.

Sei $e \in D_d$. Ist $Q(\bar{e}) = 0$, so gilt

$$o(x \cdot e) = 5 = o\big(t(\bar{x}) \cdot t(\bar{e} + \bar{d})\big) = o(x\varphi \cdot e\varphi).$$

Ist $Q(\bar{e}) = 1$, so gilt $o(x \cdot e) = 2 = o\big(t(\bar{x}) \cdot t(\bar{e})\big) = o(x\varphi \cdot e\varphi)$. Ist $Q(\bar{e}) = \alpha$, so ist $o(x \cdot e)$ eindeutig bestimmt.

Angenommen, dies wäre nicht der Fall. Dann gibt es zwei nicht vertauschbare Elemente $a, f \in D_d$ mit $Q(\bar{a}) = \alpha = Q(\bar{f})$ und $o(x \cdot a) = 3$; $o(x \cdot f) = 5$. Sei $F(d, a) = \{a, b, c\}$ und $F(d, f) = \{f, g, h\}$, wobei $b, g \in D_d \cap D_x$. Dann ist $\langle a, b, c, f, g, h\rangle \simeq A_5^*$, und $\bar{a} = \alpha^2 \cdot \bar{b}$, $\bar{c} = \alpha \cdot \bar{b}$, $\bar{f} = \alpha^2 \cdot \bar{g}$, $\bar{h} = \alpha \cdot \bar{g}$. Ist $o(b \cdot h) = 3$, so folgt $o(a \cdot f) = 3$, und $a^f \in (\langle b, g\rangle \cap D) \subseteq D_x$. Demnach ist $3 = o(x \cdot a) = o(x^{a^f} \cdot a^{a^f}) = o(x \cdot f) = 5$, was nicht sein kann. Deshalb ist $o(b \cdot h) = 5$, und ebenso $o(b \cdot f) = 5$. Es folgt $o(b \cdot g) = 3$, und $o(x \cdot a) = o(x^{bg} \cdot a^{bg}) = o(x \cdot f)$; ein Widerspruch zur Annahme.

Gegebenenfalls durch Vertauschung von α und α^2 auf $\tilde{V}$ läßt sich demnach erreichen, daß $o(x \cdot e) = 5$ für alle $e \in D_d$ mit $Q(\bar{e}) = \alpha$. Es ist dann $o(x \cdot e) = 5 = o\big(t(\bar{x}) \cdot t(\bar{e} + \alpha \cdot \bar{d})\big) = o(x\varphi \cdot e\varphi)$. Ist $Q(\bar{e}) = \alpha^2$, so gilt

$$o(x \cdot e) = 3 = o\big(t(\bar{x}) \cdot t(\bar{e} + \alpha^2 \cdot \bar{d})\big) = o(x\varphi \cdot e\varphi).$$

Es ist somit $o(r \cdot s) = o(r\varphi \cdot s\varphi)$ für $r \in \{d, x\}$ und $s \in D_d$.

Sei $e \in D_d \cap A_x$. Ist $Q(\bar{e}) = 0$, so folgt $\langle d, x, e\rangle \simeq B_5$ und $\langle d\varphi, x\varphi, e\varphi\rangle \simeq B_5$. Demnach gilt $o(y \cdot e) = 5 = o(y\varphi \cdot e\varphi)$ für $y \in \{d^x, d^{xd}, x^d\}$. Ist $Q(\bar{e}) \neq 0$, so folgt $\langle d, x, e\rangle \simeq A_5^*$ und $\langle d\varphi, x\varphi, e\varphi\rangle \simeq A_5^*$. Dann ist $o(y \cdot e)$ für $y \in \{d^x, d^{xd}, x^d\}$ durch $o(x \cdot e)$ eindeutig bestimmt, und wegen $o(x \cdot e) = o(x\varphi \cdot e\varphi)$ gilt $o(y \cdot e) = o(y\varphi \cdot e\varphi)$ für $y \in \{d^x, d^{xd}, x^d\}$. Demnach gilt $o(r \cdot s) = o(r\varphi \cdot s\varphi)$ für $r \in Y$ und $s \in D_d \cup Y$.

Sei $e \in D_d$ und $r, y \in Y$. Dann ist

$$o(r \cdot e^y) = o(r^y \cdot e) = o\big((r^y)\varphi \cdot e\varphi\big) = o\big((r\varphi)^{y\varphi} \cdot e\varphi\big) = o\big(r\varphi \cdot (e\varphi)^{y\varphi}\big) = o\big(r\varphi \cdot (e^y)\varphi\big).$$

Es gilt also

$$o(r \cdot s) = o(r\varphi \cdot s\varphi) \qquad \text{für } r \in Y \text{ und } s \in D. \tag{$*$}$$

Es bleibt zu zeigen, daß $o(e^y \cdot f^z) = o(e^y\varphi \cdot f^z\varphi)$ für alle $e, f \in D_d$ und $y, z \in Y$ ist. Ist $y = z$, so gilt $o(e^y \cdot f^z) = o(e \cdot f)$, und $o(e \cdot f)$ ist durch $(\bar{e}, \bar{f})$ eindeutig bestimmt. Da $(\bar{e}, \bar{f}) = (\bar{e} + k_1 \cdot \bar{d}, \bar{f} + k_2 \cdot \bar{d})$ für $k_1, k_2 \in GF(4)$, ist

$$o(e \cdot f) = o(e\varphi \cdot f\varphi) = o\big((e\varphi)^{y\varphi} \cdot (f\varphi)^{z\varphi}\big) = o(e^y\varphi \cdot f^z\varphi).$$

Wir können daher annehmen, daß $y \neq z$. Ebenso können wir ausschließen, daß $e \in D_x$ oder $f \in D_x$. Sonst gilt nämlich $e^y = e = e^z$ oder $f^z = f = f^y$.

Es gilt $o(e^y \cdot f^z) = o(e^{yzd} \cdot f) = o(e^w \cdot f)$ mit $w = yzd \in Y$. Da $y \neq z$, ist $w \neq d$. Wegen (*) ist

$$\langle e, w, f \rangle \simeq \langle e\varphi, w\varphi, f\varphi \rangle$$

und

$$o(e^w \cdot f) = o\big((e\varphi)^{w\varphi} \cdot f\varphi\big) = o\big((e\varphi)^{(y\varphi)(z\varphi)(d\varphi)} \cdot f\varphi\big) = o\big((e\varphi)^{y\varphi} \cdot (f\varphi)^{z\varphi}\big)$$
$$= o(e^y\varphi \cdot f^z\varphi).$$

(6.6) Satz. *Es ist* $G \simeq O^{\pm}(2n, 4)$.

Beweis. Nach (6.5) gibt es einen Isomorphismus φ zwischen dem Graphen von D und dem Graphen der Transvektionen einer $O^{\pm}(2n, 4)$. Sei $d \in D$ und $e \in \mathbf{C}_D(d)$. Dann gilt $e^d\varphi = e\varphi = (e\varphi)^{d\varphi}$. Sei $f \in A_{d,3}$. Offenbar werden die Involutionen einer D_6 oder D_{10} wiederum in eine Diedergruppe abgebildet. Deshalb ist $f^d\varphi = (f\varphi)^{d\varphi}$. Sei $x \in A_{d,5}$ und $g \in A_{x,3} \cap D_d$. Wegen $(\langle d, x \rangle \cap D)\varphi = \langle d\varphi, x\varphi \rangle \cap D\varphi$ gilt $x^d\varphi \in \{(x\varphi)^{d\varphi}, (d\varphi)^{x\varphi}, (d\varphi)^{(x\varphi)(d\varphi)}\}$. Da $o(g \cdot x^d) = 3 = o(g\varphi \cdot x^d\varphi)$, muß gelten $x^d\varphi = (x\varphi)^{d\varphi}$. Es ist also $x^d\varphi = (x\varphi)^{d\varphi}$ für alle $x \in D$, und d operiert auf D genauso wie die Transvektion $d\varphi$ auf $D\varphi$. Wegen $\mathbb{Z}(G) = 1$ ist demnach $G \simeq O^{\pm}(2n, 4)$.

Literatur

1. Artin, E.: The orders of the classical simple groups. Commun. Pure Appl. Math. 8, 455–472 (1955).
2. Coxeter, H., Moser, W.: Generators and relations for discrete groups. Ergebnisse der Math. 14, 2. edit. Berlin-Göttingen-Heidelberg-New York: Springer 1964.
3. Dieudonné, J.: Sur les groupes classiques. Act. Sci. Industr. No. 1040, Paris 1948.
4. Fischer, B.: Finite groups generated by 3-transpositions. Erscheint demnächst in Inventiones math.
5. Huppert, B.: Endliche Gruppen I. Berlin-Heidelberg-New York: Springer 1967.
6. Todd, J., Coxeter, H.: A practical method for enumerating cosets of a finite abstract group. Proc. Edinburgh Math. Soc. 5(2), 26–34 (1936).

Dr. K.-J. Fleischer
Fakultät für Mathematik
Universität Bielefeld
D-4800 Bielefeld
Kurt-Schumacher-Str. 6

(Eingegangen am 31. März 1971)

Math. Z. 124, 119–132 (1972)

Intersections of Primary Powers of a Group

DEREK J. S. ROBINSON*

1. Introduction

If G is a group and m is a non-negative integer, G^m denotes *the m-th power of G*, that is, the subgroup of G generated by all g^m with g in G. We shall be interested in situations where the intersection of an infinite set of primary powers of G is the identity subgroup. For example, G might be *residually of finite p-exponent* for some prime p, that is to say

$$\bigcap_{i=0,1,2,\ldots} G^{p^i} = 1;$$

in fact we shall study groups which have this property for infinitely many primes p.

Let i be a function from the set of all primes to the set of all non-negative integers. A group G is said to have the property

$$\mathscr{I}(i)$$

if

$$\bigcap_p G^{p^{i(p)}} = 1, \tag{1}$$

the intersection being formed over all primes p: such properties will be referred to as *$\mathscr{I}$-properties*. Of course, $\mathscr{I}(i)$ is interesting only if i has infinite support: for if i has finite support, a group G has $\mathscr{I}(i)$ if and only if it is a direct product of p-groups of exponent dividing $p^{i(p)}$, one for each prime p.

At this point let us recall some known results. If G is a finitely generated torsion-free nilpotent group, then G is residually a finite p-group for every prime p (Gruenberg [5], Theorem 2.1), and G has the property $\mathscr{I}(i)$ where $i(p)=1$ for all p: the last result simply says that

$$\bigcap_p G^p = 1$$

and is due to Higman ([8])—see also Baer [1], p. 301. On the other hand, Seksenbaev ([15]) has proved that a polycyclic group which is residually a finite p-group for infinitely many primes p is nilpotent (see also [13], p. 170). It is not hard to adapt Seksenbaev's argument to show that *a polycyclic group with an $\mathscr{I}$-property is nilpotent.*

Here we shall furnish generalizations of these theorems, aiming to show that a group which is residually of finite p-exponent for infinitely many primes p

* The author acknowledges support from the National Science Foundation.

or which has an $\mathscr{I}$-property enjoys some degree of nilpotence. Now Gruenberg ([5]) has proved that a free polynilpotent group is residually a finite p-group for all primes p, and it will be shown that such groups also have the property $\mathscr{I}(i)$ where $i(p)=1$ for all p (Theorem 3). Thus it is necessary to restrict our attention to particular classes of groups. Here we choose to study groups which have finite *torsion-free rank* (cf. Zassenhaus [18], pp. 241–242): a group G has finite torsion-free rank r if there is a series

$$1=G_0<G_1<\cdots<G_n=G$$

such that G_i is normal in G_{i+1} for each i and exactly r of the factors G_{i+1}/G_i are infinite cyclic while the remaining $n-r$ are periodic. That r is an invariant of G follows from the Schreier Refinement Theorem.

For example, Theorem 1 and Corollary 3 combine to show that a *group with finite torsion-free rank is residually of finite p-exponent for an infinity of primes p if and only if it is a torsion-free nilpotent group whose centre is without elements of infinite p-height for all relevant p.* This implies the theorems of Gruenberg and Seksenbaev. According to Theorem 2, *if G is a group with finite torsion-free rank which has an $\mathscr{I}$-property, the elements of G with finite order form a subgroup T such that G/T is nilpotent.* A converse to this is provided by Theorem 4. Some of our results are valid for a class of groups $\mathfrak{X}$ which is significantly wider than the class of groups with finite torsion-free rank: this is defined in Section 2.

Notation

$H \lhd K$: H is a normal subgroup of K,
$\langle X_\lambda : \lambda \in \Lambda \rangle$: subgroup generated by the subsets X_λ,
$[x, {}_{i+1}y]=[[x, {}_iy], y]$
$[x, {}_1y]=[x, y]=x^{-1}y^{-1}xy$
$[H, {}_{i+1}K]=[[H, {}_iK], K]$
$[H, {}_1K]=[H, K]=\langle [h, k]: h\in H, k\in K\rangle$,
$\gamma_n(G)$: n-th term of the lower central series of G,
$H^K=\langle h^k=k^{-1}hk: k\in K, h\in H\rangle$, the normal closure of H in K,
$C_G(N)$: centralizer of the G-operator group N in G,
$\mathrm{Aut}_G(N)$: automorphism group of the G-operator group N induced by G in the natural way,
π': the complement of the set of primes π.

2. Groups with Finite Torsion-Free Rank and the Class $\mathfrak{X}$

If $\mathscr{P}$ is a group theoretical property, *a hyper-$\mathscr{P}$ group* is a group which has an ascending series of normal subgroups the factors of which all have $\mathscr{P}$. When $\mathscr{P}$ is inherited by homomorphic images, hyper-$\mathscr{P}$ is equivalent to the requirement that non-trivial homomorphic images possess non-trivial normal subgroups with $\mathscr{P}$.

We define

$$\mathfrak{X}$$

to be the class of groups that are hyper-(abelian of finite rank or periodic). Here a group is said to have *finite rank* r if each finitely generated subgroup can be generated by r elements and r is the least such integer. Observe that $\mathfrak{X}$ and the class of groups with finite torsion-free rank are closed with respect to forming subgroups and homomorphic images.

Lemma 1. *A locally nilpotent group belonging to the class $\mathfrak{X}$ is an extension of a periodic group by a hypercentral (or ZA-)group.*

This is essentially due to Čarin ([2], Theorem 2). Here is a brief proof.

Proof. Let G be a non-trivial torsion-free locally nilpotent group and let $G\in\mathfrak{X}$. Let us show that the centre of G is non-trivial. By hypothesis there is a non-trivial normal subgroup A which is abelian and has finite rank, say r. Let $a\in A$ and let $g_i\in G$, $i=1, \ldots, n$; then $H=\langle a, g_1, \ldots, g_n\rangle$ is nilpotent of class c, say, Set $B=a^H$; then $B\leqq A$ and $[B, {}_cH]=1$. Now the action of H can be extended from B to a rational vector space of dimension r and this affords a representation of H by a unipotent group of linear transformations. Hence $[B, {}_rH]=1$, and this implies that A lies in the r-th term of the upper central series of G. It follows that the centre of G is not 1.

To complete the proof it is sufficient to show that G/Z_α is torsion-free where Z_α is the α-th term of the upper central series of G. Suppose that α is the first ordinal for which this is false and observe that α cannot be a limit ordinal. Let $x^m\in Z_\alpha$ where $m>0$, and let g be any element of G. Then $K=\langle x, g\rangle$ is nilpotent and $KZ_{\alpha-1}/Z_{\alpha-1}$ is torsion-free. Now $(x\,Z_{\alpha-1})^m$ lies in the centre of $KZ_{\alpha-1}/Z_{\alpha-1}$; by a well known theorem of Mal'cev ([10]) the element $x\,Z_{\alpha-1}$ must belong to the centre of $KZ_{\alpha-1}/Z_{\alpha-1}$ for all g. Hence $x\in Z_\alpha$ and G/Z_α is torsion-free.

Corollary 1. *A group G with finite torsion-free rank belongs to the class $\mathfrak{X}$.*

Proof. Let $1=G_0<G_1<\cdots<G_n=G$ be a series in G with infinite cyclic or periodic factors. If $n\leqq 1$, it is obvious that $G\in\mathfrak{X}$. Let $n>1$ and write $H=G_1$. Suppose that G contains no normal periodic subgroups other than 1. If H is periodic, then so is H^G since H is subnormal in G (see [18], p. 246). Thus H is infinite cyclic and H^G is a Baer group; hence H^G is a locally nilpotent. Now $H^G\leqq G_{n-1}$ and $G_{n-1}\in\mathfrak{X}$ by induction on n. Thus H^G is hypercentral by the lemma. The centre of H^G is a non-trivial normal abelian subgroup of G with finite rank (since it has finite torsion-free rank). This implies that $G\in\mathfrak{X}$.

Corollary 2. *A locally nilpotent group with finite torsion-free rank is an extension of a periodic group by a nilpotent group.*

Proof. Lemma 1 and Corollary 1 permit us to deal only with a torsion-free hypercentral group G with finite torsion-free rank. Let A be a maximal normal abelian subgroup of G. Then $A=C_G(A)$ and A has finite rank $r>0$, say. As in the proof of Lemma 1 we find that A lies in the r-th term of the upper central series of G. Since G/A is isomorphic with a group of unipotent $r\times r$ matrices, G is nilpotent of class at most $r+(r-1)=2r-1$.

Corollary 2 is also a consequence of a more general theorem of Mal'cev ([11], Theorem 5).

3. Groups that are Residually of Finite p-Exponent for an Infinity of Primes p

As our first main result we shall prove

Theorem 1. *Let G be a group and let π be an infinite set of primes. Assume that G is residually of finite p-exponent for all $p \in \pi$.*

(i) *If $G \in \mathfrak{X}$, then G is torsion-free and hypercentral.*

(ii) *If G has finite torsion-free rank, then G is torsion-free and nilpotent.*

Proof. (a) Corollary 2 shows that it is sufficient to establish (i): assume therefore that $G \in \mathfrak{X}$. Now G is obviously torsion-free, so by Lemma 1 we need only prove that G is locally nilpotent.

By hypothesis there is an ascending series of normal subgroups $\{G_\beta : \beta \leqq \alpha\}$ in which each $G_{\beta+1}/G_\beta$ is periodic or abelian of finite rank. We use transfinite induction on α. Thus we may assume that α is not a limit ordinal and that $N = G_{\alpha-1}$ is locally nilpotent. From now on G will be supposed finitely generated, with say

$$G = \langle x_1, \dots, x_n \rangle \tag{2}$$

(b) First of all consider the case G/N periodic. Then $x_i^{m_i} \in N$ for some positive integer m_i and the $x_i^{m_i}$ generate a nilpotent subgroup X of N. Let π_1 be the set of all primes in π which do not divide $m_1 m_2 \dots m_n$: certainly π_1 is nonempty. For any $j \geqq 0$ and $p \in \pi_1$ we have $x_i^{m_i} \in X G^{p^j}$, which, by definition of π_1, implies that $x_i \in X G^{p^j}$. This is true for all i, so $G = X G^{p^j}$. If c is the nilpotent class of X, then $\gamma_{c+1}(G) \leqq G^{p^j}$ and

$$\gamma_{c+1}(G) \leqq \bigcap_{j=0,1,2,\dots} G^{p^j} = 1 .$$

Thus is G nilpotent.

(c) Therefore we can assume that G/N is abelian. N is locally nilpotent and therefore hypercentral: write Z_β for the β-th term of the upper central series of N. Let $a \in N$ and $g \in G$: our immediate aim is to prove that $\langle a, g \rangle$ is nilpotent. If this is false, $a \neq 1$, so $a \in Z_{\alpha+1} \setminus Z_\alpha$ for some ordinal α. Choose a so that α is a minimal subject to the non-nilpotence of some $\langle a, g \rangle$.

Clearly $Z_{\alpha+1} \lhd G$ and $a^{\langle g \rangle} \leqq Z_{\alpha+1}$, so that $H = \langle a, g \rangle Z_\alpha / Z_\alpha$ is finitely generated and metabelian. By a theorem of P. Hall ([7], Theorem 3), H satisfies the maximal condition on normal subgroups. Since $H \in \mathfrak{X}$, we see that H has finite torsion-free rank. Now $Z_{\alpha+1}/Z_\alpha$ is torsion-free because N is — this is essentially due to Mal'cev ([10]) but see also McLain [9]. Therefore $a^{\langle g \rangle} Z_\alpha / Z_\alpha$ is a torsion-free abelian group with finite rank r, say. Write

$$a_i = a^{g^i} .$$

The elements $a_0, a_1, \dots, a_r$ must be linearly dependent modulo Z_α; this implies that there are positive integers s and t such that a_{-1}^s and a_r^t belong to a subgroup of the form

$$Y = \langle a_0, a_1, \dots, a_{r-1}, b^{\langle g \rangle}, c^{\langle g \rangle} \rangle$$

where b and c belong to Z_α: to achieve this we need to transform by powers of g. By minimality of α the subgroups $\langle b, g\rangle$ and $\langle c, g\rangle$ are nilpotent: consequently $b^{\langle g\rangle}$ and $c^{\langle g\rangle}$ are finitely generated and Y is finitely generated. Since $Y \leqq N$, the subgroup Y is nilpotent, of class d let us say. Define π_2 to be the set of all primes in π which do not divide st; then π_2 is not empty. For any $j \geqq 0$ and $p \in \pi_2$ the elements a_{-1}^s and a_r^t belong to YG^{p^j}. The choice of p leads to a_{-1} and a_r belonging to YG^{p^j}. From this one finds through transformation by powers of g that every a_i belongs to YG^{p^j}. Hence $a^{\langle g\rangle} \leqq YG^{p^j}$. Thus $\gamma_{d+1}(a^{\langle g\rangle}) \leqq G^{p^j}$ and

$$\gamma_{d+1}(a^{\langle g\rangle}) \leqq \bigcap_{j=0,1,2,\ldots} G^{p^j} = 1.$$

Hence $A = a^{\langle g\rangle}$ is nilpotent and the group $K = \langle a, g\rangle$ is soluble because $a^K = A$. The group K/A' satisfies the maximal condition on normal subgroup by P. Hall's theorem and this leads to A/A' having finite torsion-free rank. Now A is nilpotent and one easily verifies that a tensor product of two abelian groups with finite torsion-free rank itself has finite torsion-free rank. By Corollary 1 of [14] the group A, and hence $K = \langle g, A\rangle$, has finite torsion-free rank.

It follows that $\gamma_i(K)/\gamma_{i+1}(K)$ is periodic for some integer i: let $L = \gamma_i(K)$ and note that $L \neq 1$ because K is not nilpotent. For any $p \in \pi$ and any $j \geqq 0$ the group K/K^{p^j} is a finite p-group, for it is a finitely generated soluble p-group: thus K/K^{p^j} is nilpotent. It follows that there is a descending central series of K

$$K = K_1 > K_2 > \cdots$$

such that $|K_j : K_{j+1}| = p$ for all j and the intersection of the K_j is 1. Since $L \neq 1$, there is a least integer t such that $L \not\leqq K_t$. Naturally $t > 1$ and $L \leqq K_{t-1}$. Since $|K_{t-1} : K_t| = p$, it follows that $K_{t-1} = LK_t$ and

$$|L : L \cap K_t| = |K_{t-1} : K_t| = p.$$

Now

$$[L, K] \leqq L \cap [K_{t-1}, K] \leqq L \cap K_t.$$

Therefore $L/[L, K]$ has a subgroup of index p for all $p \in \pi$. But $K/[L, K]$ is a finitely generated nilpotent group: hence $L/[L, K]$ cannot be periodic, which contradicts our choice of L. (This is the central argument of Seksenbaev's theorem.)

(d) Finally let us complete the proof of the nilpotence of G. Since G/N is abelian, G will be hypercentral — and hence nilpotent — if each Z_α lies in the hypercentre of G. If this is not true, let α be the first offending ordinal: then α cannot be a limit ordinal. Let $a \in Z_\alpha$. Now $G' \leqq N$, so by (2)

$$G = \langle x_1, \ldots, x_{n-1}\rangle \langle x_n\rangle N$$

and

$$B = a^G \equiv (a^{\langle x_1, \ldots, x_{n-1}\rangle})^{\langle x_n\rangle} \bmod Z_{\alpha-1}.$$

Since by (c) every $a^{\langle g\rangle}$ is finitely generated, induction on n shows that $BZ_{\alpha-1}/Z_{\alpha-1}$ is finitely generated; there is no loss in supposing this group to be either free abelian or an elementary abelian p-group. Let u be the rank of $BZ_{\alpha-1}/Z_{\alpha-1}$;

then $[b, {}_u g] \in Z_{\alpha-1}$ for all $b \in B$ and $g \in G$ by (c). From this we obtain $[B, {}_u G] \leqq Z_{\alpha-1}$ by a simple induction on u, using $G' \leqq N$: alternatively, apply Wedderburn's theorem on algebras with a basis of nilpotent elements ([17]). Thus $BZ_{\alpha-1}/Z_{\alpha-1}$ lies on the hypercentre of $G/Z_{\alpha-1}$ and Z_α lies in the hypercentre of G, which is not so.

Sufficient Conditions

Any group which is to be residually of finite p-exponent for an infinity of primes p must, of course, be torsion-free. Clearly an abelian group possesses the former property if and only if it contains no elements of infinite p-height for any of the relevant primes p. If we widen our enquiry to nilpotent groups, the following result is decisive.

Lemma 2. *Let G be a nilpotent group of positive class c and let I be a set of positive integers. Assume that Z_1, the centre of G, satisfies*

$$\bigcap_{m \in I} (Z_1)^m = 1 .$$

(i) *Every upper central factor of G has this property.*

(ii) *If in addition $a \in Z_1$, $m \in I$ and $a^{m^t} = 1$ always imply that $a^m = 1$, then*

$$\bigcap_{m \in I} G^{m^{c^*}} = 1$$

where

$$c^* = 1 + (c-1) + (c-1)(c-2) + \cdots + (c-1)(c-2)\ldots 2.1 .$$

Proof. Let Z_i denote the i-th term of the upper central series of G and write D/Z_1 for the intersection of all the $(Z_2/Z_1)^m$ with $m \in I$. To prove (i) it is enough to show that $D = Z_1$. If $g \in G$, the map ϕ_g sending $a Z_1$ to $[a, g]$ is a homomorphism of Z_2/Z_1 into Z_1. Therefore

$$(D/Z_1)^{\phi_g} \leqq \bigcap_{m \in I} ((Z_2/Z_1)^{\phi_g})^m \leqq \bigcap_{m \in I} Z_1^m = 1 .$$

Consequently $[D, g] = 1$ for all $g \in G$, and $D = Z_1$.

Assume now that G satisfies the condition specified in (ii). First observe that this condition is also valid in the second, and hence in every, upper central factor of G. For let $a \in Z_2$ and $m \in I$ and assume that $a^{m^t} \in Z_1$; then for any $g \in G$

$$1 = [a^{m^t}, g] = [a, g]^{m^t},$$

since $[a, g] \in Z_1$. Thus

$$1 = [a, g]^m = [a^m, g]$$

and $a^m \in Z_1$.

Write $M = Z_r$ and $N = Z_{r+1}$ where $0 < r < c$, and assume that there is a power m' of m such that

$$\bigcap_{m \in I} M^{m'} = 1 .$$

We shall prove that

$$\bigcap_{m \in I} N^{m''} = 1 \quad \text{where} \quad m'' = m(m')^r . \tag{3}$$

Suppose that $x \neq 1$ belongs to $N^{m''}$ for all $m \in I$. Since m divides m'' and by (i) the intersection of all the $(N/M)^m$ is trivial, $x \in M$. Hence $x \notin M^{m'}$ for some $m \in I$. Let

$$C = C_N(M/M^{m'}) \quad \text{and} \quad L = C^{mm'} M^{m'}.$$

Clearly $C \lhd N$ and $L \lhd N$. Now $[M, {}_r N] = 1$ since $M = Z_r$: from this it is straightforward to deduce that N/C has exponent dividing $(m')^{r-1}$ (see [12], Lemma 4.12). Therefore N/L has exponent dividing $(m m')(m')^{r-1} = m''$ and $x \in N^{m''} \leqq L$. Now N/M is abelian, so

$$[C', C] \leqq [M, C] \leqq M^{m'}$$

and $C/M^{m'}$ has nilpotent class at most 2. Use of the identity

$$(a b)^d = a^d b^d [b, a]^{\binom{d}{2}},$$

which is valid in any nilpotent group of class at most 2, leads to

$$x \equiv y^{mm'} \bmod (C')^{\binom{mm'}{2}} M^{m'}$$

for some y in C. Now $C' \leqq M$ and m' divides $\binom{m m'}{2}$ because m' is a power of m. Therefore

$$x \equiv y^{mm'} \bmod M^{m'}.$$

Since $x \in M$, the element yM of N/M has order dividing $m m'$, a power of m. Hence $y^m \in M$ and consequently $x \in (M)^{m'}$, a contradiction. The required result can now be obtained by repeated use of (3).

Corollary 3. *A nilpotent group is residually of finite p-exponent for an infinity of primes p if and only if its centre has no elements of infinite p-height for all relevant p and is torsion-free.*

Theorem 1 and Corollary 3 provide necessary and sufficient conditions for a group with finite torsion-free rank to be residually of finite p-exponent for an infinite set of primes p. Whether there is a criterion similar to Corollary 3 valid for hypercentral groups seems unclear.

4. Groups with an $\mathscr{I}$-Property

From now on i will denote a function from the set of primes to the set of non-negative integers and $\mathscr{I}(i)$ will be the property defined by Eq. (1): this property is inherited by subgroups but not by homomorphic images (unless i has finite support).

It is obvious that $\mathscr{I}(i)$ and $\mathscr{I}(i')$ coincide if and only if $i = i'$. For torsion-free groups, however, the situation is different: let G be a torsion-free group with $\mathscr{I}(i)$ and let p be any prime. Writing D for the intersection of all the $G^{q^{i(q)}}$ with $q \neq p$, we find that $D \cap G^{p^{i(p)}} = 1$ and $D^{p^{i(p)}} = 1$. Hence $D = 1$ and G has the stronger property $\mathscr{I}(i_1)$ where $i_1(q) = i(q)$ if $q \neq p$ and $i_1(p) = 0$. Thus $\mathscr{I}(i)$ and $\mathscr{I}(i')$ coincide for torsion-free groups if $i(p) = i'(p)$ for all but a finite number of p, that is, if i and i' belong to the same *type* (see [4], § 42 for terminology).

Now let A be a torsion-free abelian group with rank 1. Then it is easy to see that A has $\mathscr{I}(i)$ if and only if $T(A) \ngeqq \bar{i}$ where $T(A)$ is the type of A and $\bar{i}$ is the type containing i. Moreover there exist groups of arbitrary type.

The results of the last two paragraphs are summarised in

Lemma 3. (i) *The properties $\mathscr{I}(i)$ and $\mathscr{I}(i')$ coincide if and only if $i=i'$.*

(ii) *The properties $\mathscr{I}(i)$ and $\mathscr{I}(i')$ coincide for torsion free groups if and only if i and i' are of the same type.*

The next lemma follows easily from the fact that a group G with $\mathscr{I}(i)$ can be embedded in the cartesian product of the groups $G/G^{p^{i(p)}}$.

Lemma 4. *Let G be a group with $\mathscr{I}(i)$: then*

(i) *the elements of G with finite orders form a subgroup T which is the direct product of its Sylow subgroups;*

(ii) *if T_p is the Sylow p-subgroup of T, then T_p and* $\mathrm{Aut}_G(T_p)$ have *finite exponent dividing $p^{i(p)}$.*

Our first main result about the property $\mathscr{I}(i)$ is

Theorem 2. *Let G be a group with $\mathscr{I}(i)$ and let T be the subgroup of all elements of G with finite order.*

(i) *If $G \in \mathfrak{X}$, then G/T is torsion-free and hypercentral.*

(ii) *If $G \in \mathfrak{X}$ and T is locally finite, then G is locally nilpotent.*

(iii) *If G has finite torsion-free rank, then G/T is torsion-free and nilpotent with finite rank.*

Corollary 4. *Let G have an $\mathscr{I}$-property. If G is a soluble group with finite rank, it is hypercentral. If G is polycyclic, it is nilpotent.*

Of these statements the second is an immediate consequence of Theorem 2, while the first follows with the aid of arguments used in proving Lemma 1.

Proof of Theorem 2. Assume that $G \in \mathfrak{X}$: we must show that G/T is hypercentral and this will be the case if it is locally nilpotent, by Lemma 1. We will also assume that G can be generated by finitely many elements $x_1, \ldots, x_n$.

By hypothesis there is an ascending series of normal subgroups in G each of whose factors is either a periodic group or is abelian of finite rank. Arguing by transfinite induction on the length of this series, we find a normal subgroup N such that NT/T is locally nilpotent and G/NT is either periodic or abelian. Evidently there is nothing to be lost in assuming that $T \leqq N$.

Let us make the following observation once and for all. Suppose that π is a finite set of primes and P is the direct product of the Sylow p-subgroups of T for $p \in \pi$: *then G/P has $\mathscr{I}(i)$.* For if D/P is the intersection of all the $(G/P)^{p^{i(p)}}$ for $p \notin \pi$, then $D \leqq G^{p^{i(p)}} P = G^{p^{i(p)}}$ if $p \notin \pi$. Therefore

$$D \cap \Big(\bigcap_{p \in \pi} G^{p^{i(p)}}\Big) = 1,$$

which shows that D has finite exponent dividing $\prod_{p \in \pi} p^{i(p)}$. Thus $D = P$.

Suppose now that G/N is periodic. Then there are positive integers m_i such that $X=\langle x_1^{m_1}, \ldots, x_n^{m_n}\rangle \leqq N$. Now XT/T is finitely generated and nilpotent, so $X/X \cap T$ is finitely presented. Thus $X \cap T$ is the normal closure in X of a finite subset (for these facts see [7], p. 421). This implies that X contains elements of only finitely many distinct prime orders. Let π denote the set of all such primes and all prime divisors of $m_1 m_2 \ldots m_n$. If P is the direct product of the Sylow p-subgroups of T for $p \in \pi$, then G/P has $\mathscr{I}(i)$. There is, therefore, no loss in supposing that $P=1$. Hence X is nilpotent, say of class c.

Let $p \notin \pi$: then $x_i^{m_i} \in XG^{p^{i(p)}}$ implies that $x_i \in XG^{p^{i(p)}}$, so $G=XG^{p^{i(p)}}$. Hence $\gamma_{c+1}(G) \leqq G^{p^{i(p)}}$ for all $p \notin \pi$. It follows that $\gamma_{c+1}(G)$ has exponent equal to a π-number. Hence $\gamma_{c+1}(G) \leqq P=1$ and G is nilpotent.

Assume from now on that G/N is abelian. The next step is to prove that $\langle a, g\rangle T/T$ is nilpotent if $a \in N$ and $g \in G$. The proof is similar to part (c) of the proof of Theorem 1: moreover the hypercentrality of G/T follows as in part (d) of the latter proof. Suppose the assertion is false and let $a \in Z_{\alpha+1} \setminus Z_\alpha$ where Z_α/T is the α-th term of the upper central series of N/T. Choose a so that α is minimal subject to $\langle a, g\rangle T/T$ being non-nilpotent for some g. Writing $a_i=a^{g^i}$, we can find positive integers r, s and t and elements b and c of Z_α such that a_{-1}^s and a_r^t belong to

$$Y=\langle a_0, a_1, \ldots, a_{r-1}, b^{\langle g\rangle}, c^{\langle g\rangle}\rangle .$$

By minimality of α both $\langle b, g\rangle T/T$ and $\langle c, g\rangle T/T$ are finitely generated nilpotent groups, whence $\langle b, g\rangle$ and $\langle c, g\rangle$ have elements of only finitely many different prime orders. We may therefore suppose that $\langle b, g\rangle$ and $\langle c, g\rangle$ are torsion-free, which implies that they are nilpotent. Hence Y is finitely generated and can contain elements of only finitely many distinct prime orders. If π is the set of all such primes and all prime divisors of st, we can assume that G contains no non-trivial element with order a π-number. One then argues that a_{-1} and a_r must belong to $YG^{p^{i(p)}}$ and hence that $a^{\langle g\rangle} \leqq YP^{p^{i(p)}}$ if $p \notin \pi$. This leads to $\gamma_{d+1}(a^{\langle g\rangle})$ having exponent equal to a π-number, where d is the nilpotent class of Y, and hence to the nilpotency of $a^{\langle g\rangle}$. The group $\langle a, g\rangle$ is therefore soluble, and it has finite torsion-free rank by the argument of the proof of Theorem 1.

The nilpotence of $\langle a, g\rangle$ is now a consequence of the following result. *If K is a finitely generated group with finite torsion-free rank, if K has $\mathscr{I}(i)$ and if each $K/K^{p^{i(p)}}$ is finite, then K is nilpotent.* We shall suppose that this assertion is false.

Since $K/K^{p^{i(p)}}$ is a finite p-group and K has $\mathscr{I}(i)$, there is an infinite descending central series of K

$$K=K_1>K_2>\cdots$$

such that the intersection of all the K_j is 1, each $|K_i:K_{i+1}|$ is prime and a prime p can occur only finitely often as the order of a factor of the series. Since K has finite torsion-free rank, all but a finite number of the lower central factors of K must be periodic. Suppose that $\gamma_j(K)/\gamma_{j+1}(K)$ is periodic and set $L=\gamma_j(K)$.

Then $L/[L, K]$ is finite since it is a periodic subgroup of the finitely generated nilpotent group $K/[L, K]$. Define

$$L_t = L \cap K_t$$

and observe that $|L_t : L_{t+1}| = 1$ or $|K_t : K_{k+1}|$. Now K/L and K/K_t are nilpotent, so $L_t \neq 1$ for all t. Hence there occur among the $|L_t : L_{t+1}|$ an infinite number of primes. Choose one such prime p which does not divide $|L : [L, K]|$ and suppose that $p = |L_t : L_{t+1}|$, with t as small as possible. L/L_{t+1} is a finite nilpotent group with Sylow p-subgroup of order p. Hence there exists a normal subgroup M of K such that $L_{t+1} \leqq M < L$ and L/M has order p. Then L/M is a minimal normal subgroup of K/M, and K/M is nilpotent because $L_{t+1} \leqq M$. Hence L/M lies in the centre of K/M and $[L, K] \leqq M < L$. Thus $L/[L, K]$ has a subgroup of index p: but this is impossible by our choice of p.

So far (i) has been established. The result of the previous paragraph, together with (i), implies (ii), while (iii) follows from (i) and Corollary 2. This completes the proof.

Remark. The group G/T need not have an $\mathscr{I}$-property nor need G be nilpotent, even if G is metabelian and of torsion-free rank 1. For example, let H_p be a finite metabelian p-group with exponent p and nilpotent class $p-1$; here H_p could be the semi-direct product of an elementary abelian p-group with basis $\{a_1, \ldots, a_{p-1}\}$ by the cyclic group $\langle t \rangle$ of order p where

$$t^{-1} a_{p-1} t = a_{p-1} \quad \text{and} \quad t^{-1} a_i t = a_i a_{i+1}, \qquad 1 \leqq i < p-1.$$

Let H be the cartesian product of the groups H_p for all primes p. The subgroup T of elements of finite order in H is just the direct product of the H_p. Clearly

$$\bigcap_p H^p = 1.$$

It is easy to see that H/T is radicable; in fact H/T contains a subgroup G/T which is isomorphic with the additive group of rational numbers. G is a group with the properties mentioned.

Free Polynilpotent Groups

To indicate that the hypothesis "$G \in \mathfrak{X}$" cannot be omitted from Theorem 2 we shall prove the following theorem, which is doubtless well-known.

Theorem 3. *If G is a subgroup of a free polynilpotent group and π is any infinite set of primes, then*

$$\bigcap_{p \in \pi} G^p = 1.$$

Proof. Evidently we may assume that G is a finitely generated free polynilpotent group. By Gruenberg's theorem[1] ([5], Theorem 7.1), G is residually a finite p-group and hence is residually nilpotent. Now a theorem of P. Hall

[1] A simple proof of Gruenberg's theorem is provided by work of Dunwoody [3].

(see also Šmel'kin [16]) asserts that each $\gamma_j(G)/\gamma_{j+1}(G)$ is free abelian. Hence $G/\gamma_{j+1}(G)$ is a finitely generated torsion-free nilpotent group: by Higman's theorem ([8]; see also Lemma 6 below)

$$\bigcap_{p\in\pi} G^p \leqq \bigcap_{j=0,1,2,\ldots} \gamma_{j+1}(G)=1.$$

For similar reasons every free group has the above property. Theorem 3 shows that it is impossible in Theorem 2 to assume merely that G has an ascending series of *subnormal* subgroups whose factors are infinite cyclic.

5. Sufficient Conditions for an $\mathscr{I}$-Property

First let us consider which abelian and nilpotent groups have $\mathscr{I}(i)$.

Lemma 5. *Let A be an abelian group and write A_p for its p-component. Then A has the property $\mathscr{I}(i)$ if and only if A_p has exponent dividing $p^{i(p)}$ for each prime p and i does not belong to the type of any torsion-free rank* 1 *subgroup of A.*

Proof. Since only the sufficiency of these conditions is in doubt, let us assume that they are satisfied in A. Let D be the intersection of all the $A^{p^{i(p)}}$ and denote the torsion-subgroup of A by T. Then

$$D\cap T=\bigcap_p (A^{p^{i(p)}}\cap T)=\bigcap_p T^{p^{i(p)}}=\bigcap_p A_{p'}=1,$$

where $A_{p'}$ is the p'-component of A. Thus D is torsion-free. Let $1\neq a\in D$; then for each prime p there is an element a_p in A such that

$$a=a_p^{p^{i(p)}}.$$

Let X be the subgroup generated by all the a_p and let $1\neq x\in X$. Write $x=b_1^{t_1}\ldots b_r^{t_r}$ where $t_j\neq 0$ and $b_j=a_{p_j}$. A simple induction on r shows that x has infinite order. Thus X is torsion-free. Finally the mapping $a_p\rightarrow 1/p^{i(p)}$ determines an isomorphism of X with the additive group generated by the rational numbers $1/p^{i(p)}$; the latter, however, is a torsion-free group with rank 1 and type containing i. Hence $D=1$.

If i_r is a function from the set of primes to the set of non-negative integers and n_r is a non-negative integer, then as usual

$$\sum_{r=1}^{k} n_r i_r$$

denotes the function which maps p to $\sum_{r=1}^{k} n_r i_r(p)$.

Lemma 6. *Let G be a nilpotent group with positive class c and let the centre Z_1 of G have $\mathscr{I}(i)$. Then*

(i) *G has $\mathscr{I}(i^*)$ where*

$$i^*=\big(1+(c-1)+(c-1)(c-2)+\cdots+(c-1)(c-2)\ldots 2.1\big)\,i.$$

(ii) *If in addition G is torsion-free, then G has $\mathscr{I}(i)$.*

Proof. The first part follows directly from Lemma 2 if I is taken to be the set of all $p^{i(p)}$.

Now let G be torsion-free; then G/Z_1 is torsion-free. Let p be a prime greater than c. Philip Hall's commutator collection process ([6], §§ 12.3 and 12.4) shows that $G^{p^{i(p)}}$ actually consists of $p^{i(p)}$-th powers. Hence, if D is the intersection of all the $G^{p^{i(p)}}$ for $p>c$,

$$D \cap Z_1 = \bigcap_{p>c} Z_1^{p^{i(p)}}. \tag{4}$$

Define $i'(p)=i(p)$ if $p>c$ and $i'(p)=0$ otherwise. By Lemma 3 the properties $\mathscr{I}(i)$ and $\mathscr{I}(i')$ are identical for torsion-free groups. From (4) one now concludes that $D \cap Z_1 = 1$. Since G is nilpotent and $D \lhd G$, it follows that $D=1$.

Theorem 4. *Let G be a group and let T be a normal subgroup of G satisfying the following conditions.*

(i) *T is the direct product of its Sylow p-subgroups T_p.*

(ii) $\mathrm{Aut}_G(T_p)$ *has finite exponent $p^{j(p)}$ for all p.*

(iii) *The centre of the centralizer of T in G has the property $\mathscr{I}(k)$.*

(iv) *G/T is a torsion-free nilpotent group of class c.*

Then G has the property $\mathscr{I}(i)$ where i depends only on j, k and c.

Lemma 4 and Theorems 2 and 4 provide necessary and sufficient conditions for a group with finite torsion-free rank to have an $\mathscr{I}$-property. Notice that the validity of condition (iii) of Theorem 4 can be decided by reference to Lemma 5.

The value of i furnished by the following proof is without doubt a crude one: however it is linear in j and k.

Proof of Theorem 4. (a) Suppose that H/K is a G-admissible factor of G with finite exponent p^e. Our first move is to prove that $\mathrm{Aut}_G(H/K)$ *has finite exponent dividing* p^{e^*} where

$$e^* = \max\{e(c-1), j(p)\} + \min\{e, j(p)+k(p)\}. \tag{5}$$

In the first place hypotheses (ii) and (iii) imply that T_p has exponent dividing $p^{j(p)+k(p)}$. Since G/T is nilpotent of class c,

$$\mathrm{Aut}_G(HT/KT) = \mathrm{Aut}_{G/T}(HT/KT)$$

has exponent dividing $p^{e(c-1)}$ ([12], Lemma 4.12). Also $H \cap T/K \cap T$ and $(H \cap T_p)/(K \cap T_p)$ are isomorphic as G-operator groups, so $\mathrm{Aut}_G(H \cap T/K \cap T)$ is a homomorphic image of $\mathrm{Aut}_G(T_p)$ and therefore has exponent dividing $p^{j(p)}$. Let L be the subgroup consisting of all elements of G which centralize both HT/KT and $H \cap T/K \cap T$. Then G/L has exponent dividing p^a where

$$a = \max\{e(c-1), j(p)\}.$$

Now L centralizes both $H/(H \cap T)K$ and $(H \cap T)K/K$. Thus if $h \in H$, the mapping $l \to [h, l]K$ is a homomorphism of L into $(H \cap T)K/K$. The intersection of the kernels of all such homomorphisms is precisely $C_G(H/K)$ and the exponent of $(H \cap T)K/K$ divides both p^e and $p^{j(p)+k(p)}$. Hence $\mathrm{Aut}_G(H/K)$ has exponent dividing p^{e^*}.

(b) From now on write
$$C = C_G(T).$$
Then $\gamma_{c+1}(C) \leqq T$, so $\gamma_{c+2}(C) = 1$ and C is nilpotent of class at most $c+1$. By hypothesis the centre of C has $\mathscr{I}(k)$, so by Lemma 6 the group C has $\mathscr{I}(i_1)$ where
$$i_1 = \big(1 + c + c(c-1) + \cdots + c(c-1)\ldots 2.1\big)k. \tag{6}$$
Since T is periodic and G/T is nilpotent, the elements of G/C with finite order form a subgroup S/C such that $T \leqq S$. We shall prove next that S has $\mathscr{I}(i_2)$ where
$$i_2 = i_1 + j. \tag{7}$$
Suppose that this is false and let x be a non-trivial element lying in $S^{p^{i_2(p)}}$ for every p. Since $G/C_G(T_p)$ has exponent $p^{j(p)}$, the group G/C has $\mathscr{I}(j)$. Now $j(p) \leqq i_2(p)$, so x must belong to C. Since C has $\mathscr{I}(i_1)$, there is a prime p such that
$$x \notin X = C^{p^{i_1(p)}}.$$
The group TX/X is the direct product of its Sylow subgroups: let P/X and Q/X denote respectively the p- and p'-components of TX/X. Clearly X, P and Q are normal subgroups of G. Let D be the centraliser of P/X in S: then S/D is a p-group by hypothesis (ii). Now $\gamma_{c+1}(D) \leqq T$ and $[T, D] \leqq Q \leqq D$; hence $\gamma_{c+2}(D) \leqq Q$ and D/Q is nilpotent of class at most $c+1$. Consequently the elements of S/Q whose order is prime to p form a subgroup E/Q such that $E \leqq D$. Now S/C and C/X are periodic, so S/X, and therefore S/Q, is periodic; hence S/E is a p-group. But CE/E has exponent dividing $p^{i_1(p)}$ and S/CE has its unique Sylow p-subgroup of exponent dividing $p^{j(p)}$ because S/C is periodic and G/C has $\mathscr{I}(j)$. Hence S/E has exponent dividing
$$p^{i_1(p)+j(p)} = p^{i_2(p)}.$$
Thus $x \in S^{p^{i_2(p)}} \leqq E$. But E/Q and Q/X are p'-groups, so E/X is a p'-group, while $x \in C$ and C/X is a p-group. Thus $x \in X$, a contradiction.

(c) The group G/S is torsion-free and nilpotent of class at most c since $T \leqq S$. Let G_t/S be the t-th term of the upper central series of G/S, so that
$$S = G_0 \leqq G_1 \leqq \cdots \leqq G_c = G.$$
Let $M = G_r$ and $N = G_{r+1}$ where $0 \leqq r < c$: notice that M/N is torsion-free. Assume that M has $\mathscr{I}(i')$ where $i'(p) \geqq i_2(p)$ for all p. We will prove that N has $\mathscr{I}(i'')$ where, in the notation of (5),
$$i''(p) = \big(i'(p)\big)^* + i'(p) + 1. \tag{8}$$
Suppose that x is a non-identity element belonging to $N^{p^{i''(p)}}$ for all p. Since G/C has $\mathscr{I}(j)$ and $j(p) \leqq i_2(p) \leqq i'(p) \leqq i''(p)$, it follows that $x \in C$: hence $x \in M$. Therefore $x \notin M^{p^{i'(p)}}$ for some p. Now define
$$F = C_N(M/M^{p^{i'(p)}}) \quad \text{and} \quad L = F^{p^{i'(p)+1}} M^{p^{i'(p)}}.$$

By (a) the group N/L has exponent dividing $p^{i''(p)}$, so $x \in L$. However, just as in the proof of Lemma 2, this leads to the contradiction $x \in M^{p^{i'(p)}}$.

From Eqs. (5), (6), (7), and (8) we obtain

$$i'' = c\,i' + j + k + 1 \quad \text{or} \quad i' + 2j + k + 1 \tag{9}$$

according as $c > 1$ or $c = 1$. Eq. (9) allows us to conclude that G has $\mathscr{I}(i)$ where

$$i = i_2\, c^c + (1 + c + c^2 + \cdots + c^{c-1})(j + k + 1) \quad \text{or} \quad i_2 + 2j + k + 1$$

according as $c > 1$ or $c = 1$. It is clear that i depends only on j, k and c.

The problem remains to find suitable conditions for a hypercentral group to have an $\mathscr{I}$-property and to extend Theorem 4 to the case where G/T is hypercentral. This would provide necessary and sufficient conditions for a group in the class $\mathfrak{X}$ to have an $\mathscr{I}$-property.

References

1. Baer, R.: Das Hyperzentrum einer Gruppe III. Math. Z. **59**, 299–338 (1953).
2. Čarin, V.S.: On groups possessing a soluble ascending invariant series. Mat. Sbornik 41, 297–316 (1957).
3. Dunwoody, M.J.: On verbal subgroups of free groups. Arch. der Math. **16**, 153–157 (1965).
4. Fuchs, L.: Abelian groups. Oxford: Pergamon Press 1960.
5. Gruenberg, K.W.: Residual properties of infinite soluble groups. Proc. London Math. Soc. **7** (3), 29–62 (1957).
6. Hall, M.: The theory of groups. New York: MacMillan 1959.
7. Hall, P.: Finiteness conditions for soluble groups. Proc. London Math. Soc. **4** (3), 419–436 (1954).
8. Higman, G.: A remark on finitely generated nilpotent groups. Proc. Amer. Math. Soc. **6**, 284–285 (1955).
9. McLain, D.H.: Remarks on the upper central series of a group. Proc. Glasgow Math. Assoc. 3, 38–44 (1956).
10. Mal'cev, A.I.: Nilpotent torsion-free groups. Izvestija Akad. Nauk. SSSR, Ser. Mat. 13, 201–212 (1949).
11. — On certain classes of infinite soluble groups. Mat. Sbornik **28**, 567–588 (1951); Amer. Math. Soc. Translat. **2** (2), 1–21 (1956).
12. Robinson, D.J.S.: Residual properties of some classes of infinite soluble groups. Proc. London Math. Soc. **18** (3), 495–520 (1968).
13. — Infinite soluble and nilpotent groups. Queen Mary College Mathematics Notes, London 1968.
14. — A property of the lower central series of a group. Math. Z. **107**, 225–231 (1968).
15. Seksenbaev, K.: On the theory of polycyclic groups. Algebra i Logika, Sem. **4**, 79–83 (1965).
16. Šmel'kin, A.L.: Free polynilpotent groups. Izvestija Akad. Nauk SSSR, Ser. Mat. **28**, 91–122 (1964).
17. Wedderburn, J.H.M.: Note on algebras. Ann. of Math. **38** (2), 854–856 (1937).
18. Zassenhaus, H.: The theory of groups, 2nd ed. New York: Chelsea 1958.

Dr. D.J.S. Robinson
Department of Mathematics
University of Illinois
Urbana, Illinois 61801
USA

(Received April 13, 1971)

Math. Z. 124, 133–140 (1972)

Homologies in Collineation Groups of Finite Projective Planes. I

JULIA M. NOWLIN BROWN

§ 1. Introduction

Let $\mathfrak{G}$ be a collineation group of a finite projective plane $\pi=(\mathscr{P}, \mathscr{L}, I)$. By orbits, point orbits, line orbits and point-line orbits we shall mean $\mathfrak{G}$-orbits, $\mathfrak{G}$-orbits in $\mathscr{P}$, $\mathfrak{G}$-orbits in $\mathscr{L}$, and $\mathfrak{G}$-orbits in $\mathscr{P}\times\mathscr{L}$ respectively. By homologies and elations we shall mean homologies and elations in $\mathfrak{G}$.

In [6, pp. 183–184], Fred Piper has proved that if $\mathfrak{G}$ fixes no point or line and if $\mathfrak{G}$ contains order p elations for $p\geqq 2$, then the centers of order p elations form a point orbit, the axes of order p elations form a line orbit, and the center-axis pairs of order p elations form a point-line orbit. Our main result, Theorem 6, is the exact analogue of Piper's result for order q homologies provided that $q>2$ and $\mathfrak{G}$ fixes no point or line. The restriction that $\mathfrak{G}$ fix no point or line is necessary. Otherwise more orbits can occur and in Theorem 4 we easily prove that at most three orbits (of each type) can occur if $q>2$. The restriction that $q>2$ is also necessary in Theorem 6. The main argument in the proof of Theorem 5 gives a partial result (which is included in the statement of Theorem 5), in the case $q=2$. In a later paper [3] (or see Chapter III, Section 6, pp. 95–103 of [2]), we shall use this partial result together with some further arguments to prove that if $q=2$ and $\mathfrak{G}$ fixes no point or line then there are at most two orbits of centers, axes and center-axis pairs of order 2 homologies. The two orbit case does occur, examples are described in Mitchell [4, pp. 213 –214].

I wish to thank Professor Andrew Gleason for his helpful suggestions while I was doing this research.

§ 2. Definitions, Notations, and Results Used

$\pi=(\mathscr{P}, \mathscr{L}, I)$ is a finite projective plane defined by the usual three axioms. Points are elements of $\mathscr{P}$ and will be denoted by capital letters $P, Q, \ldots$. Lines are elements of $\mathscr{L}$ and will be denoted by small letters $k, m, \ldots$. The incidence relation I will be used symmetrically, i.e., both $P\,I\,k$, $k\,I\,P$ will be used for "P and k are incident". The usual geometric language will also be used for incidence. AB means the unique line incident with the distinct points A and B. km means the unique point incident with the distinct lines k and m. (k) is the set of points incident with k. (P) is the set of lines incident with P.

A *collineation* is a pair of bijections of the points and lines which preserves incidence. The same letter will be used to denote both bijections and the induced map on $\mathscr{P} \times \mathscr{L}$.

Elations and *homologies* and their *centers* and *axes* have the usual definitions (and the identity is an elation and a homology). A *center-axis pair* is a point-line pair $(K, k) \in \mathscr{P} \times \mathscr{L}$ such that there is a nonidentity elation or homology in $\mathfrak{G}$ with center K and axis k. *Homology (elation) centers, axes,* and *center-axis pairs* have the obvious meaning. *K is a homology (elation) center of k* and *k is a homology (elation) axis of K* if (K, k) is a homology (elation) center-axis pair.

If $\mathfrak{G}$ is a collineation group of π and $S \subseteq \mathscr{P} \cup \mathscr{L}$, then $\mathfrak{G}_S$ is the subgroup of $\mathfrak{G}$ which fixes S elementwise. Also, $\mathfrak{G}_{(A)(a)} = \mathfrak{G}_{(A) \cup (a)}$ and $\mathfrak{G}_a = \mathfrak{G}_{\{a\}}$.

The following previously known results will be used:

Result 1. Let g be a collineation. If A is a homology (elation) center of a, then $g(A)$ is a homology (elation) center of $g(a)$. (This follows from the fact that $g\mathfrak{G}_{(A)(a)} g^{-1} = \mathfrak{G}_{(g(A))(g(a))}$.)

Result 2. The set of homology centers of a is either empty or it is a $\mathfrak{G}_{(a)}$-orbit. (Satz 3, André [1].)

Result 3. The set of homology centers of a is either empty or it is a $\mathfrak{G}_a$-orbit. (This follows from Results 1 and 2 and the fact that $\mathfrak{G}_{(a)} \lhd \mathfrak{G}_a$.)

Result 4. If $\mathfrak{G}$ fixes no point or line and $p \geqq 2$, then $\mathfrak{G}$ is transitive on the centers of order p elations, on the axes of order p elations, and on the center-axis pairs of order p elations. (Lemmas 1 and 2, Piper [6].)

§ 3. Correspondence of Center, Axis and Center-Axis Pair Orbits

In this section we establish natural 1-1 correspondences between orbits of homology centers, axes and center-axis pairs. These correspondences "preserve the orders" of the related homologies.

Proposition 1. *Let ρ be defined on homology center-axis pairs by $\rho((X, y)) = y$. Then ρ induces a bijection between orbits of homology center-axis pairs and orbits of homology axes.*

Proof. (i) *If $\mathcal{O}$ is an orbit of homology center-axis pairs, then $\rho(\mathcal{O})$ is contained in a line orbit of homology axes.*

If $s, t \in \rho(\mathcal{O})$, then $(S, s), (T, t) \in \mathcal{O}$ for some points S and T. Then, since $\mathcal{O}$ is an orbit, there is a $g \in \mathfrak{G}$ with $g((S, s)) = (T, t)$. Thus, $g(s) = t$.

(ii) *If $\mathscr{A}$ is an orbit of homology axes, then $\rho^{-1}(\mathscr{A})$ is contained in an orbit of homology center-axis pairs.*

Let $(S, s), (T, t) \in \rho^{-1}(\mathscr{A})$. Then $s, t \in \mathscr{A}$. Because $\mathscr{A}$ is an orbit, there is a $g \in \mathfrak{G}$ with $g(s) = t$. Then by Result 1, $(g(S), g(s)) = (g(S), t)$ (as well as (T, t)) is a homology center-axis pair. By André's Result 3, there is an $h \in \mathfrak{G}_t$ which maps $g(S)$ to T. Then $(h g)((S, s)) = h((g(S), t)) = (T, t)$.

(iii) *If $\mathcal{O}$ is an orbit of homology center-axis pairs, then $\mathcal{O}=\rho^{-1}\rho(\mathcal{O})$.*

$\mathcal{O}\subseteq\rho^{-1}\rho(\mathcal{O})$ (by the definition of ρ^{-1})$\subseteq\rho^{-1}(\mathcal{A})$ (for some homology axis orbit $\mathcal{A}$, by (i))$\subseteq\mathcal{O}_1$ (for some homology center-axis pair orbit $\mathcal{O}_1$, by (ii)). But a containment between orbits must be an equality. So $\mathcal{O}=\rho^{-1}\rho(\mathcal{O})$.

(iv) *If $\mathcal{A}$ is an orbit of homology axes, then $\mathcal{A}=\rho\,\rho^{-1}(\mathcal{A})$.*

This is immediate from the definition of ρ^{-1} and the fact that the range of ρ is the set of all homology axes.

The proposition follows immediately from (iii) and (iv).

Definition. A homology center-axis pair orbit and a homology axis orbit are called *corresponding* if they correspond under the bijection of the above proposition. A homology center-axis pair orbit and a homology center orbit are called *corresponding* if they correspond under the bijection defined dually to the one above. A homology center orbit and a homology axis orbit are called *corresponding* if they correspond to the same homology center-axis pair orbit. (Thus, if $\mathcal{C}$ is a homology center orbit, the corresponding homology axis orbit is $\mathcal{A}=\{\ell\in\mathcal{L}\,|\,\ell$ is a homology axis of C for some $C\in\mathcal{C}\}$.)

Note that if an orbit is an orbit of centers, axes, or center-axis pairs of order q homologies, then so are the corresponding orbits. Thus we have the following:

Theorem 2. *Let π be a finite projective plane with collineation group $\mathfrak{G}$. Then the number of $\mathfrak{G}$-orbits of centers of homologies in $\mathfrak{G}=$ the number of $\mathfrak{G}$-orbits of axes of homologies in $\mathfrak{G}=$ the number of $\mathfrak{G}$-orbits of center-axis pairs of homologies in $\mathfrak{G}$. Also, the number of $\mathfrak{G}$-orbits of centers of order q homologies in $\mathfrak{G}=$ the number of $\mathfrak{G}$-orbits of axes of order q homologies in $\mathfrak{G}=$ the number of $\mathfrak{G}$-orbits of center-axis pairs of order q homologies in $\mathfrak{G}$.*

Remark. If $\mathfrak{G}$ fixes no point or line, then a result analogous to Theorem 2 holds for elations. This follows either directly from Result 4 of Piper or from the above discussion modified by replacing homologies by elations throughout and by replacing, in part (ii) of the proof of Proposition 1, André's Result 3 by the last part of Piper's Result 4. So there are natural 1-1 correspondences between elation center, axis, and center-axis pair orbits and between order p elation center, axis, and center-axis pair orbits if $\mathfrak{G}$ fixes no point or line. This restriction on $\mathfrak{G}$ (which does not occur in the analogous homology result) is necessary here. This is easily shown by the example of a Desarguesian plane together with the collineation group consisting of all elations with a fixed axis.

§ 4. The Number of Order q Homology Center, Axis, and Center-Axis Pair Orbits for $q>2$

The main result here is Theorem 6 which asserts that the number mentioned in the title of this section is at most one if $\mathfrak{G}$ fixes no point or line. The analogous result in the case that $\mathfrak{G}$ fixes a point or line is much easier and is included here as Theorem 4. The arguments we use in proving Theorem 6 also prove a partial result in the case $q=2$. This accounts for the alternate "$q=2$ hypothesis" in Lemma 3 and Theorem 5. The $q=2$ part of Theorem 5 will be used in a later paper [3] (or see [2, pp. 95–103]) to obtain a best possible bound on the number of involutory homology center, axis, and center-axis pair orbits in the case that $\mathfrak{G}$ fixes no point or line.

Lemma 3. *Let π be a finite projective plane with collineation group $\mathfrak{G}$. Let $\mathscr{C}_1$ and $\mathscr{C}_2$ be distinct $\mathfrak{G}$-orbits of centers of order q homologies for $q \geqq 2$. Let (C_1, c_1) and (C_2, c_2) be order q homology center-axis pairs with $C_1 \in \mathscr{C}_1$ and $C_2 \in \mathscr{C}_2$. If $q=2$, let at least one of the orbits $\mathscr{C}_1$ and $\mathscr{C}_2$ have the property that none of its points lie on a line of the corresponding line orbit of homology axes. Let $k = C_1 C_2$. Then $c_1 k \in \mathscr{C}_2$ and $c_2 k \in \mathscr{C}_1$.*

Proof. We use proof by contradiction. Suppose $c_2 k \notin \mathscr{C}_1$.

For $i=1,2$, let α_i be an order q homology with center-axis pair (C_i, c_i). Because $\mathscr{C}_1$ is an orbit and by the semi-regularity of α_i on the set of points of k other than C_i and $c_i k$ (i.e. on $(k) - \{C_i\} \cup \{c_i k\}$):

$$|\mathscr{C}_1 \cap (k)| \equiv |\mathscr{C}_1 \cap \{C_i\}| + |\mathscr{C}_1 \cap \{c_i k\}| \pmod q$$

for $i=1,2$. By setting the right sides congruent for $i=1,2$:

$$1 + |\mathscr{C}_1 \cap \{c_1 k\}| \equiv 0 + |\mathscr{C}_1 \cap \{c_2 k\}| \pmod q.$$

So if $c_2 k \notin \mathscr{C}_1$, then $1 + |\mathscr{C}_1 \cap \{c_1 k\}| \equiv 0 \pmod q$. But $|\mathscr{C}_1 \cap \{c_1 k\}| = 0$ or 1. So 1 or $2 \equiv 0 \pmod q$, according as $c_1 k \notin \mathscr{C}_1$ or $c_1 k \in \mathscr{C}_1$. Thus $q=2$ and $c_1 k \in \mathscr{C}_1$.

But now $c_1 k \notin \mathscr{C}_2$ (because $c_1 k \in \mathscr{C}_1$), so the above argument with the indices 1 and 2 interchanged can be applied. Thus $c_2 k \in \mathscr{C}_2$.

So we have $q=2$, $c_1 k \in \mathscr{C}_1$, $c_2 k \in \mathscr{C}_2$. This set of conditions contradicts the hypothesis of the Lemma.

So $c_2 k \in \mathscr{C}_1$. Analogously, $c_1 k \in \mathscr{C}_2$.

Theorem 4. *Let π be a finite projective plane and let $\mathfrak{G}$ be a collineation group of π which fixes a point or line. Let $q > 2$. Then there are at most three $\mathfrak{G}$-orbits of centers of order q homologies in $\mathfrak{G}$, at most three $\mathfrak{G}$-orbits of axes of order q homologies in $\mathfrak{G}$, and at most three $\mathfrak{G}$-orbits of center-axis pairs of order q homologies in $\mathfrak{G}$ and the number of each of these three types of orbits is the same.*

Proof. By duality and Theorem 2, it is sufficient to prove that there are at most three orbits of centers of order q homologies in the case where $\mathfrak{G}$ fixes a line, say k.

We first count the center orbits which meet k. (Compare Ostrom [5, Theorem 3.4, p. 67].) By Lemma 3, if (C_1, c_1), with $C_1 \mathrm{I} k$, is a center-axis pair of an order q homology in $\mathfrak{G}$, then $\mathfrak{G}(C_1)$ and $\mathfrak{G}(c_1 k)$ are the only possible orbits of centers of order q homologies in $\mathfrak{G}$ which meet k.

We now count the center orbits which do not meet k. The set of homology centers off k is the set of homology centers of k because $\mathfrak{G}$ fixes k. Thus, by

André's Result 3 and because $\mathfrak{G}=\mathfrak{G}_k$, the set of centers of order q homologies off k is either empty or consists of exactly one $\mathfrak{G}$-orbit.

So there are at most three orbits of centers of order q homologies in $\mathfrak{G}$ (at most two which meet k and at most one which does not).

Remark. The three orbit case can occur. Consider the example of a finite Desarguesian plane π and the group of all collineations of π which fix two distinct points A and B. Then $\{A\}$, $\{B\}$, and the set of points off the line AB are three orbits of centers of order q homologies for any q which divides the order of the plane minus one.

Theorem 5. *Let π be a finite projective plane and let $\mathfrak{G}$ be a collineation group of π which fixes no point or line of π. Then there is at most one $\mathfrak{G}$-point orbit of centers of order q homologies in $\mathfrak{G}$ if $q>2$. If there is an involutory homology in $\mathfrak{G}$ with no $\mathfrak{G}$-image of its center lying on its axis, then there is only one $\mathfrak{G}$-point orbit of centers of involutory homologies in $\mathfrak{G}$.*

Proof. We use proof by contradiction. Suppose that $\mathscr{C}_1$ and $\mathscr{C}_2$ are two distinct point orbits of centers of order q homologies in $\mathfrak{G}$ with respectively corresponding homology axis orbits $\mathscr{A}_1$ and $\mathscr{A}_2$, and suppose either that $q>2$ or that $q=2$ and at least one of the orbits $\mathscr{C}_1$ and $\mathscr{C}_2$ has the property that none of its points lie on a line of the corresponding line orbit of homology axes.

For $i, j\in\{1, 2\}$, let:

c_{ij} = the number of points of $\mathscr{C}_i$ on a line of $\mathscr{A}_j$,

a_{ij} = the number of lines of $\mathscr{A}_i$ through a point of $\mathscr{C}_j$.

The numbers c_{ij} and a_{ij} are well defined because the $\mathscr{C}_i$'s and $\mathscr{A}_j$'s are orbits. It is clear from the definitions of c_{ij} and a_{ji} that $a_{ij}=0$ iff $c_{ji}=0$. We will deduce a contradiction by showing in (iv) below that either $a_{11}=c_{11}=0$ or $a_{22}=c_{22}=0$ and by showing in (xv) that neither $a_{11}=c_{11}=0$ nor $a_{22}=c_{22}=0$.

Note that our set of hypotheses is selfdual and symmetric in the indices 1 and 2. At no point in our arguments will we make any assumption as to which orbit has the "special property" in the case $q=2$. (In fact, this special property is used directly only in applications of Lemma 3. The conclusions of Lemma 3 are symmetric in the indices.) Thus any of our intermediate results (i)–(xv) can be used as originally stated, dualized, with the indices 1 and 2 interchanged or both dualized and with the indices interchanged.

(i) *Neither $\mathscr{C}_1$ nor $\mathscr{C}_2$ is collinear; neither $\mathscr{A}_1$ nor $\mathscr{A}_2$ is concurrent.*

$\mathfrak{G}$ fixes no line, so no $\mathfrak{G}$-point orbit is collinear unless it consists of a single point. But $\mathfrak{G}$ fixes no point, so no $\mathfrak{G}$-point orbit is collinear. By duality, no $\mathfrak{G}$-line orbit is concurrent.

(ii) *Let $P_1\in\mathscr{C}_1$. Then there are exactly c_{21} lines k through P_1 which meet $\mathscr{C}_2$.*

Let p_1 be an homology axis of P_1. Let $k\,\mathrm{I}\,P_1$. Clearly, k meets $\mathscr{C}_2$ if $p_1 k\in\mathscr{C}_2$. By Lemma 3, k meets $\mathscr{C}_2$ only if $p_1 k\in\mathscr{C}_2$. So k meets $\mathscr{C}_2$ if and only if $p_1 k\in\mathscr{C}_2$.

But since there are exactly c_{21} points of $\mathscr{C}_2$ on p_1, there are exactly c_{21} lines k through P_1 satisfying $p_1 k \in \mathscr{C}_2$ and therefore meeting $\mathscr{C}_2$.

(iii) $c_{21} \geqq 2$, $c_{12} \geqq 2$, $a_{12} \geqq 2$, $a_{21} \geqq 2$.

By (ii), $c_{21} \leqq 1$ implies $\mathscr{C}_2$ is empty or collinear. This violates (i). Thus $c_{21} \geqq 2$. The other assertions follow by interchanging the indices 1 and 2, dualizing, or both.

(iv) *Either* $a_{22} = c_{22} = 0$ *or* $a_{11} = c_{11} = 0$.

Suppose $a_{22} = c_{22} = 0$ is false. Then $c_{22} > 0$, because a_{22} and c_{22} are either both zero or both positive. $c_{21} > 0$ by (iii). Thus, by the definitions of c_{22} and c_{21}, there is a point of $\mathscr{C}_2$ on each of the $a_{21} + a_{11}$ lines of $\mathscr{A}_2 \cup \mathscr{A}_1$ through a point of $\mathscr{C}_1$. Thus, by (ii), $a_{21} + a_{11} \leqq c_{21}$. Dually $c_{21} + c_{11} \leqq a_{21}$. These two inequalities imply that $a_{11} \leqq -c_{11}$. But $a_{11}, c_{11} \geqq 0$. Thus $a_{11} = c_{11} = 0$.

Analogously, if $a_{11} = c_{11} = 0$ is false then $a_{22} = c_{22} = 0$.

(v) *For any line* m, $|(m) \cap \mathscr{C}_2| \leqq c_{21}$.

By the noncollinearity of $\mathscr{C}_1$, there is a point $P_1 \in \mathscr{C}_1$ with $P_1 \not{I}\, m$. For every $M \in (m) \cap \mathscr{C}_2$, $P_1 M$ is a line through P_1 which meets $\mathscr{C}_2$. Thus, $|(m) \cap \mathscr{C}_2| \leqq$ the number of lines through P_1 which meet $\mathscr{C}_2 = c_{21}$ (by (ii)).

(vi) *Let* $P_1 \in \mathscr{C}_1$, $P_1 I k$, $(k) \cap \mathscr{C}_2 \neq \emptyset$. *Let* $a_1 \in \mathscr{A}_1$ *with* $a_1 \not{I}\, P_1$. *Then* $a_1 k \in \mathscr{C}_2$.

Let $t = c_{21}$. By (ii), there is a set of t lines, say $\{k_1, \ldots, k_t\}$, with $k_i I P_1$ and $\mathscr{C}_2 \subseteq \bigcup_{i=1,2,\ldots,t} (k_i)$. Then $(a_1) \cap \mathscr{C}_2 \subseteq (a_1) \cap \bigcup_{i=1,2,\ldots,t} (k_i) = \bigcup_{i=1,2,\ldots,t} \{a_1 k_i\}$ (because $a_1 \neq k_i$ for all i). But $|(a_1) \cap \mathscr{C}_2| = c_{21} = t$ by the definition of c_{21}. So we have a containment between two finite sets of equal size. Thus, the containment is an equality: $(a_1) \cap \mathscr{C}_2 = \bigcup_{i=1,2,\ldots,t} \{a_1 k_i\}$. So $a_1 k_i \in \mathscr{C}_2$ for all i. But k is one of the k_i. So $a_1 k \in \mathscr{C}_2$.

(vii) $a_{12} \leqq c_{21} + 1$. *Moreover*, $a_{12} \leqq c_{21}$ *if there are* $P_1 \in \mathscr{C}_1$, $P_2 \in \mathscr{C}_2$ *with* $P_1 P_2 \notin \mathscr{A}_1$.

Let $P_1 \in \mathscr{C}_1$, $P_2 \in \mathscr{C}_2$. Choose these points with $P_1 P_2 \notin \mathscr{A}_1$ if this is possible. Choose a line k with $k I P_1$, $k \not{I}\, P_2$, and $(k) \cap \mathscr{C}_2 \neq \emptyset$. (Such a k exists because, by (i), $\mathscr{C}_2 \nsubseteq (P_1 P_2)$.) Then $c_{21} \geqq |(k) \cap \mathscr{C}_2|$ (by (v)) $\geqq \left|\bigcup_{x \in \mathscr{A}_1 \cap (P_2)} \{x k\} \cap \mathscr{C}_2\right|$ (because $P_2 \not{I}\, k) = |\mathscr{A}_1 \cap (P_2) - \{P_1 P_2\}|$ (by (vi)) $= a_{12} - 1$ or a_{12} according as $P_1 P_2 \in \mathscr{A}_1$ or $P_1 P_2 \notin \mathscr{A}_1$.

(viii) *If* 1) $a_{12} = c_{21}$ *and* $a_{11} = 0$ *or* 2) $a_{12} = c_{21} + 1$; *then for any two distinct points* $P_2, Q_2 \in \mathscr{C}_2$, $P_2 Q_2 \in \mathscr{A}_1$.

By (i), $\mathscr{C}_1 \nsubseteq (P_2 Q_2)$. So we can choose $P_1 \in \mathscr{C}_1$ with $P_1 \not{I}\, P_2 Q_2$. Let $k = P_1 Q_2$. By (vi), $a_1 k \in \mathscr{C}_2$ for every $a_1 \in \mathscr{A}_1$ with $a_1 \not{I}\, P_1$. Thus:

$$(\mathscr{A}_1 \cap (P_2)) - \{P_1 P_2\} \subseteq \bigcup_{X \in (k) \cap \mathscr{C}_2} X P_2.$$

The line set on the right has size $|(k) \cap \mathscr{C}_2|$ which, by (v), is no larger than c_{21}. The line set on the left has size $a_{12} - e$ where $e = 1$ or 0 according as $P_1 P_2 \in \mathscr{A}_1$

or $P_1 P_2 \notin \mathscr{A}_1$. In case 1), $a_{11}=0$ and so $P_1 P_2 \notin \mathscr{A}_1$. So, in case 1), the line set on the left has size $a_{12}-0=c_{21}$. In case 2), the line set on the left has size at least $a_{12}-1=c_{21}$. So in either case we have a containment between finite sets of equal size. Thus equality prevails and $Q_2 P_2 \in \mathscr{A}_1 \cap (P_2) - \{P_1 P_2\} \subseteq \mathscr{A}_1$.

(ix) $a_{11}=0 \Rightarrow a_{12} \leqq c_{21}-1$.

$a_{11}=0 \Rightarrow P_1 P_2 \notin \mathscr{A}_1$ for all $P_1 \in \mathscr{C}_1$ and $P_2 \in \mathscr{C}_2$. Thus, by (vii), $a_{12} \leqq c_{21}$.

Suppose $a_{12}=c_{21}$. Then, by (viii), $P_2 Q_2 \in \mathscr{A}_1$ for any two distinct points $P_2, Q_2 \in \mathscr{C}_2$. Because of this and the fact that $a_{11}=0$, no line through a point P_1 of $\mathscr{C}_1$ can contain two points of $\mathscr{C}_2$. But then, by Lemma 3, $\mathscr{C}_2$ is contained in any homology axis p_1 of P_1. But this contradicts (i) which asserts the non-collinearity of $\mathscr{C}_2$. Thus $a_{12} \neq c_{21}$.

So $a_{12} \leqq c_{21}$ and $a_{12} \neq c_{21}$. Thus $a_{12} < c_{21}$ or $a_{12} \leqq c_{21}-1$.

(x) $c_{22}=0 \Rightarrow a_{12}=c_{21}+1$.

$c_{21}+1 \leqq a_{12}$ (by (ix) dualized and with the indices 1 and 2 interchanged) $\leqq c_{21}+1$ (by (vii)). Thus equality prevails and $c_{21}+1=a_{12}$.

(xi) $c_{22}=0 \Rightarrow P_2 Q_2 \in \mathscr{A}_1$ *for any two distinct* $P_2, Q_2 \in \mathscr{C}_2$.

This follows immediately from (x) and (viii).

(xii) $c_{22}=0 \Rightarrow P_1 P_2 \in \mathscr{A}_1$ *for any* $P_1 \in \mathscr{C}_1$ *and* $P_2 \in \mathscr{C}_2$.

By (x), $a_{12}=c_{21}+1 \nleqq c_{21}$. So by (vii) there cannot be $P_1 \in \mathscr{C}_1$, $P_2 \in \mathscr{C}_2$ with $P_1 P_2 \notin \mathscr{A}_1$.

(xiii) $c_{22}=0 \Rightarrow c_{11} \leqq 1$.

Let $k \in \mathscr{A}_1$, $P_2 \in \mathscr{C}_2$ with $P_2 \notin k$. Then, by (xi) and (xii), the $c_{21}+c_{11}$ lines XP_2 where $X \in (\mathscr{C}_2 \cup \mathscr{C}_1) \cap (k)$ are lines of $\mathscr{A}_1$. Thus, $c_{21}+c_{11} \leqq a_{12}$ (by the definition of a_{12}) $=c_{21}+1$ (by (x)). Thus $c_{11} \leqq 1$.

(xiv) $c_{22}=0 \Rightarrow c_{21} \leqq a_{11}$.

Let $P_1 \in \mathscr{C}_1$. By (ii), there are exactly c_{21} lines through P_1 which meet $\mathscr{C}_2$. By (xii) these c_{21} lines all belong to $\mathscr{A}_1$. Thus $c_{21} \leqq a_{11}$.

(xv) *Neither* $a_{22}=c_{22}=0$ *nor* $a_{11}=c_{11}=0$ *is possible.*

If $a_{22}=c_{22}=0$, then $2 \leqq c_{21}$ (by (iii)) $\leqq a_{11}$ (by (xiv)) $\leqq 1$ (by the dual of (xiii)). So $a_{22}=c_{22}=0$ is impossible.

Analogously, $a_{11}=c_{11}=0$ *is impossible.*

This proves the theorem.

Theorem 6. *Let π be a finite projective plane and let $\mathfrak{G}$ be a collineation group of π which fixes no point or line of π. Let $q>2$. Then $\mathfrak{G}$ is transitive on the centers of order q homologies in $\mathfrak{G}$, on the axes of order q homologies in $\mathfrak{G}$ and on the center-axis pairs of order q homologies in $\mathfrak{G}$.*

Proof. This follows immediately from Theorem 5 and the last part of Theorem 2.

References

1. André, J.: Über Perspektivitäten in endlichen projektiven Ebenen. Arch. der Math. **6**, 29-32 (1954).
2. Brown, J. M. N.: Homologies and elations of finite projective planes. Thesis. Harvard University: Cambridge, Massachusetts, U.S.A. 1970.
3. — Homologies in collineation groups of finite projective planes. II. (In Preparation.)
4. Mitchell, H. H.: Determination of the ordinary and modular ternary linear groups. Trans. Amer. Math. Soc. **12**, 207-242 (1911).
5. Ostrom, T. G.: Finite translation planes. Berlin-Heidelberg-New York: Springer 1970.
6. Piper, F. C.: Collineation groups containing elations. I. Math. Z. **89**, 181-191 (1965).

Dr. Julia M. Nowlin Brown
Department of Mathematics
York University
Toronto, Ontario, Canada

(Received June 8, 1971)

Math. Z. 124, 141–168 (1972)

Globale Moduln

KLAUS LANGMANN

§ 0. Einleitung

Zu Beginn der Arbeit beweisen wir als Hilfssätze für die in §2 zu zeigenden funktionentheoretischen Aussagen einige rein algebraische Sätze, die auch zum Teil ohne die funktionentheoretische Anwendungen von Interesse sind. 1.1 bzw. 1.2 zeigt, daß unter der Voraussetzung, daß alle Untermoduln eines Moduls M eine (nicht notwendig endliche) Primärzerlegung haben, für M auch „Krullsche Durchschnittsätze" gelten. Hilfssatz 1.3 verallgemeinert den Satz von Cohen [16] auf Moduln. Anschließend (1.3.1) beweisen wir, daß bei Moduln über J-adisch kompletten Ringen besonders leicht die Maximalbedingung zu testen ist. Hilfssatz 1.4 verallgemeinert ein Theorem aus [16]: Moduln M, in denen der J-adische Abschluß aller endlich erzeugten Untermoduln von M wieder endlich erzeugt ist, sind genau dann noethersch, wenn alle Untermoduln $\supset JM$ endlich erzeugt sind. Satz 1.7 zeigt, daß unter schwachen Voraussetzungen für einen Modul M die Eigenschaften „lokal noethersch" (d.h. $M_{\mathfrak{m}}$ ist noetherscher $R_{\mathfrak{m}}$-Modul für alle maximalen Ideale $\mathfrak{m} \subset R$), „Jeder endliche Untermodul von M hat eine Primärzerlegung", „Jeder Untermodul hat eine Primärzerlegung" und „Krullsche Durchschnittssätze für M" äquivalent sind. Ein lokal noetherscher Ring braucht jedoch nicht noethersch zu sein (es gibt solche Integritätsringe, in denen es sogar aufsteigende Hauptidealketten gibt, s. [13]); der folgende Satz 1.8 gibt darüber nähere Auskunft: Ein lokal noetherscher R-Modul, in dem jeder endliche Untermodul eine endliche Primärzerlegung hat, ist noethersch. Satz 1.9 zeigt, daß ein Modul M genau dann noethersch (bzw. lokal noethersch) ist, wenn alle endlichen (bzw. alle) Untermoduln eine endliche (bzw. beliebige) Primärzerlegung haben und M sowie $\mathfrak{m}M$ für alle maximalen Ideale $\mathfrak{m}$ endlich erzeugt (bzw. lokal endlich erzeugt) ist. Speziell ist ein Ring genau dann noethersch, wenn alle maximalen Ideale endlich erzeugt sind und jedes endliche Ideal eine endliche Primärzerlegung hat. Anschließend wird an Hand von Gegenbeispielen gezeigt, daß von diesen Voraussetzungen keine überflüssig ist.

In §2 wenden wir diese Ergebnisse funktionentheoretisch an. Im allgemeinen behandelt man in der Funktionentheorie die Ringe aller in einem k-analytischen Raum Z holomorphen Funktionen oder aber die Ringe aller in einer Umgebung einer kompakten Teilmenge $G \subset Z$ holomorphen Funktionen. Wir wollen hier etwas allgemeiner für beliebige Mengen $G \subset Z$ den Ring $R_Z(G)$ aller in einer Umgebung von G holomorphen Funktionen (auch im nichtarchimedischen) betrachten und insbesondere dabei auch noch Unterringe $R \subset R_Z(G)$ behandeln. Dabei sind natürlich nicht alle Unterringe funktionentheoretisch sinnvoll; zu „kleine" Unterringe etwa haben gar keine Beziehung mehr zu G und Z und können besser als Unterringe des Ringes $R_{Z'}(G')$ für gewisse $Z' \subset Z$, $G' \subset G$ behandelt werden. Aus diesem Grunde fordern wir, daß R gewisse Funktionen z_i enthält, die ähnliche Eigenschaften wie die Koordinatenfunktionen des k^m haben und die damit R zu einem sinnvollen, nicht zu kleinen Unterring von $R_Z(G)$ machen.

Sei nun M ein R-Modul, der aus einer Menge von Schnitten einer kohärenten Garbe $\mathcal{M}$ über G besteht. Dann ist für $N \subset M$ der Ausdruck $N\mathcal{O}_z \subset \mathcal{M}_z$ wohldefiniert, wenn $\mathcal{O}_z$ die Strukturgarbe von G bedeutet. Wir sagen nun, daß ein *R-Untermodul* $N \subset M$ die Eigenschaft (I) bzw. (II) bzw. (III) erfüllt, wenn es beliebig bzw. abzählbar bzw. endlich viele Punkte $z_i \in G$ gibt, so daß aus $m \in M$, $m \in N\mathcal{O}_{z_i}$ $\forall z_i$ schon $m \in N$ folgt. Wir definieren weiter: Der R-Modul M erfüllt die Eigenschaften I bzw. II bzw. III, wenn *alle endlichen Untermoduln* $N \subset M$ die Eigenschaft (I) bzw. (II) bzw. (III) erfüllen. Ist diese letzte Bedingung sogar für *alle Untermoduln* $N \subset M$ erfüllt, so erfülle M die Eigenschaften $\bar{\mathrm{I}}$ bzw. $\overline{\mathrm{II}}$ bzw. $\overline{\mathrm{III}}$.

M erfüllt genau dann I, wenn für alle endlichen R-Untermoduln $N \subset M$ auch M/N auf kanonische Weise als eine Menge von Schnitten über einer (von N abhängigen) Garbe $\mathcal{M}'$ aufgefaßt werden kann. Speziell ist für Steinsche Mengen $G \subset Z$ (d.h. für Mengen G, die ein Umgebungssystem von Steinschen Räumen G_i haben) und für jede kohärente Garbe $\mathcal{M}$ auf G der $R_Z(G)$-Modul $\Gamma(G, \mathcal{M})$ ein Modul mit I bezüglich G. Damit gelten alle im folgenden bewiesenen Sätze auch für die üblicherweise betrachteten globalen Ringe und Moduln.

Satz 2.6 charakterisiert schon die Moduln mit $\bar{\mathrm{I}}$ bzw. $\overline{\mathrm{II}}$ bzw. $\overline{\mathrm{III}}$: Dies sind genau diejenigen R-Moduln, in denen jeder Untermodul $N \subset M$ eine Primärzerlegung $N = \bigcap_{i \in I_*} p_i$ mit $\operatorname{Var} p_i := \{z \in G \mid p_i \mathcal{O}_z \neq M\mathcal{O}_z\} \neq \emptyset$ sowie beliebiger bzw. abzählbarer bzw. endlicher Indexmenge I_* hat und für die die Untermoduln der Form $(z-z^0)^k M$ die Eigenschaft (I) erfüllen (diese letzte Bedingung ist ziemlich unwesentlich, da sie meistens leicht zu verifizieren ist). Satz 2.7 zeigt, daß die etwas schwächeren (aber in der Anwendung wichtigen) Eigenschaften I bzw. II bzw. III für M genau dann erfüllt sind, wenn wir statt der Primärzerlegung $N = \bigcap_{i \in I_*} p_i$ für alle Untermoduln $N \subset M$ nur eine solche Primärzerlegung für alle endlichen Untermoduln $N \subset M$ fordern.

Mit dem folgenden Satz 2.8 können wir sehr viele funktionentheoretische Probleme leicht algebraisch behandeln: Wenn M ein Modul mit I bezüglich G ist und $\mathfrak{m} := \{f \in R \mid f(z_0) = 0\}$ das maximale Ideal aller in $z_0 \in G$ verschwindenden Funktionen aus R bedeutet, so ist der lokalisierte Modul $M_{\mathfrak{m}}$ noethersch über $R_{\mathfrak{m}}$.

Satz 2.9 gibt nun ein handlicheres Kriterium für Moduln mit $\bar{\mathrm{I}}$: Dies sind genau die lokal noetherschen Moduln M, in denen für $N \subsetneq M$ stets $\operatorname{Var} N \neq \emptyset$

ist und die Untermoduln $(z-z^0)^k M$ die Eigenschaft (I) erfüllen. Der sich anschließende Satz 2.10 ist die funktionentheoretische Umformulierung des Satzes 1.8: Ein Modul M erfüllt genau dann $\overline{\mathrm{III}}$, wenn er III erfüllt oder äquivalent dazu, wenn er I erfüllt und noethersch ist. Damit können wir die Moduln mit III leichter algebraisch charakterisieren: Dies sind die noetherschen Moduln M, in denen für $N \subsetneqq M$ stets $\operatorname{Var} N \neq \emptyset$ ist und in denen die Untermoduln $(z-z^0)^k M$ die Eigenschaft (I) erfüllen. Satz 2.11 zeigt dann, daß für diejenigen G, bei denen der volle Ring $R(G)$ noethersch ist (also etwa für semianalytisch kompakte Steinsche Mengen), die Eigenschaften I, $\bar{\mathrm{I}}$, II, $\overline{\mathrm{II}}$, III und $\overline{\mathrm{III}}$ äquivalent sind.

Satz 2.13 besagt, daß die Eigenschaften I und II bzw. $\bar{\mathrm{I}}$ und $\overline{\mathrm{II}}$ (nicht jedoch I und $\bar{\mathrm{I}}$) immer äquivalent sind; bei offenem G gibt es überdies in einem Modul M mit I zu jedem endlichen Untermodul $N \subset M$ eine *diskrete* Menge z_i, so daß aus $m \in M$, $m \in N\mathcal{O}_{z_i}$ schon $m \in N$ folgt. Wenn G nicht offen ist, gibt es nicht immer eine solche diskrete Menge; wir können jedoch immer eine in $\mathring{G}$ diskrete Menge von solchen „erzeugenden Punkten" finden.

Satz 2.14 schließlich zeigt, daß bei kompaktem G (wobei im nichtarchimedischen wir G „kompakt" nennen, wenn G ein Umgebungssystem von affinoiden Bereichen hat) die Eigenschaften I, $\bar{\mathrm{I}}$, II, $\overline{\mathrm{II}}$ (nicht aber III) äquivalent sind, und 2.14 gibt ferner algebraische Kriterien dafür, daß diese funktionentheoretischen Eigenschaften erfüllt sind. Für spezielle kompakte G (die in der Funktionentheorie den G mit $G=\hat{G}$ entsprechen) können wir mit Satz 2.19 bzw. Satz 2.21 die Moduln M mit I bzw. $\bar{\mathrm{I}}$ bzw. III noch einfacher algebraisch charakterisieren, da die manchmal nicht so leicht zu überprüfende Voraussetzung $\operatorname{Var} N \neq \emptyset$ für $N \subsetneqq M$ hier sehr einfach zu verifizieren ist. Mit 2.19 bzw. 2.21 brauchen wir zum Testen der Eigenschaften I bzw. III im wesentlichen nur zu prüfen, ob M noethersch bzw. lokal noethersch ist.

§ 1. Maximalbedingung, Primärzerlegung und Durchschnittssätze

Zunächst sollen die algebraischen Hilfsmittel für die in § 2 zu zeigenden funktionentheoretischen Sätze bereitgestellt werden.

Sei R ein beliebiger (kommutativer) Ring, M ein R-Modul. Wir nennen – etwas abweichend von der üblichen Definition – einen Untermodul $N \subset M$ primär, wenn jede Homotheti a_r: $M/N \to M/N$ $x \to rx$ entweder injektiv oder nilpotent (d. h. $a_{(r^n)}=0$) ist. Dann gilt:

Hilfssatz 1.1. *Sei M ein R-Modul. Jeder Untermodul N von M sei Durchschnitt von primären Untermoduln; $J \subset R$ sei ein endliches (d. h. endlich erzeugtes) Ideal. Dann gilt für alle Untermoduln $N \subset M$:*

$$\bigcap_k N + J^k M = \{f \in M \mid f(1-j) \in N \text{ für ein } j \in J\}.$$

Beweis. „$\supset$" Ist $f(1-j) \in N$, so ist $f = n_1 + jf = n_1 + j(n_1 + jf) = n_2 + j^2 f$ usw., also $f = n_k + j^k f$.

„$\subset$" Wenn diese Inklusion falsch wäre, gäbe es ein $x \in M$, $x \in \bigcap_k N + J^k M$, aber $x \notin xJ + N$. Nun hat der Untermodul $xJ + N$ eine Primärzerlegung; es gibt somit einen primären Untermodul p mit $p \supset xJ + N$, $x \notin p$. Sei $J = (x_1, \ldots, x_n)$. Es folgt dann aus $x \notin p$ und $x_i x \subset p$, daß $x_i^e M \subset p$ ist; somit ist also $J^{ne} M \subset p$ und damit auch $N + J^{ne} M \subset p$. Da $x \in \bigcap N + J^k M$ ist, folgte $x \in p$ im Widerspruch zur Annahme.

Diesen Hilfssatz brauchen wir später in einer etwas schärferen Form:

Bemerkung 1.1. Statt „J endlich" genügt es zu fordern: Es gibt zu jedem maximalen $\mathfrak{m}$ ein endliches Ideal $I \subset J$, so daß $IMR_{\mathfrak{m}} = JMR_{\mathfrak{m}}$ ist.

Beweis. Sei wie in Beweis 1.1 $p \supset xJ + N$, $x \notin p$. Sei $\mathfrak{m}$ ein maximales Ideal, das alle Nullteiler mod p enthält und sei I das vorausgesetzte endliche Ideal. Aus $x \notin p$ und $p \supset xJ + N$ folgt $I^k M \subset p$. Sei nun $y \in J^k M$. Wegen $JMR_{\mathfrak{m}} = IMR_{\mathfrak{m}}$ ist $J^k MR_{\mathfrak{m}} = I^k MR_{\mathfrak{m}}$ und damit $yR_{\mathfrak{m}} \subset I^k MR_{\mathfrak{m}} = (i_1, \ldots, i_n)^k MR_{\mathfrak{m}} \subset pR_{\mathfrak{m}}$. Somit ist $y \cdot s \in p$ für ein $s \notin \mathfrak{m}$; also ist nach Definition von $\mathfrak{m}$ schon $y \in p$. Somit ist $J^k M \subset p$. Der Beweis geht dann wie oben weiter.

Genauso wie 1.1 beweisen wir den folgenden Hilfssatz (hierbei ist zu beachten, daß für endliche N der Untermodul $xJ + N$ wieder endlich ist):

Hilfssatz 1.2. *Sei M ein R-Modul, in dem alle* endlichen *Untermoduln N eine Primärzerlegung haben. Ist J ein endliches Ideal, so gilt für jeden endlichen Untermodul N*

$$\bigcap_k N + J^k M = \{f \mid f(1-j) \in N \text{ für ein } j \in J\}.$$

Bemerkung 1.2. Statt „J endlich" brauchen wir in 1.2 nur zu fordern: Es gibt ein endliches Ideal $I \subset J$ mit $IM = JM$.

Beweis. Aus $IM = JM$ folgt $J^k M = I^k M$. Daraus folgt

$$\bigcap_k N + J^k M = \bigcap_k N + I^k M = \{f \in M \mid f(1-i) \in N\} \subset \{f \in M \mid f(1-j) \in N\}.$$

Wir nennen nun einen Untermodul $N \subset M$ prim, wenn jede Homotheti a_r: $M/N \to M/N$ $x \to xr$ entweder injektiv oder die Nullabbildung ist.

Ferner definieren wir für einen Modul M als $|M|$ die kleinstmöglichste Anzahl von Erzeugenden für M.

Dann gilt:

Hilfssatz 1.3. *Sei M ein R-Modul.*

a) *Ist N ein maximaler nicht endlich erzeugter Untermodul von M, so gilt für jedes $x \in R$ mit $xM \not\subset N$: Es gibt einen endlichen Untermodul $Q \subset N$ mit $N = Q + xN$.*

b) *Ist N ein nicht primer, aber sonst beliebiger Untermodul von M, so gibt es einen Untermodul $N' \supsetneq N$ mit $|N| \leqq 2|N'|$. Insbesondere ist also ein maximaler nicht endlich erzeugter Untermodul immer prim.*

c) *M ist noethersch genau dann, wenn alle primen Untermoduln endlich erzeugt sind.*

Beweis. a) Sei $xM \not\subset N$. Dann ist $N+xM$ endlicher Untermodul; also $N+xM=(x_1, \dots, x_r)$. Die x_i haben die Gestalt $x_i=n_i+m_i$ mit $n_i \in N$, $m_i \in xM$. Es folgt $N \subset (n_1, \dots, n_r)+xM$ und somit $N=(n_1, \dots, n_r)+xM'$ für einen Untermodul M', der o.B.d.A. $\supset N$ angenommen werden kann. Wäre $M' \supsetneqq N$, so wäre M' endlicher Untermodul und damit auch $N=(n_1, \dots, n_r)+xM'$ endlich.

b) Da N nicht prim ist, gibt es $x \in R$ und $m \in M-N$ mit $xm \in N$, $xM \not\subset N$. Wie unter a) folgt $N=(n_1, \dots, n_r)+xM'$ mit $r=|N+xM|$, $M' \supset N$. Dann folgt aus $N=(n_1, \dots, n_r)+x(M'+mR)$ die Behauptung.

c) Nach dem Zornschen Lemma existierte ein maximaler nicht endlich erzeugter Untermodul, wenn M nicht noethersch wäre. Dieser wäre nach b) prim.

Als leichte Anwendung von 1.3 sei folgender Satz genannt, den wir jedoch nicht weiter brauchen werden:

1.3.1. *Sei R ein Ring, komplett bezüglich der J-adischen Topologie für ein Ideal $J \subset R$. Ein separierter Modul M ist genau dann noethersch, wenn alle Untermoduln $N \supset JM$ endlich erzeugt sind.*

Beweis. Sei p ein maximaler nicht endlicher Untermodul. Es ist $p \not\supset JM$; somit gibt es $i \in J$ mit $p \not\supset iM$. Nach 1.3.a) gibt es $p_i \in p$ mit $p=(p_1, \dots, p_m)+ip$. Zu irgendeinem $x_n \in p$ existieren also ein $x_{n+1} \in p$ und $r_k^n \in R$ mit

$$x_n = \sum_1^m r_k^n p_k + i x_{n+1}.$$

Wir konstruieren dann zu einem festen x_0 die so definierte Folge x_n und die daraus resultierenden r_k^n und setzen weiter $r_k := \sum_{n=0}^{\infty} r_k^n i^n$. Dann ist $x_0 - \sum_{k=0}^{m} r_k p_k$ aus $i^n M$ für alle n; somit ist also $x_0 \in (p_1, \dots, p_k)$ und damit p endlich.

Bemerkung. Der Beweis zeigt genauer, daß gilt:

1.3.2. *R sei komplett und M separiert bezüglich der (i)-adischen Topologie für jedes $i \in J$. M ist noethersch genau dann, wenn alle $N \supset JM$ endlich erzeugt sind.*

Daraus folgt dann sofort, wenn J das Nilradikal des Ringes R ist:

1.3.3. *M ist noethersch genau dann, wenn alle Untermoduln $N \supset JM$ endlich erzeugt sind.*

1.3 brauchen wir auch zum Beweis des nun folgenden Hilfssatzes 1.4. Dabei bedeute $\operatorname{Jak} R$ das Jakobsonradikal des Ringes R:

Hilfssatz 1.4. a) *Sei M ein R-Modul; $J \supset \operatorname{Jak} R$ sei ein festes Ideal. Ist für alle endlichen Untermoduln $N \subset M$ wieder $\bigcap_k N+J^k M$ endlich und sind alle Untermoduln $Q \supset JM$ endlich, so ist M noethersch.*

b) *Ist speziell R quasisemilokal mit dem endlich erzeugten Jakobsonradikal J, so gilt: Ist M ein endlicher R-Modul, so daß für alle endlichen N wieder $\bigcap_k N+J^k M$ endlich ist, so ist M noethersch.*

Beweis. a) Sei p ein maximaler nicht endlich erzeugter Untermodul. Es ist dann $p \not\supset JM$; somit ist für ein $x \in J$ auch $p \not\supset xM$. Nach 1.3 folgt dann $p = (f_1, \dots, f_r) + xp$ und daraus durch Iteration $p = (f_1, \dots, f_r) + x^k p$; somit ist $p \subset \bigcap_k (f_1, \dots, f_r) + J^k M$. Es ist aber auch $p \supset \bigcap_k (f_1, \dots, f_r) + J^k M$: Denn $\overline{M} := M/p$ ist noetherscher R-Modul; deshalb ist nach 1.1 $\bigcap_k J^k \overline{M} = \{\overline{0}\}$, also $\bigcap_k p + J^k M = p$ und damit $\bigcap_k (f_1, \dots, f_r) + J^k M \subset p$. Damit haben wir

$$\bigcap_k (f_1, \dots, f_r) + J^k M = p$$

gezeigt; dies ist ein Widerspruch zur Voraussetzung.

b) Es ist JM ein endlicher Untermodul von M. Ferner ist für jedes maximale Ideal $\mathfrak{m}$ der Quotient $M/\mathfrak{m}M$ Artinsch, somit ist auch M/JM Artinsch. Daraus folgt, daß alle Untermoduln $Q \supset JM$ endlich sind.

Bemerkung 1.4. Wird in 1.4.b) das Jakobsonradikal nicht als endlich vorausgesetzt, so ist diese Aussage falsch. Es genügt jedoch – wie der Beweis zeigt – vorauszusetzen, daß ein endliches Ideal I existiert, so daß $IM = JM$ ist. Weiter können die Durchschnittseigenschaften in a) und b) auch gelten, wenn M nicht endlich erzeugt ist; somit ist diese Voraussetzung auch wesentlich (s. Beispiele nach 1.9).

Hilfssatz 1.5. *Sei M ein R-Modul, $\mathfrak{m} \subset R$ ein maximales Ideal. Dann ist jeder Untermodul $N \supset \mathfrak{m}^k M$ primär; die Menge der Nullteiler* mod N *ist gerade das maximale Ideal $\mathfrak{m}$.*

Beweis. Betrachten wir die Homotheti a_r: $M/N \to M/N$: Ist $r \in \mathfrak{m}$, so ist $r^k M \subset N$, somit ist die Homotheti nilpotent. Ist $r \notin \mathfrak{m}$, so muß a_r injektiv sein: Sonst gäbe es $x \notin N$ mit $rx \in N$. Es ist aber auch $\mathfrak{m}^k \cdot x \subset N$, somit wäre $(rR + \mathfrak{m}^k)\, x \subset N$ und damit $x \in N$.

Hilfssatz 1.6. *Sei M ein R-Modul, $\mathfrak{m} \subset R$ sei ein endliches maximales Ideal, so daß $M_{\mathfrak{m}}$ endlicher $R_{\mathfrak{m}}$-Modul ist. Dann gilt für jeden Untermodul $N \supset \mathfrak{m}^k M$: M/N Artinsch (wobei hier wie im folgenden ein Modul Artinsch heiße, wenn er von endlicher Länge ist).*

Beweis. Sei $N_1 \supsetneqq N_2 \supsetneqq \cdots \supsetneqq N \supset \mathfrak{m}^k M$. Nach 1.5 sind die N_i primär; somit ist $(N_1)_{\mathfrak{m}} \supsetneqq (N_2)_{\mathfrak{m}} \supsetneqq \cdots \supsetneqq N_{\mathfrak{m}} \supset (\mathfrak{m}^k M)_{\mathfrak{m}}$. Da $M_{\mathfrak{m}}$ endlicher $R_{\mathfrak{m}}$-Modul ist und deshalb $\mathfrak{m}^v M_{\mathfrak{m}}$ endlich erzeugt ist, ist $M_{\mathfrak{m}}/(\mathfrak{m}^k M)_{\mathfrak{m}}$ endlicher $R/\mathfrak{m}$-Vektorraum; somit $M_{\mathfrak{m}}/(\mathfrak{m}^k M)_{\mathfrak{m}}$ Artinscher Modul.

Bemerkung 1.6. Auch in 1.6 sind die Voraussetzungen „$\mathfrak{m}$ endlich" und „$M_{\mathfrak{m}}$ endlich über $R_{\mathfrak{m}}$" nicht unnötig. Es genügt jedoch wieder – wie der Beweis zeigt – zu fordern, daß es ein endliches Ideal $I \subset \mathfrak{m}$ gibt, so daß $IM = \mathfrak{m}M$ ist.

Wir nennen nun einen R-Modul M *lokal endlich*, wenn $M_{\mathfrak{m}}$ endlicher $R_{\mathfrak{m}}$-Modul für alle maximalen Ideale $\mathfrak{m}$ ist. M heiße *lokal noethersch*, wenn jeder Untermodul lokal endlich ist. Dann gilt:

Satz 1.7. *Sei R ein Ring mit endlich erzeugten maximalen Idealen, M ein lokal endlicher R-Modul. Dann ist äquivalent*

1) *Jeder Untermodul N besitzt eine Primärzerlegung.*
2) *Jeder endliche Untermodul N besitzt eine Primärzerlegung.*
3) *M ist lokal noethersch.*
4) $\bigcap_{\mathfrak{m}\in \operatorname{Max\,Spec} R} \bigcap_k N+\mathfrak{m}^k M = N$ *für alle $N \subset M$.*
5) $\bigcap_{\mathfrak{m}} \bigcap_k N+\mathfrak{m}^k M = N$ *für alle endlich erzeugten $N\subset M$.*
6) $\bigcap_k N+J^k M=\{f\in M \mid f\cdot(1-j)\in N$ *für ein* $j\in J\}$ *für alle $N\subset M$ und alle endlichen Ideale $J\subset R$.*
7) $\bigcap_k N+J^k M=\{f\in M \mid f(1-j)\in N$ *für ein* $j\in J\}$ *für alle endlichen $N\subset M$ und für alle endlichen Ideale $J\subset R$.*
8) *Jeder Untermodul N ist Durchschnitt von Untermoduln N_i mit M/N_i Artinsch.*
9) *Jeder endliche Untermodul N ist Durchschnitt von Untermoduln N_i mit M/N_i Artinsch.*
10) *Jeder endliche Untermodul N ist Durchschnitt von Untermoduln N_i mit M/N_i noethersch.*
11) *Jeder endliche Untermodul N ist Durchschnitt von Untermoduln N_i mit M/N_i lokal noethersch.*

Beweis. 2)→3) Sei N ein endlicher Untermodul, $x\in\bigcap_k N_{\mathfrak{m}}+\mathfrak{m}^k M_{\mathfrak{m}}$. Es ist also $x=y/s$ mit $s\notin\mathfrak{m}$, $y\in M$ und $y\cdot s_k\in N+\mathfrak{m}^k M$ für ein $s_k\notin\mathfrak{m}$. Nach 1.5 folgt daraus $y\in N+\mathfrak{m}^k M$; somit nach 1.2 $y(1-m)\in N$ für ein $m\in\mathfrak{m}$. Daraus folgt, daß $x\in N_{\mathfrak{m}}$ ist. Es ist also $\bigcap_k N_{\mathfrak{m}}+\mathfrak{m}^k M_{\mathfrak{m}}=N_{\mathfrak{m}}$ für alle endlichen $R_{\mathfrak{m}}$-Moduln $N_{\mathfrak{m}}$. Nach 1.4 folgt die Behauptung

3)→1) Da $M_{\mathfrak{m}}$ noethersch ist, besitzt $N_{\mathfrak{m}}$ eine Primärzerlegung in $M_{\mathfrak{m}}$: $N_{\mathfrak{m}}=\bigcap_i (N_i^{\mathfrak{m}})_{\mathfrak{m}}$ (s. [23]; $R_{\mathfrak{m}}$ braucht nicht noethersch zu sein). Dann ist aber auch $N=\bigcap_{\mathfrak{m}}\bigcap_i N_i^{\mathfrak{m}}\cap M$ eine Primärzerlegung in M für N.

1)→4) Sei $x\in\bigcap_{\mathfrak{m}}\bigcap_k N+\mathfrak{m}^k M$, $x\notin N$. Es gibt dann ein maximales Ideal $\mathfrak{m}_0$ mit folgender Eigenschaft: Ist $r\in R$, $x r\in N$, so ist $r\in\mathfrak{m}_0$. Es ist nun insbesondere $x\in\bigcap N+\mathfrak{m}_0^k M$, somit nach 1.1 $x(1-\mathfrak{m}_0)\in N$ für ein $m_0\in\mathfrak{m}_0$. $1-m_0$ ist aber nicht aus $\mathfrak{m}_0$, im Widerspruch zur Definition von $\mathfrak{m}_0$.

5)→2) Nach 1.5 ist $N+\mathfrak{m}^k M$ primär.

1)→6) Folgt aus Hilfssatz 1.1.

7)$\to$3) Es gilt insbesondere für jedes maximale $\mathfrak{m}$:

$$\bigcap_k N + \mathfrak{m}^k M = N_{\mathfrak{m}} \cap M.$$

Daraus folgt nach 1.5 wie unter 2)$\to$3):

$$\bigcap_k N_{\mathfrak{m}} + \mathfrak{m}^k M_{\mathfrak{m}} = N_{\mathfrak{m}}.$$

Somit ist $M_{\mathfrak{m}}$ nach 1.4 noethersch.

4)$\to$8) Nach 1.6 ist $M/N + \mathfrak{m}^k M$ Artinsch.

11)$\to$2) Da für alle lokal endlichen Moduln 1) und 3) äquivalent sind, ist jedes N_i und somit auch N Durchschnitt von primären Moduln.

Nach Satz 1.3 werden wir an Hand von Gegenbeispielen zeigen, daß die Voraussetzungen „M lokal endlich" und „alle maximalen Ideale endlich" nötig sind. Es gilt jedoch:

Bemerkung 1.7.1. Ist M lokal endlich und gibt es zu jedem maximalen $\mathfrak{m}$ ein endliches Ideal $I \subset \mathfrak{m}$ mit $IM = \mathfrak{m}M$, so ist Satz 1.7 richtig.

Beweis. Die im Beweis von 1.7 gebrauchten Sätze 1.1, 1.2, 1.4, 1.5 und 1.6 sind auch unter dieser abgeschwächten Voraussetzung gültig (s. Bemerkungen 1.1, 1.2, 1.4 und 1.6).

Wir haben in 2)$\to$3) mehr bewiesen:

Bemerkung 1.7.2. Ist M ein R-Modul, in dem jeder endliche Untermodul eine Primärzerlegung hat, und ist $\mathfrak{m}$ ein maximales Ideal, so daß es ein endliches $I \subset \mathfrak{m}$ gibt mit $IM = \mathfrak{m}M$, so ist äquivalent:

1) $M_{\mathfrak{m}}$ ist noetherscher $R_{\mathfrak{m}}$-Modul.
2) $M_{\mathfrak{m}}$ ist endlicher $R_{\mathfrak{m}}$-Modul.

Es sei noch erwähnt, daß entsprechend Hilfssatz 1.4 gilt:

Sei R ein Ring mit endlich erzeugten maximalen Idealen. Ein R-Modul M ist genau dann lokal noethersch, wenn für jeden endlichen Untermodul N der Durchschnitt aller umfassenden primären Moduln wieder ein endlicher Modul ist.

1.7 gibt Kriterien dafür, daß ein Modul M lokal noethersch ist. Wir wollen nun untersuchen, wie eng die Eigenschaften „lokal noethersch" und „noethersch" miteinander verknüpft sind. Es gilt hier:

Satz 1.8. *Sei M ein lokal noetherscher R-Modul; jeder endliche Untermodul $N \subset M$ habe eine endliche Primärzerlegung. Dann ist M noethersch.*

Beweis. Für die Dauer dieses Beweises definieren wir: Ist $p_i^j \subset M$ ein Primärmodul, so bezeichne q_i^j das Primideal aller Nullteiler von M/p_i^j.

Sei nun p ein maximaler nicht endlicher Untermodul von M. p ist nach 1.3 prim. Sei $N_0 \subset p$ ein beliebiger endlich erzeugter Untermodul. Es ist nach Voraussetzung $N_0 = \bigcap_{i=1}^{n_0} p_i^0 \cap p$ mit primären $p_i^0 \subset M$, $p_i^0 \not\supset p$. Es gibt sodann, da M lokal noethersch ist, einen endlichen Untermodul N_1 mit $N_0 \subset N_1 \subset p$, $N_1 R_{q_i^0} = p R_{q_i^0}$ für $i = 1, \ldots, n_0$.

Es ist dann nach Voraussetzung $N_1=\bigcap\limits_1^{n_1} p_j^1\cap p$ mit $p_j^1\not\supset p$. Es kann nun für kein Paar (i,j) $q_i^0=q_j^1$ sein: Denn wäre dies erfüllt, so folgte aus

$$pR_{q_i^0}=N_1R_{q_i^0}=\left(\bigcap_1^{n_1}p_i^1\cap p\right)R_{q_i^0}\subset p_j^1R_{q_i^0},$$

daß $p\subset p_j^1$ ist, im Widerspruch zur Definition von p_j^1.

Es gibt aber zu jedem q_j^1 ein q_i^0 mit $q_i^0\subset q_j^1$. Wäre nämlich $q_j^1\not\supset q_i^0$ für alle i, so wäre $p_i^0\otimes R_{q_j^1}=M\otimes R_{q_j^1}$ für alle i, also folgte aus

$$p_j^1\supset N_1\supset N_0=p\cap\bigcap_{i=1}^{n_0}p_i^0,$$

daß

$$p_j^1\otimes R_{q_j^1}\supset p\otimes R_{q_j^1}\cap\bigcap_{i=1}^{n_0}p_i^0\otimes R_{q_j^1}=p\otimes R_{q_j^1}$$

und somit $p_j^1\supset p$ wäre.

Aus den letzten beiden Absätzen folgt also: Zu jedem q_j^1 gibt es ein $q_{i(j)}^0$ mit $q_j^1\supsetneqq q_{i(j)}^0$; wir führen dasselbe Verfahren iterativ durch und erhalten nach dem m. Schritt primäre Moduln p_j^m mit

$$q_j^m\supsetneqq q_{i(j)}^{m-1}. \tag{$*$}$$

Angenommen, dies Verfahren würde nicht abbrechen. Dann erhielten wir eine unendliche Primidealkette $q_{i_0}^0\subsetneqq q_{i_1}^1\subsetneqq q_{i_2}^2\subsetneqq\cdots$. (Zum Beweis betrachte man etwa die Menge I_0 der i_0, so daß die Länge aller solcher Ketten, angefangen mit $q_{i_0}^0$, nicht beschränkt ist. Diese Menge ist nicht leer: Denn sonst hätte, da es nur endlich viele q_i^0 gibt, jede Kette höchstens die Länge m für ein $m\in\mathbb{N}$. Daß es aber beliebig lange Ketten gibt, folgt sofort aus ($*$). Dann betrachte man zu festem $i_0\in I_0$ die Menge I_1 der i_1 mit $q_{i_1}^1\supset q_{i_0}^0$, so daß auch hier die Länge aller Ketten, angefangen mit $q_{i_1}^1$, nicht beschränkt ist. Auch diese Menge ist nicht leer: Sonst wäre nämlich, da es nur endlich viele $q_i^1\supset q_{i_0}^0$ gibt, die Ketten $q_{i_0}^0\subsetneqq\cdots$ von beschränkter Länge. Wenn man so weiterfährt, erhält man eine aufsteigende Primidealkette.)

Sei dann $q:=\bigcup\limits_k q_{i_k}^k$. q ist prim in R. Da $q_{i_0}^0$ zu dem primären Untermodul $p_{i_0}^0$ assoziiert ist, ist $R/q_{i_0}^0\subset M/p_{i_0}^0$; da M_q noetherscher R_q-Modul ist, ist also auch $(R/q_{i_0}^0)_q$ noethersch. Wir hätten somit in $(R/q_{i_0}^0)_q$ eine echt aufsteigende Kette von Primidealen $(q_{i_k}^k/q_{i_0}^0)_q$.

Für das Weitere wird folgende Bemerkung wichtig sein:

Bemerkung 1.8. Ist jeder endliche Untermodul $N\subset M$ endlicher Durchschnitt von primären Untermoduln p_i, so daß für $\{q_i\}=\operatorname{Ass} M/p_i$ der Modul M_{q_i} noetherscher R_{q_i}-Modul ist, und gibt es keine echt aufsteigende Primidealkette von solchen q_i, so ist M noethersch.

Beweis. Nur diese Voraussetzungen wurden im Beweis von Satz 1.8 benutzt.

Es sei noch bemerkt, daß es sehr wohl lokal noethersche Moduln gibt (in denen jeder Untermodul eine Primärzerlegung hat), die aber nicht noethersch sind (s. Gegenbeispiel nach 1.9).

Satz 1.9. *Sei R ein beliebiger kommutativer Ring, M ein R-Modul.*

A. *Dann ist äquivalent*

1) *M ist lokal noethersch.*
2) a) *Jeder Untermodul hat eine Primärzerlegung.*
   b) *M und $\mathfrak{m}M$ ist lokal endlich für alle maximalen Ideale $\mathfrak{m}$.*
3) a) *In jedem Untermodul $N \subset M$ existieren maximale Ideale $\mathfrak{m}_i \subset R$ mit*

$$N = \bigcap_{i \in I} \bigcap_{k} N + \mathfrak{m}_i^k M .$$

   b) *M und $\mathfrak{m}M$ ist lokal endlich für alle maximalen Ideale $\mathfrak{m}$.*

B. *Ferner ist äquivalent:*

1) *M ist noethersch.*
2) a) *Jeder endliche Untermodul hat eine endliche Primärzerlegung.*
   b) *M und $\mathfrak{m}M$ ist endlich erzeugt für alle maximalen Ideale $\mathfrak{m} \subset R$ (es genügt, lokal endlich erzeugt vorauszusetzen).*
3) a) *Zu jedem endlichen Untermodul $N \subset M$ gibt es endlich viele maximale Ideale $\mathfrak{m}_i$ mit*

$$N = \bigcap_{i=1}^{m} \bigcap_{k} N + \mathfrak{m}_i^k M .$$

   b) *M und $\mathfrak{m}M$ ist endlich erzeugt für alle maximalen Ideale $\mathfrak{m} \subset R$.*

Ist R nicht lokal, so ist die Voraussetzung „M endlich" in B. 2b) *bzw.* B. 3b) *überflüssig.*

Beweis. Aus $\mathfrak{m}M$ lokal endlich erzeugt folgt: Zu jedem $\mathfrak{m}$ gibt es ein endliches Ideal $I \subset \mathfrak{m}$ mit $IMR_{\mathfrak{m}} = \mathfrak{m}MR_{\mathfrak{m}}$.

A. 1)→2) a) wie 1.7 3)→1).

2)→3) a) vgl. 1.7 1)→4): Wir benutzen statt 1.1 die Bemerkung 1.1, deren Voraussetzungen nach obigem erfüllt sind.

3)→2) Nach 1.5 ist $N + \mathfrak{m}_i^k M$ primär.

2)→1) Vergleiche 1.7 2)→3): Wir benutzen statt 1.2 wieder die Bemerkung 1.1, statt 1.4 Bemerkung 1.4.

B. 2)→1) Es hat wegen a) jeder endliche Untermodul $NR_{\mathfrak{m}} \subset MR_{\mathfrak{m}}$ eine Primärzerlegung; weiter gibts nach obigem zu $\mathfrak{m}$ ein endliches Ideal $I \subset \mathfrak{m}$ mit $IMR_{\mathfrak{m}} = mMR_{\mathfrak{m}}$, und es ist $MR_{\mathfrak{m}}$ endlich für alle $\mathfrak{m}$. Dann ist nach Bemerkung 1.7.1 schon $MR_{\mathfrak{m}}$ noethersch. Nach 1.8 folgt dann die Behauptung

1)→3) Sei $N = \bigcap_{1}^{m} N_i$, N_i primär. Nach Bemerkung 1.1 folgt dann, wenn $\mathfrak{m}_i$ ein maximales Ideal ist, das alle Nullteiler mod N_i enthält: $N_i = \bigcap_{k} N_i + \mathfrak{m}_i^k M$. Damit folgt sofort $N = \bigcap_{1}^{m} \bigcap_{k} N + \mathfrak{m}_i^k M$.

3)$\to$2) Aus $\mathfrak{m}M$ endlich erzeugt folgt sofort: Es gibt ein endliches Ideal $I \subset \mathfrak{m}$ mit $IM = \mathfrak{m}M$. Aus Bemerkung 1.7.1 folgt dann mittels a), daß $M_{\mathfrak{m}}$ noethersch ist für alle $\mathfrak{m}$. Es ist also $N_{\mathfrak{m}_i}$ endlicher Durchschnitt von primären Moduln $p_j^i \subset M_{\mathfrak{m}_i}$; dann ist $\bigcap_i N_{\mathfrak{m}_i} \cap M = \bigcap_i \bigcap_j p_j^i \cap M$. Es ist aber $N_{\mathfrak{m}_i} \cap M \subset \bigcap_n N + \mathfrak{m}_i^n M$: Denn ist $X s \in N$ für ein $x \in M$, $s \notin \mathfrak{m}_i$, so ist $X s \in N + \mathfrak{m}_i^n M$ und wegen 1.5 $x \in N + \mathfrak{m}_i^n M$.

Dann folgt aus 3a), daß $\bigcap N_{\mathfrak{m}_i} \cap M = N$ und somit $\bigcap_i \bigcap_j p_j^i \cap M = N$ ist.

Satz 1.9.1. *Ein Ring ist genau dann noethersch, wenn die maximalen Ideale endlich erzeugt sind und jedes endliche Ideal endlicher Durchschnitt von Primäridealen ist.*

Gegenbeispiele zu Satz 1.9.

1.9.1) Die Forderung, daß jeder Untermodul *endlicher* Durchschnitt von primären Untermoduln ist, ist wichtig: Es gibt lokal noethersche Ringe R mit nur endlich erzeugten maximalen Idealen, die nicht noethersch sind, bei denen aber jeder Untermodul Durchschnitt von primären Untermoduln ist.

Beispiel. Sei k nichtarchimedisch lokal kompakt,

$$R \equiv \{f \mid f\colon E_n \to k,\ f \text{ lokal holomorph}\}.$$

Wie man leicht sieht, erfüllt R die Eigenschaft $\bar{\mathrm{I}}$ (s. [13]); somit ist R lokal noethersch, die maximalen Ideale sind endlich erzeugt und jedes Ideal ist Durchschnitt von primären Idealen.

1.9.2) Die Forderung, daß $\mathfrak{m}M$ endlich erzeugt für alle $\mathfrak{m}$ ist, ist wesentlich: Es gibt nämlich Ringe R, bei denen jeder Untermodul endlicher Durchschnitt von primären Moduln ist, die aber nicht noethersch sind.

Beispiel. In $R \equiv k[x_1, x_2, \ldots]/\mathfrak{m}^2$ mit $\mathfrak{m} = (x_1, x_2, \ldots)$ ist jedes Ideal primär. R ist jedoch nicht noethersch.

1.9.3) Für den Fall, daß R lokal ist, ist die Forderung „M endlich" nötig: Es gibt nicht noethersche Moduln M über lokalen noetherschen Ringen, so daß jeder Untermodul eine endliche Primärzerlegung hat und $\mathfrak{m}M$ endlich ist.

Beispiel. M wie in 2), $R := k[x_1]/(x_1)^2$.

Es sei noch erwähnt, daß unter Benutzung des Hilfssatzes 1.4 der Beweis von 1.8 und 1.9 leicht zum Beweis des folgenden Satzes abgeändert werden kann:

Sei R ein Ring mit endlich erzeugten maximalen Idealen. Ein lokal endlicher Modul M ist genau dann noethersch, wenn es zu jedem Untermodul $N \subset M$ und zu jedem maximalen Ideal $\mathfrak{m}_0 \subset R$ endlich viele maximale Ideale $\mathfrak{m}_1, \ldots, \mathfrak{m}_m \subset R$ gibt, so daß $\bigcap_{i=0}^{m} \bigcap_k N + \mathfrak{m}_i^k$ wieder ein endlicher Modul ist.

Folgenden Hilfssatz werden wir an anderer Stelle noch brauchen:

Hilfssatz 1.10. *Sei M ein R-Modul, $N \subset M$ ein Untermodul. Es gebe ein nicht-maximales Ideal $I \subset R$, das alle Nullteiler von M/N enthält. Dann gilt für jeden Untermodul $N_0 \subset M$:*

Ist $N \supsetneqq N_0$, so ist auch $N \supset \bigcup_{\substack{N_i/N_0 \\ \text{Artinsch}}} N_i$.

Beweis. Wir müssen zeigen: Ist $N_1 \supsetneqq N_0$, Länge $(N_1/N_0)=1$, so ist auch $N \supset N_1$. Sei $\mathfrak{m} \supset I$ ein maximales Ideal und $x \in \mathfrak{m} - I$. Wäre $N \not\supset N_1$, so wäre für $N^* := (N+N_1)/N$ auch Länge $N^*=1$. Es müßte deshalb nach Definition von x auch $x N^* = N^*$ sein; somit $x n_1 = n_2$ für $n_1, n_2 \in N^* - \{0\}$. Da $n_2 R = N^*$ ist, folgte $n_1 = n_2 r$ und somit $x \cdot n_2 \cdot r = n_2$, also $n_2(1 - x r) = 0$. Dann wäre also $1 - x r \in I \subset \mathfrak{m}$, also auch $1 \in \mathfrak{m}$.

Dieser Hilfssatz ist insbesondere für primäre Moduln $p \subset M$ mit nicht maximalem Ass M/p richtig; ganz speziell gilt also: Ist $q \supset p$ und q/p Artinsch, so ist $q = p$.

Hieraus folgt z.B. sofort: Hat in dem Ring R für jedes Ideal I die Menge $\{J \mid J \supsetneqq I\}$ ein minimales Element, so ist R nulldimensional. R braucht jedoch nicht Artinsch zu sein, wie Beispiel 1.9.2) zeigt.

§ 2. Charakterisierung einer Klasse von Moduln

Sei k ein komplett bewerteter Körper. Es sollen nun globale Ringe (d.h. Ringe von global holomorphen Funktionen) betrachtet werden. Die häufigsten behandelten Ringe dieser Art sind etwa die Ringe aller in einem Gebiet G holomorpher Funktionen oder die Ringe der in einer Umgebung eines kompakten Polyzylinders G holomorphen Funktion. Hier sollen darüber hinaus auch noch Ringe von Funktionen über Mengen G, die weder offen noch abgeschlossen sind, betrachtet werden. Ferner können wir noch folgende Verallgemeinerungen betrachten: z.B. kann ein und derselbe m-dimensionale Polyzylinder G als Teilmenge des k^m, aber auch des k^{m+1} betrachtet werden. Wir können somit auch die Ringe der in einer in k^{m+1} offenen Umgebung von G holomorphen Funktionen behandeln. Der Unterschied dieser Ringe zu dem üblicherweise betrachteten Ring entspricht dem Unterschied des Ringes $\mathcal{O}_z^{m+1}$ (das ist der Ring der in einer $m+1$ dimensionalen Umgebung von z holomorphen Funktionen) zu dem Ring $\mathcal{O}_z^m$. Diese Ringe $\mathcal{O}_z^m$ sind dann auch Spezialfälle der in dieser Arbeit untersuchten Ringe.

Wenn also G eine beliebige Teilmenge eines k-analytischen Raumes Z ist, betrachten wir folgende Ringe:

$$R_Z(G) := \{f \mid \text{Es gibt offenes } U \subset Z,\ U \supset G, \text{ so daß } f \text{ in } U \text{ holomorph ist}\}.$$

Oben haben wir gesehen, daß $R_Z(G)$ sowohl von Z als auch von G abhängt. Ist jedoch Z eine Mannigfaltigkeit, so hängt $R_Z(G)$ nur von der Dimension $n = \dim Z$ der Mannigfaltigkeit Z ab; wir schreiben dann statt $R_Z(G)$ einfach $R_n(G)$ und lassen meistens sogar – wenn es aus dem Zusammenhang klar ist – den Index n weg.

Wir betrachten nun Unterringe $R \subset R_Z(G)$. Nicht alle Unterringe sind funktionentheoretisch sinnvoll: Ist z.B. $G = Z = \mathbb{C}^2$, so ist $R := \mathbb{C}[x_1] \subset R(\mathbb{C}^2)$ ein solcher Unterring. Die Struktur dieses Ringes steht jedoch in keinem engen Zusammenhang mit der Struktur des $\mathbb{C}^2$: Es gibt z.B. keine Funktionen f_i aus R, die im Nullpunkt ein lokales Koordinatensystem in $\mathbb{C}^2$ bilden. Der Ring R ist dagegen als Unterring von $R(\mathbb{C}^1)$ auch funktionentheoretisch sinnvoll. Es erscheint also nicht geraten, alle Unterringe R von $R(\mathbb{C}^2)$ zu betrachten; vielmehr behandelt man z.B. „zu kleine" Unterringe R besser als Unterringe des $R(\mathbb{C}^1)$. Damit der Ring R „groß genug" ist, fordern wir, wenn Z Teilmenge des k^n ist: R enthalte neben dem Grundkörper k die auf Z eingeschränkten holomorphen Funktionen z_i, wobei z_i die Koordinaten des k^n sein sollen. Ist nun Z irgendein k-analytischer Raum endlicher Einbettungsdimension, so bette man Z vermöge einer Abbildung φ in einen k^n ein und verlange, daß die Projektionen $\varphi_i(z)$ von $\varphi(z)$ auf die einzelnen Koordinaten des k^n auch in R sind. Dies braucht nur für eine einzige feste Abbildung φ gelten; somit gehört also z.B. auch der für viele f, g ebenfalls funktionentheoretische sinnvolle Ring $R := \mathbb{C}[f(x_1), g(x_2)] \subset R(\mathbb{C}^2)$ zu dieser Klasse. Allgemein wird man in einem Ring $R \subset R_Z(G)$ mit hinreichend vielen Funktionen eine solche Abbildung $\varphi: Z \to k^n$ finden können. Es erscheint deshalb sinnvoll, zu definieren:

Ein Ring R mit $k \subset R \subset R_Z(G)$ ist ein analytischer Ring zur Menge $G \subset Z$, wenn es eine Einbettung $\varphi: Z \to k^m$ gibt, so daß die in $R_Z(G)$ liegenden Projektionen φ_i von φ auf die einzelnen Koordinaten des k^m auch in R liegen.

In analytischen Ringen gibt es nach Definition zu jedem $z^0 \in G$ endlich viele Funktionen $f_i \in R$ mit $(f_1, \ldots, f_n)\mathcal{O}_z = \mathcal{O}_z \ \forall z \neq z^0$ und $(f_1, \ldots, f_n)\mathcal{O}_{z^0} = \mathfrak{m}(\mathcal{O}_{z^0})$, nämlich gerade die Funktionen $\varphi_i(z) - \varphi_i(z^0)$, wobei φ_i die nach Definition in R liegenden Projektionen von $\varphi: Z \to k^m$ auf die einzelnen Koordinaten sein sollte. In fast allen folgenden Sätzen genügt es, die Existenz solcher Funktionen f_i aus R zu gegebenem $z^0 \in G$ zu fordern, ohne daß wir eine Einbettung $\varphi: Z \to k^m$ voraussetzen müssen. Somit gehören also für alle Steinschen Räume Z und alle Mengen $G \subset Z$ die Ringe $R_Z(G)$ zu unserer betrachteten Kategorie; ebenso Ringe, die Lokalisationen und Faktorringe davon sind (wird die stärkere Voraussetzung gebraucht, so gilt dies nur für alle Steinschen Räume Z mit beschränkter Einbettungsdimension, also insbesondere für alle Steinschen Mannigfaltigkeiten Z). Speziell aber gehören bei $G \subset Z \subset k^m$ für beliebiges G, Z alle Ringe $R \subset R_Z(G)$, die den Polynomring $k[z_1, \ldots, z_m]$ enthalten, zu dieser Klasse.

Wir bezeichnen nun die zu $z^0 \in G$ gehörenden oben erwähnten Funktionen f_i (in Analogie zu dem Fall, daß $Z \subset k^m$ ist) mit $z_i - z_i^0$; für das Ideal $(f_1, \ldots, f_n) = (z_1 - z_1^0, \ldots, z_n - z_n^0)$ schreiben wir kurz $(z - z^0)$. Dabei heißen der aus den $z_i - z_i^0$ erzeugte Unterring von R der „Polynomring in R". Man beachte jedoch, daß auch bei $G = Z \subset k^n$, selbst wenn wir jetzt für f_i wirklich die Einschränkungen der Koordinatenfunktionen z_i auf G wählen, der Ring $k[z_1, \ldots, z_m]$ nicht zum freien Polynomring in m Variablen isomorph zu sein braucht: Denn zwei Elemente aus $k[z_1, \ldots, z_m] \subset R$ sind schon gleich, wenn sie auf Z übereinstimmen; somit ist etwa bei niederdimensionalen Steinschen Mannigfaltigkeiten oder gar bei Steinschen Räumen Z, die im k^m eingebettet sind, der Polynomring aus $R_Z(Z)$ nie isomorph zum freien Polynomring in n Variabeln.

Sodann wollen wir Moduln über R betrachten. Funktionentheoretisch sinnvoll sind nur Moduln, die im Schnittmodul einer kohärenten Garbe enthalten sind. Wir definieren also:

Sei $R \subset R_Z(G)$ ein analytischer Ring. Ein R-Modul M heiße analytisch zur Menge $G \subset Z$, wenn es eine kohärente Garbe $\mathcal{M}$ über einer in Z offenen Umgebung U von G gibt, so daß M ein R-Untermodul von $\Gamma(G, \mathcal{M})$ ist.

Ist $(\mathcal{O}_z)_{z \in Z}$ die Strukturgarbe von Z, so ist für ein $f \in \Gamma(G, \mathcal{M})$ der Ausdruck $f\mathcal{O}_z \subset \mathcal{M}_z$ wohldefiniert; somit ist also $M\mathcal{O}_z$ erklärt (schon aus diesem Grunde mußten wir $M \subset \Gamma(G, \mathcal{M})$ voraussetzen). Wir definieren nun:

Definition 2.1. Sei M ein analytischer R-Modul zur Menge $G \subset Z$. Ein R-Untermodul $N \subset M$ hat die Eigenschaft (I) bzw. (II) bzw. (III) bezüglich $G \subset Z$, wenn es in G Punkte $\{z_i\}_{i \in I}$ gibt (wobei I eine beliebige bzw. abzählbare bzw. endliche Indexmenge ist), so daß gilt:

$$\{f \in M \mid f \in N\mathcal{O}_{z_i} \text{ für alle } i \in I\} = N.$$

Ist $M \subset \Gamma(G, \mathcal{M})$ ein analytischer Modul, so sei für eine Teilmenge $\mathcal{N}_z \subset \mathcal{M}_z$ der Durchschnitt $M \cap \mathcal{N}_z$ als die Menge der $f \in M$ mit $f\mathcal{O}_z \subset \mathcal{N}_z$ definiert; ist die Abbildung $M \to \mathcal{M}_z$ injektiv, so ist dies die übliche Durchschnittsdefinition.

Die Varietät $\operatorname{Var} N$ eines Untermoduls $N \subset M$ sei die Menge der z aus G mit $N\mathcal{O}_z \neq M\mathcal{O}_z$; ist $M = R$, so ist dies natürlich die gemeinsame Nullstellenmenge aller Funktionen aus N. Speziell ist noch $\operatorname{Var}\{0\} = \operatorname{Tr} M$.

Dann gilt zunächst:

Hilfssatz 2.2. *Ist M ein analytischer Modul zur Menge $G \subset Z$ und ist $N \subset M$ primär, so gilt:*

Ist $f \in R$ Nullteiler in M/N, so ist $f(z) = 0$ $\forall z \in \operatorname{Var} N$.

Beweis. Wäre $f(z_0) \neq 0$ für ein $z_0 \in \operatorname{Var} N$, so wäre für alle k $f^k M\mathcal{O}_{z_0} = M\mathcal{O}_{z_0}$. Da f Nullteiler mod N ist und N primär vorausgesetzt ist, ist $f^k M \subset N$ für großes k. Somit ist $M\mathcal{O}_{z_0} = f^k M\mathcal{O}_{z_0} \subset N\mathcal{O}_{z_0}$, also nicht $N\mathcal{O}_{z_0} \subsetneqq M\mathcal{O}_{z_0}$.

Hilfssatz 2.3. *Sei M ein analytischer Modul zu $G \subset Z$. Ist $N \subset M$ ein Untermodul mit* (I) *bzw.* (II) *bzw.* (III), *so ist $N = \bigcap_{i \in I} N_i$ mit primären Untermoduln N_i, die alle eine nicht leere Varietät haben. Dabei ist I eine beliebige bzw. abzählbare bzw. endliche Indexmenge.*

Beweis. Es ist $N\mathcal{O}_{z_i}$ ein Untermodul von $M\mathcal{O}_{z_i} \subset \mathcal{M}_{z_i}$. Da $\mathcal{O}_{z_i}$ noethersch ist und $\mathcal{M}_{z_i}$ endlicher $\mathcal{O}_{z_i}$ Modul ist, ist $N\mathcal{O}_{z_i} = \bigcap_{j=1}^{n_i} N_j^i$, N_j^i primär in $M\mathcal{O}_{z_i}$. Dann ist nach Voraussetzung $N = \bigcap_{i \in I} \bigcap_{j=1}^{n_i} N_j^i \cap M$, und weiter ist $N_j^i \cap M$ primär Untermodul von M. Hat $N_j^i \cap M$ eine leere Varietät, so ist $N_j^i \cap M = M$ und kann daher beim Durchschnitt weggelassen werden.

Hilfssatz 2.3′. *Sei $R \subset R(G)$ ein Unterring. Ist $J \subset R$ ein reduziertes Ideal mit* (I), *so ist J Durchschnitt von Primidealen mit Nullstelle in G.*

Beweis. Ist $z \in \operatorname{Var} J$ und $\mathfrak{m}(z)$ das maximale Ideal aller in $z \in G$ verschwindenden Funktionen, so ist $JR_{\mathfrak{m}(z)} \subset R_{\mathfrak{m}(z)}$ reduziert und $JR_{\mathfrak{m}(z)} = \bigcap_{i \in I} p_i^z R_{\mathfrak{m}(z)}$ mit o. B. d. A. $p_i^z \subset R$ prim und $z \in \operatorname{Var} p_i^z$. Dann ist $J = \bigcap_{z \in \operatorname{Var} J} \bigcap_{i \in I} p_i^z$.

Weiter sei noch folgender leicht zu beweisender Satz genannt:

Hilfssatz 2.4. *Sei M ein analytischer R-Modul zu $G \subset Z$. Seien $N_i \subset M$ Untermoduln mit* (I) *bzw.* (II) *bzw.* (III). *Dann ist für jede beliebige bzw. jede abzählbare bzw. jede endliche Indexmenge I auch $\bigcap_{i \in I} N_i$ ein Untermodul mit* (I) *bzw.* (II) *bzw.* (III).

Beweis. Es gibt $(z_j^i)_{j \in J_i}$ mit $\{f \in M \mid f \in N_i \mathcal{O}_{z_j^i} \ \forall j \in J_i\} = N_i$. Sei dann $f \in M$, $f \in N \mathcal{O}_{z_j^i}$ für alle $j \in J_i$ und $i \in I$. Dann ist erst recht $f \in N_i \mathcal{O}_{z_j^i} \ \forall j \in J_i$, woraus $f \in N_i$ folgt. Somit ist $f \in N$.

Wir können nun wegen 2.4 zu jedem Untermodul $N \subset M$ den Abschluß $\overline{N}$ als den Durchschnitt aller größeren Untermoduln mit (I) definieren. Ein Untermodul ist also genau dann abgeschlossen, wenn er (I) erfüllt; somit ist also M und $\{0\}$ abgeschlossen; und der Durchschnitt einer Familie von abgeschlossenen Untermoduln ist wieder abgeschlossen.

Die Umkehrung von 2.1 ist natürlich falsch. Um sie geeignet formulieren zu können, definieren wir:

Definition 2.5. Ein analytischer R-Modul M hat die Eigenschaft I bzw. II bzw. III bezüglich $G \subset Z$, wenn alle *endlichen* R-Untermoduln $N \subset M$ die Eigenschaft (I) bzw. (II) bzw. (III) haben.

Ein analytischer R-Modul M hat die Eigenschaft $\bar{\mathrm{I}}$ bzw. $\overline{\mathrm{II}}$ bzw. $\overline{\mathrm{III}}$ bezüglich $G \subset Z$, wenn *alle* R-Untermoduln $N \subset M$ die Eigenschaft (I) bzw. (II) bzw. (III) haben.

Ein R-Modul M, dessen Elemente Schnitte einer kohärenten Garbe $\mathscr{M}$ über G sind, erfüllt also genau dann I bezüglich G, wenn jeder Restklassenmodul nach einem endlichen Untermodul wieder als Menge von Schnitten einer kohärenten Garbe $\mathscr{M}'$ über G aufgefaßt werden kann.

Beispiel. Ist $R = R_Z(G)$, wobei G eine Steinsche Menge ist (d.h. G besitzt ein Umgebungssystem von Steinschen Räumen $G_i \subset Z$), und ist $\mathscr{M}$ eine kohärente Garbe in einer Umgebung von G, so erfüllt $M := \Gamma(G, \mathscr{M}) \equiv \lim \Gamma(G_i, \mathscr{M})$ die Bedingung I.

Beweis. Ist $N \subset M$ ein endlicher Untermodul, so gibt es eine Umgebung $U \supset G$ und endlich viele Schnitte $(f_i) \in N$ von $\mathscr{M}$ über U, so daß $(f_i)\mathcal{O}_z = N\mathcal{O}_z$ $\forall z \in G$ ist. Ist dann $f \in M$, $f \in N\mathcal{O}_z$ $\forall z \in G$, so ist $f \in (f_i)\mathcal{O}_z$ $\forall z \in G$; dann gibt es ein offenes U', $G \subset U' \subset U$ mit $f \in (f_i)\mathcal{O}_z$ $\forall z \in U'$ und somit auch einen Steinschen Raum U'' mit $f \in (f_i)\mathcal{O}_z$ $\forall z \in U''$. Dann folgt nach Theorem B, daß $f \in (f_i) \cdot R(U'')$ ist und damit $f \in (f_i) R \subset NR$ ist.

Weiter brauchen wir noch einen trivialen Hilfssatz

Hilfssatz 2.6. *Sei M ein analytischer Modul zu $G \subset Z$.*

1. *Ist* $m \in N\mathcal{O}_{z^0} \cap M$, *so ist für jedes* k *schon* $f \in N + M_k$, *wobei* $M_k := (z-z^0)^k M \mathcal{O}_{z^0} \cap M$ *ist.*

2. *Erfüllt* $(z-z^0)^k M$ *zusätzlich* (I), *so folgt sogar* $f \in N + (z-z^0)^k M$.

Beweis 1. Sei $m \in N\mathcal{O}_{z^0} \cap M$. Es ist also $m = \sum_1^q n_i o_i$ mit $n_i \in N$ und $o_i \in \mathcal{O}_{z^0}$. Wir schreiben $o_i = o_i' + o_i''$, wobei $o_i' \in R$ ist und $o_i'' \in (\mathfrak{m}(\mathcal{O}_{z^0}))^k$ ist. Dann ist $m = m' + m''$, wobei $m' = \sum n_i o_i' \in N$ ist und $m'' = \sum n_i o_i'' \in (z-z^0)^k M \mathcal{O}_{z^0}$ ist.

Beweis 2. Wegen $(z-z^0)^k M \mathcal{O}_z = M\mathcal{O}_z$ $\forall z \neq z^0$ folgt die Behauptung dann aus 1.

Jetzt können wir zeigen:

Satz 2.6. *Sei* M *ein analytischer* R*-Modul zur Menge* $G \subset Z$. *Dann ist äquivalent*

1) M *erfüllt* $\bar{\mathrm{I}}$ *bzw.* $\overline{\mathrm{II}}$ *bzw.* $\overline{\mathrm{III}}$.

2) a) *Jeder Untermodul* $N \subset M$ *ist beliebiger bzw. abzählbarer bzw. endlicher Durchschnitt von primären Untermoduln mit nichtleerer Varietät.*

b) *Jeder Untermodul* $N \subset M$ *der Form* $N = (z-z^0)^m M$ *erfüllt* (I).

Sind diese Bedingungen erfüllt, so gilt weiter: Ist N *primärer Untermodul,* $z^0 \in \operatorname{Var} N$, *so folgt aus* $f \in M$, $f \in N\mathcal{O}_{z^0}$, *daß schon* $f \in N$ *ist. Insbesondere erfüllt also jeder primäre Untermodul* (III).

Beweis. 1) $\to$ 2): Folgt aus Satz 2.3.

2) $\to$ 1): Wir zeigen zunächst, daß die zusätzliche Behauptung richtig ist: Sei $J = (z-z^0)R$. Nach Hilfssatz 1.1 ist

$$\bigcap_k N + J^k M = \{f \in M \mid f(1-i) \in N \text{ für ein } i \in J\}.$$

Nach 2.2 folgt daraus, da $1-i$ in z^0 keine Nullstelle hat, daß $\bigcap_k N + J^k M = N$ ist.

Nach 2.6.2 folgt aus $f \in N\mathcal{O}_{z^0} \cap M$, daß schon $f \in N + (z-z^0)^k M$ für jedes k ist, woraus nach obigem $f \in N$ folgt. Damit ist die Zusatzbehauptung gezeigt.

Die Behauptung 1) ergibt sich nun direkt aus 2.4.

Wir wollen nun untersuchen, wann ein Modul M die Eigenschaften I, II und III erfüllt. Diese Frage ist wichtiger als die gerade behandelte.

Satz 2.7. *Sei* M *ein analytischer* R*-Modul. Dann ist äquivalent:*

1) M *erfüllt* I *bzw.* II *bzw.* III.

2) a) *Jeder endliche Untermodul* $N \subset M$ *ist beliebiger bzw. abzählbarer bzw. endlicher Durchschnitt von primären Untermoduln mit nichtleerer Varietät.*

b) *Jeder Untermodul* $N \subset M$ *der Form* $N = (z-z^0)^k M$ *erfüllt* (I).

Beweis. 1) $\to$ 2) a) folgt aus Satz 2.3.

b) Es ist $M\mathcal{O}_z$ ein Untermodul von $\mathcal{M}_z$. Es gibt also, da $\mathcal{M}_z$ noetherscher $\mathcal{O}_z$-Modul ist, $f_1, \ldots, f_m \in M$ mit $(f_1, \ldots, f_m)\mathcal{O}_z = M\mathcal{O}_z$. Sei wieder $J = (z-z^0)R$.

Ist $f \in J^k M \mathcal{O}_{z^0}$, so ist also $f \in J^k(f_1, \ldots, f_m)\mathcal{O}_{z^0}$. Klar ist weiter $f \in f \cdot J^k \mathcal{O}_z \ \forall z \neq z^0$. Somit ist $f \in (f, f_1, \ldots, f_m) J^k \mathcal{O}_z \ \forall z \in G$, woraus $f \in (f, f_1, \ldots, f_m) J^k$ und damit $f \in M J^k$ folgt.

2)$\to$1) Sei $N = \bigcap_{i \in I} N_i$, $z^i \in \operatorname{Var} N_i$. Sei $f \in N \mathcal{O}_{z^i} \cap M$ für alle $i \in I$. Aus 2.6.2 folgt daraus für alle i und alle k die Beziehung $f \in N + J_i^k M$ mit $J_i \equiv (z - z^i)$. Aus $f \in \bigcap_k N + J_i^k M$ folgt aber nach 1.2: $f \cdot (1 - j_i) \in N$ für ein $j_i \in J_i$. Somit ist $f \cdot (1 - j_i) \in N_i$, also nach Hilfssatz 2.2 $f \in N_i$ und damit $f \in N$. Also ist

$$\bigcap_{i \in I} N \mathcal{O}_{z^i} \cap M = N.$$

Bemerkung. Wie aus dem Beweis hervorging, können wir im Satz 2.6 *und Satz* 2.7 *statt* 2a) *auch fordern:*

Es gilt $\bigcap_{i \in I} \bigcap_k N + (z - z^i)^k M = N$, *wobei I eine beliebige bzw. abzählbare bzw. endliche Indexmenge ist.*

Die Methoden des Beweises von 2.6 und 2.7 führen zu einem etwas allgemeineren Ergebnis, das später gebraucht werden wird. Dazu wollen wir nun immer abkürzend folgende Definition verwenden:

Ist $z^0 \in G$, so sei $\mathfrak{m}(z^0)$ das maximale Ideal aller in z^0 verschwindenden Funktionen aus R und M_{z^0} sei der nach $\mathfrak{m}(z^0)$ lokalisierte $R_{\mathfrak{m}(z^0)}$-Modul $M_{\mathfrak{m}(z^0)}$.

Dann gilt:

Hilfssatz 2.7. *Sei M ein analytischer Modul von* $G \subset Z$.

1. *Ist* $z^0 \in G$, M_{z^0} *noetherscher* R_{z^0}*-Modul und gibt es zu jedem* $k \in \mathbb{N}$ *ein* q, *so daß* $\overline{(z - z^0)^q M} \subset (\mathfrak{m}(z^0))^k M$ *ist (erfüllt also insbesondere* $(z - z^0)^k M$ *die Eigenschaft* (I) *bezüglich G), so ist für alle* $N \subset M$ *schon* $N \mathcal{O}_{z^0} \cap M = N R_{z^0} \cap M$.

2. *Ist* $G' \subset G$ *und sind für alle* $z \in G'$ *die obigen Voraussetzungen erfüllt, so folgt aus* $m \in N \mathcal{O}_z \ \forall z \in G'$, *daß* $m J \subset N$ *ist für ein Ideal* $J \subset R$ *mit* $\operatorname{Var} J \cap G' = \emptyset$ *und* $J \supset \operatorname{Ann} M$.

Beweis 2.7.1. Sei $m \in N \mathcal{O}_{z^0}$. Nach 2.6.1 folgt daraus für jede Zahl q eine Darstellung $m = n_q + m_q$ mit $n_q \in N$, $m_q \in \overline{(z - z^0)^q M \mathcal{O}_{z^0} \cap M}$. Es gibt somit zu jedem k ein $n \in N$ und $m' \in \mathfrak{m}(z^0)^k M$ mit $m = n + m'$. Somit ist $m \in N + (\mathfrak{m}(z^0))^k M$, also $m R_{z^0} \subset \bigcap_k N R_{z^0} + (\mathfrak{m}(z^0))^k M_{z^0}$. Da M_{z^0} noethersch ist, gibt es ein endliches Ideal $J \subset \mathfrak{m}(z^0)$ mit $J M_{z^0} = \mathfrak{m}(z^0) M_{z^0}$. Dann ist $m R_{z^0} \subset \bigcap_k N R_{z^0} + J^k M_{z^0}$; nach 1.1 folgt $m \in N R_{z^0}$.

2.7.2. Nach 2.7.1 gibt es zu jedem $z \in G'$ ein $s \in R$, $s(z) \neq 0$ mit $m s \in N$. Dann sind für die Summe J aller dieser Hauptideale sR die Beziehungen $J m \subset N$ und $\operatorname{Var} J \cap G' = \emptyset$ erfüllt.

Es sei hier bemerkt, daß wir bei allen unseren Sätzen immer statt der Forderung, daß $(z - z^0)^k M$ (I) erfüllt, nur die schwächere Bedingung $\overline{(z - z^0)^q M} \subset (\mathfrak{m}(z^0))^k M$ vorauszusetzen brauchten; der Beweis wird dazu entsprechend 2.7.1 abgeändert.

Für das Weitere brauchen wir noch eine leichte Verallgemeinerung von 2.7.1. Wir können nämlich leicht definieren, was $m \in N\hat{\mathcal{O}}_z$ heißt, wobei $\hat{\mathcal{O}}_z$ die Komplettierung von $\mathcal{O}_z$ ist: Ist F ein freier $\mathcal{O}_z$-Modul mit $F \xrightarrow{\varphi} M\mathcal{O}_z \to 0$, so seien m' bzw. N' irgendwelche Urbilder von m bzw. N in F; dann ist die Aussage $m' \in (N' + \operatorname{Ker}\varphi)\hat{\mathcal{O}}_z$ wohldefiniert, da $N' + \operatorname{Ker}\varphi$ Teilmenge eines freien $\mathcal{O}_z$-Moduls ist. Wenn diese Aussage für irgendeinen freien $\mathcal{O}_z$-Modul F mit $F \to M\mathcal{O}_z \to 0$ erfüllt ist, so sei per Definition $m \in N\hat{\mathcal{O}}_z$. Es gilt dann:

2.7.3. *Sei $m \in M$, $N \subset M$ und $m \in N\hat{\mathcal{O}}_{z_0}$ für ein $z_0 \in G$. Dann ist $m \in N\mathcal{O}_{z_0}$.*

Beweis. Für freie Moduln folgt aus $m \in M\mathcal{O}_{z_0}, N \subset M\mathcal{O}_{z_0}$, und $m \in N\hat{\mathcal{O}}_{z_0}$ wie in 2.6.1, daß schon $m \in N + M_k$ ist, wobei $M_k := (z - z_0)^k M\hat{\mathcal{O}}_{z_0} \cap M$ ist. Nun ist aber, da M freier $\mathcal{O}_{z_0}$-Modul ist, $M_k = (z - z_0)^k M\mathcal{O}_{z_0}$, womit nach dem Krullschen Durchschnittssatz die Behauptung folgt. Für beliebige Moduln folgt die Behauptung nach Definition von $M\hat{\mathcal{O}}_{z_0}$.

Satz 2.8. *Sei M ein analytischer R-Modul mit I. Dann ist für jedes $z_0 \in G$ der lokalisierte Modul M_{z_0} noethersch über R_{z_0}.*

Beweis. Sei $N \subset M$ ein Untermodul. Es gibt, da $M\mathcal{O}_{z_0}$ noetherscher $\mathcal{O}_{z_0}$-Modul ist, endliche viele $f_i \in N$ mit $(f_1, \ldots, f_m)\mathcal{O}_{z_0} = N\mathcal{O}_{z_0}$. Ist $f \in N\mathcal{O}_{z_0} \cap M$, so ist wegen $\mathfrak{m}(z_0)\mathcal{O}_z = \mathcal{O}_z \ \forall z \neq z_0$ auch $f \in \big((f_1, \ldots, f_m) + f \cdot \mathfrak{m}(z_0)\big)\mathcal{O}_z \ \forall z \in G$; somit ist also $f \in (f_1, \ldots, f_m) + f \cdot \mathfrak{m}(z_0)$ und damit $f(1-m) \in (f_1, \ldots, f_m)$ für ein $m \in \mathfrak{m}(z_0)$. Es ist also $N\mathcal{O}_{z_0} \cap M \subset (f_1, \ldots, f_m) R_{z_0}$, woraus sofort $NR_{z_0} = (f_1, \ldots, f_m) R_{z_0}$ folgt.

Diesen für das Weitere sehr wichtigen Satz können wir im Fall $R = R(G)$, $M = \Gamma(G, \mathcal{M})$, G Steinsch auch ohne Theorem B (das ja die Eigenschaft I für M impliziert), aber unter Zuhilfenahme von Theorem A beweisen: Sei $m_i \in \Gamma(G, \mathcal{M})$ und $m_0 = m_1 o_1 + \cdots + m_q o_q$ mit $o_i \in \mathcal{O}_{z_0}$. Dann ist

$$f := (-1, o_1, \ldots, o_q) \in \operatorname{Rel}_{\mathcal{O}_{z_0}}(m_0, m_1, \ldots, m_q).$$

Da die Relationengarbe kohärent ist, gibt es nach Theorem A endliche viele Schnitte $f_i = (f_i^0, \ldots, f_i^q)$ mit $f = \sum_1^k o_i'(f_i^0, \ldots, f_i^q)$. Insbesondere muß ein f_i^0 in z_0 keine Nullstelle haben. Für dieses i ist also $m_0 f_i^0 = m_1 f_i^1 + \cdots + m_q f_i^q$, woraus $m_0 \in (m_1, \ldots, m_q) R_{z_0}$ folgt. Somit haben wir also für jedes endliche $N \subset M$ gezeigt: $N\mathcal{O}_{z_0} \cap MR_{z_0} = NR_{z_0}$, womit aus $M\mathcal{O}_{z_0}$ noethersch die Behauptung folgt.

Mit der gerade gezeigten Formel $N\mathcal{O}_{z_0} \cap MR_{z_0} = NR_{z_0}$ können wir weiterhin sofort viele wichtige Beziehungen (z.B. [13], Satz 6 oder Satz 28) ohne Benutzung von Theorem B beweisen.

Insbesondere gilt also, daß $\big(R(G)\big)_z$ für alle $z \in G$ noethersch ist. Dies ist auch richtig, wenn $G \subset k^m$ *nicht* Steinsch ist; auch in diesem Fall gilt folgende schärfere Aussage: Ist $f_i \in R(G)$, $f_0 \in (f_1, \ldots\ldots, f_n)\mathcal{O}_{z_0}$, so ist $f_0 \in (f_1, \ldots, f_n) R_{z_0}$. Für Steinsche Mengen G wurde dies gerade in 2.8 mitbewiesen; wie in [13], Hilfssatz 34 folgt dann die allgemeine Behauptung.

Satz 2.9. *Sei M ein analytischer R-Modul. Dann ist äquivalent:*

1) *M erfüllt* $\bar{\mathrm{I}}$.

2) a) *M erfüllt* I.
   b) *Für jedes $N \subsetneqq M$ ist* $\operatorname{Var} N \neq \emptyset$.
   c) *M ist lokal endlicher R-Modul.*

3) a) *M ist lokal noethersch.*
   b) *Für jedes $N \subsetneqq M$ ist* $\operatorname{Var} N \neq \emptyset$.
   c) *Die Untermoduln der Form $(z-z_0)^k M$ erfüllen* (I).

4) a) *M_z ist noethersch $\forall z \in G$.*
   b) *Ist $J \subset R$ ein Ideal ohne Nullstelle in G, so ist $m \cdot J = m \cdot R$ $\forall m \in M$.*
   c) *Die Untermoduln $(z-z_0)^k M$ erfüllen* (I).

Beweis. 2)→1) M erfülle I. Sei $\mathfrak{m}$ ein maximales Ideal mit einer Nullstelle $z_0 \in G$. Dann ist nach 2.8 $M_{\mathfrak{m}}$ noethersch. Ist aber $\mathfrak{m}$ ein maximales Ideal ohne Nullstelle in G, so ist $\operatorname{Var} \mathfrak{m} M = \emptyset$, woraus nach Voraussetzung $\mathfrak{m} M = M$ folgt. Daraus folgt nach dem Nakayama-Lemma, da M lokal endlich ist, die Beziehung $M_{\mathfrak{m}} = 0$. Insgesamt folgt also, daß M lokal noethersch ist; wie 1.7. 3)→1) folgt dann, daß *jeder* Untermodul eine Primärzerlegung hat. Nach Voraussetzung ist außerdem $\operatorname{Var} N_i \neq \emptyset$. Dann folgt nach 2.6, daß M $\bar{\mathrm{I}}$ erfüllt.

1)→2) b) Wäre $\operatorname{Var} N = \emptyset$, so wäre $N \mathcal{O}_z = M \mathcal{O}_z$ für alle $z \in G$, also wegen der Eigenschaft $\bar{\mathrm{I}}$ auch $N = M$.

c) Ist $\mathfrak{m}$ ein maximales Ideal ohne Nullstelle, so ist $\mathfrak{m} \cdot x \cdot \mathcal{O}_z = x \cdot \mathcal{O}_z$ für alle $z \in G$ und für alle $x \in M$. Wegen der Eigenschaft $\bar{\mathrm{I}}$ folgt daraus $\mathfrak{m} \cdot x = x$ für alle $x \in M$. Das heißt aber, daß $M_{\mathfrak{m}} = 0$ ist. Für maximale Ideale $\mathfrak{m}$ mit Nullstelle ist $M_{\mathfrak{m}}$ nach 2.8 endlicher $R_{\mathfrak{m}}$-Modul. Insgesamt folgt also, daß M lokal endlich ist.

2)→3) In 2)→1) wurde schon gezeigt, daß M lokal noethersch ist.

3)→1) Wenn M lokal noethersch ist, hat jeder Untermodul N eine Primärzerlegung $N = \bigcap p_i$ (s. Beweis von 1.7. 3)→1)). Dann folgt nach 2.6 die Behauptung

3)→4) 4b) wird über den Umweg 3)→1)→4b) gezeigt; 1)→4b) ist aber trivial.

4)→1) Nach 2.7.2 folgt aus $m \in N \mathcal{O}_z$ $\forall z \in G$ schon $mJ \subset N$ mit $\operatorname{Var} J = \emptyset$. Also folgt aus 4b) schon $m \in N$.

Es folgt nun eine Charakterisierung für Moduln mit III.

Satz 2.10. *Sei M ein analytischer Modul. Dann ist äquivalent:*

1) *M erfüllt* I *und ist noethersch.*

2) *M erfüllt* III.

3) *M erfüllt* $\overline{\mathrm{III}}$.

4) a) *M ist noethersch.*
   b) *Für jedes $N \subsetneqq M$ ist* $\operatorname{Var} N \neq \emptyset$.
   c) *Die Untermoduln der Form $(z-z_0)^k M$ erfüllen* (I).

5) a) *M ist noethersch.*

b) *Ist $J \subset R$ ein Ideal ohne Nullstelle in G, so ist $mJ = mR$ für alle $m \in M$.*

c) *Die Untermoduln der Form $(z - z_0)^k M$ erfüllen* (I).

Beweis. 2)→3) M erfülle III. Nach 2.6 hat jeder endliche Untermodul $N \subset M$ eine Primärzerlegung $N = \bigcap_1^m p_i$, so daß $\operatorname{Var} p_i \neq \emptyset$ ist. Sei $z_i \in \operatorname{Var} p_i$. Nach 2.8 ist M_{z_i} noetherscher R_{z_i}-Modul. Sei nun wie in Bemerkung 1.9 $\{q_i\} = \operatorname{Ass} M/p_i$ gegeben und sei $q_0 \subsetneqq q_1 \subsetneqq q_2 \subsetneqq \cdots$ eine solche echt aufsteigende Kette. Es gibt eine Zahl m, so daß alle Ideale der Form $(z - z_0)$ durch m Elemente erzeugt werden. Wir behaupten, daß auch jede solche Kette nach m Gliedern abbrechen muß:

Es ist zunächst $R/q_0 \subset M/p_0$, da q_0 zum Primärmodul p_0 assoziiert sein sollte. Es gibt somit $N \subset M$, $N \supset p_0$, so daß $N/p_0 = R/q_0$ ist. Da der Modul N trivialerweise auch I erfüllt, ist nach 2.7.2b) schon $(z - z_0) N = \mathfrak{m}(z_0) N$. Somit ist auch $(z - z_0) R/q_0 = \mathfrak{m}(z_0) R/q_0$. Lokalisieren wir nun nach dem maximalen Ideal $\mathfrak{m}(z_0)$, so wird deshalb das maximale Ideal von $(R/q_0)_{\mathfrak{m}(z_0)}$ von den Restklassen von $(z - z_0)$ erzeugt. Ferner ist nach 2.8 $M_{\mathfrak{m}(z_0)}$ noetherscher $R_{\mathfrak{m}(z_0)}$-Modul; daraus folgt, daß auch $(R/q_0)_{\mathfrak{m}(z_0)}$ noetherscher $R_{\mathfrak{m}(z_0)}$-Modul und damit noetherscher Ring ist. Da das maximale Ideal nach dem gerade Gesagten von m Elementen erzeugt wird, ist Krulldim $(R/q_0)_{\mathfrak{m}(z_0)} \leqq m$. Wäre nun z_0 eine Nullstelle von q_{m+1} (nach 2.2 ist wegen $\operatorname{Var} p_{m+1} \neq \emptyset$ auch $\operatorname{Var} q_{m+1} \neq \emptyset$), so gäbe es die Primidealkette $\overline{(q_0)}_{\mathfrak{m}(z_0)} \subsetneqq \overline{(q_1)}_{\mathfrak{m}(z_0)} \subsetneqq \cdots \subsetneqq \overline{(q_{m+1})}_{\mathfrak{m}(z_0)}$. Dies steht im Widerspruch zu Krulldim $(R/q_0)_{\mathfrak{m}(z_0)} \leqq m$, womit die Zwischenbehauptung gezeigt ist.

Die Behauptung selbst ergibt sich jetzt sehr leicht aus Bemerkung 1.8: Denn jedes endliche $N \subset M$ hat eine endliche Primärzerlegung $\bigcap p_i$; für $z_i \in \operatorname{Var} p_i$ umfassen die maximalen Ideale $\mathfrak{m}(z_i)$ die Primideale $q_i := \operatorname{Ass} M/p_i$; es ist $M_{\mathfrak{m}(z_i)}$ noetherscher $R_{\mathfrak{m}(z_i)}$-Modul und deshalb auch M_{q_i} noetherscher R_{q_i}-Modul; ferner gibt es keine aufsteigende Kette von solchen Primidealen q_i. Somit sind alle Voraussetzungen zu Bemerkung 1.8 erfüllt; M ist also noethersch, Ein noetherscher Modul mit III erfüllt aber trivialerweise auch $\overline{\text{III}}$.

3)→4) a) wurde in 2)→3) gezeigt, b) folgt aus 2.9.

4)→1) folgt aus 2.7.

1)→2) Jeder Untermodul N hat eine nichtleere Varietät: Denn wäre $N\mathcal{O}_z = M\mathcal{O}_z$ $\forall z \in G$, so wäre, da N endlich erzeugt ist und damit (I) erfüllt, schon $M \subset N$. Somit ist jeder Untermodul N endlicher Durchschnitt von primären Moduln mit nichtleerer Varietät. Nach 2.7 folgt die Behauptung.

5)→1) Nach 2.9 erfüllt M $\bar{\text{I}}$.

1)→5) M erfüllt auch $\bar{\text{I}}$; dann folgt die Behauptung mit 2.9.

Es sei noch bemerkt, daß alle drei Voraussetzungen in 2.10.4 bzw. 2.10.5 wirklich notwendig sind. Statt 5b) hätten wir wie in 2.9.4b) auch äquivalent fordern können, falls M endlich ist: Jedes Ideal $J \supset \operatorname{Ann} M$ hat eine Nullstelle in G.

Folgender Satz gibt darüber Auskunft, wann ein Modul mit I schon III erfüllt. Dabei heiße eine Steinsche Menge $G \subset Z$ im nichtarchimedischen kompakt, wenn G ein Umgebungssystem von speziellen affinoiden Bereichen besitzt (diese Definition ist auch für viele andere Sätze, die sich auf kompakte Steinsche Mengen beziehen, sinnvoll):

Satz 2.11. *Sei $G \subset Z$ eine Steinsche Menge und $R_Z(G)$ noethersch. Dann sind für jeden Ring $R \subset R(G)$ und jeden R-Modul M die Eigenschaften* I, II, III, $\bar{\mathrm{I}}$, $\bar{\mathrm{II}}$ *und* $\overline{\mathrm{III}}$ *äquivalent. Ferner ist G dann kompakt.*

Beweis. Wir zeigen zunächst die zweite Behauptung. Wir führen den Beweis nur im nichtarchimedischen durch (im archimedischen ist er noch einfacher): Da G Steinsch ist, gibt es eine Umgebungsbasis von Steinschen Räumen G_i, $G \subset G_i \subset Z$. Jeder Steinsche Raum G_i ist aufsteigende Vereinigung von speziellen affinoiden Bereichen G_i^j. Ist nun bei festem i für jedes j der Punkt $z_j \in G_i^j - G_i^{j-1}$ gewählt, so können wir wegen der Gültigkeit von Theorem A im nichtarchimedischen (s. [12]) zu fester Zahl k Funktionen f_v^k finden, die alle in z_j für $j > k$ eine Nullstelle haben und in $\{z_1, \ldots, z_k\}$ keine gemeinsame Nullstelle haben. Wäre $z_j \in G$ für alle j, so wäre demnach bei $J_k := \{f \in R(G) | f(z_j) = 0$ für $j > k\}$ die Idealkette $J_1 \subset J_2 \subset \cdots$ eine echt aufsteigende Kette. Da $R(G)$ noethersch vorausgesetzt war, kann für große j also nicht mehr $(G_i^j - G_i^{j-1}) \cap G \neq \emptyset$ sein und es ist somit schon $G \subset G_i^{j_0}$. Da in jeder Umgebung U von G ein Steinscher Raum G_i mit $G \subset G_i \subset U$ existiert, gibt es somit auch zu jedem U einen speziellen affinoiden Bereich $G_i^{j_0}$ mit $G \subset G_i^{j_0} \subset U$; also ist G kompakt.

Aus G kompakt folgt nun, daß $N := \Gamma(G, \mathscr{M})$ für jede in einer Umgebung von G kohärente Garbe $\mathscr{M}$ ein endlicher $R(G)$-Modul ist (Im nichtarchimedischen kommt N von einer Garbe auf einem speziellen affinoiden Bereich her; wegen der Eigenschaft I muß dann auch N endlicher $R(G)$-Modul sein.) Nach Definition gibt es ferner zum gegebenem R-Modul M eine kohärente Garbe $\mathscr{M}$ mit $M \subset N := \Gamma(G, \mathscr{M})$.

Ist M' ein endlicher R-Untermodul von M, so ist erst recht $M' \subset N$ und nach 2.10 erfüllt $M' R_Z(G)$ die Bedingung (III) als $R_Z(G)$-Untermodul von N: Denn N ist noetherscher R_Z(G)-Modul und erfüllt I als $R_Z(G)$-Modul. Es gibt also $z_1, \ldots, z_n \in G$, so daß

$$\{f \in N | f \in M' \mathcal{O}_{z_i}\} = \{f \in N | f \in M' \mathcal{O}_z \ \forall z \in G\}$$

ist. Daraus folgt sofort, daß

$$\{f \in M | f \in M' \mathcal{O}_{z_i}\} = \{f \in M | f \in M' \mathcal{O}_z \ \forall z \in G\}$$

erfüllt ist, und wegen der Eigenschaft I für M ist dies gleich M'. Damit erfüllt M' die Eigenschaft (III); somit gilt auch für M III.

Wird nur R als noethersch vorausgesetzt, nicht aber $R_Z(G)$, so ist 2.11 falsch. Sei z.B. $G \subset \mathbb{R}^2$ kompakt so gewählt, daß es $z_i \in G$, $i = 1, 2, 3, \ldots$ gibt, so daß die $J_n = \{f \in R(G) | f(z_i) = 0$ für $i \geqq n\}$ eine echt aufsteigende Kette bilden (s. Beispiel in [5]). Setze $M := \bigcup J_n / J_0$. Dann erfüllt M die Bedingung I über dem Polynomring: Denn ist $\bar{f} \in (\bar{f}_1, \ldots, \bar{f}_n) \mathcal{O}_z \ \forall z \in G$, so verschwindet f auf all den z_i, auf denen auch die f_i verschwinden; es gibt dann Polynome P_i, so daß für $i \leqq N_0$ stets $(f-(f_1 P_1 + \cdots + f_n P_n))(z_i)=0$ ist; für $i \geqq N_0$ ist bei großem N_0 sowieso immer $f_j(z_i)$ und $f(z_i)=0$. M ist aber nicht noethersch, somit erfüllt M nicht III.

Ist jedoch M endlich erzeugt und R noethersch, so sind trivialerweise die Eigenschaften I, II, III, $\bar{\mathrm{I}}$, $\overline{\mathrm{II}}$ und $\overline{\mathrm{III}}$ für M äquivalent.

Speziell folgt aus 2.11:

Ist G semianalytisch kompakt, $\mathcal{M}$ eine kohärente Garbe in einer Umgebung von G, $M \subset \Gamma(G, \mathcal{M})$, so ist $MR(G)$ ein $R(G)$-Modul mit III.

Wir wollen nun die analytischen Moduln mit II charakterisieren. Wir brauchen dazu folgenden Hilfssatz 2.12.1 (s. [14], 3.21).

Hilfssatz 2.12.1. *Sei $\mathcal{M}$ eine kohärente Garbe über $G \subset Z$. Zu jedem $z_0 \in \mathring{G}$ und zu jeder Teilmenge $N \subset M := \Gamma(G, \mathcal{M})$ gibt es eine Umgebung $U(z_0)$, so daß gilt:*

$$\{f \in M \mid f \in N\,\mathcal{O}_{z_0}\} = \{f \in M \mid f \in N\,\mathcal{O}_z \ \forall z \in U(z_0)\}.$$

Hilfssatz 2.12.2. *Sei $\mathcal{M}$ eine kohärente Garbe über eine in Z offene Umgebung $U \supset G$.*

a) *Ist $U_1 \supset U_2 \supset \cdots$ eine Umgebungsbasis von $G \subset U$, so hat jede Menge $M \subset \Gamma(G, \mathcal{M})$ eine Darstellung,*

$$M = \bigcup_1^\infty M_i \quad \text{mit} \quad M_1 \subset M_2 \subset M_3 \subset \cdots, \ M_i \subset \Gamma(U_i, \mathcal{M}).$$

b) *Es gibt zu jedem M_i und zu jeder Untermenge $N \subset M$ eine* diskrete *Menge $\{z_j^i\} \subset G$, so daß*

$$\{f \in M_i \mid f \in N\,\mathcal{O}_{z_j^i}\} = \{f \in M_i \mid f \in N\,\mathcal{O}_z \ \forall z \in G\}$$

ist.

Beweis. a) Es ist $\Gamma(G, \mathcal{M}) = \bigcup_1^\infty \Gamma(U_i, \mathcal{M})$. Dann definieren wir

$$N_i := M \cap \Gamma(U_i, \mathcal{M}).$$

b) Sei für irgendein $z_j \in G$ und für ein $f \in M_i$ die Relation $f \in N\,\mathcal{O}_{z_j}$ erfüllt. Es ist $N\,\mathcal{O}_{z_j} = (f_1^j, \ldots, f_n^j)\,\mathcal{O}_{z_j}$ mit $f_k^j \in N$; und wegen $M = \bigcup M_i$ sind alle diese f_k^j in einem M_m enthalten. Somit ist $\{f, f_k^j\} \subset M_{m+i}$ und $f \in \{f_k^j\}\,\mathcal{O}_{z_j}$. Nach 2.12.1 folgt dann, wenn wir $G^* \equiv U_{m+i}$ und $N^* \equiv \{f_k^j\}$ setzen, daß es eine *von f unabhängige* Umgebung $U(z_j) \subset U_{m+i}$ gibt, so daß $f \in \{f_k^j\}\,\mathcal{O}_z$ für alle $z \in U(z_j)$ ist.

Überdecken wir nun G mit einer lokalendlichen Überdeckung von solchen $U(z_j)$, so folgt aus $f \in M_i$, $f \in N\,\mathcal{O}_{z_j}\ \forall j$, daß $f \in (f_1^j, \ldots f_n^j)\,\mathcal{O}_{z_j}$ ist und damit nach dem gerade Gesagten auch $f \in (f_1^j, \ldots, f_n^j)\,\mathcal{O}_z$ und somit $f \in N\,\mathcal{O}_z\ \forall z \in U(z_j)$ ist. Es gibt also eine diskrete Menge z_j, so daß $\{f \in M_i \mid f \in N\,\mathcal{O}_{z_j}\} = \{f \in M_i \mid f \in N\,\mathcal{O}_z\ \forall z \in G\}$ ist.

Damit folgt dann leicht:

Hilfssatz 2.13. *Sei M ein analytischer Modul. Erfüllt $N \subset M$ die Bedingung* (I), *so erfüllt N auch die Bedingung* (II) *und umgekehrt. Es gibt dann sogar eine abzählbare und in $\mathring{G}$ diskrete Menge $\{z_i\}$, so daß $\{f \in M \mid f \in N\,\mathcal{O}_{z_i}\} = N$ ist.*

Beweis. Definiere nach 2.12.2 die Punkte z_j^i. Ist dann $f \in N\mathcal{O}_{z_j^i}\ \forall i\ \forall j$, so ist, da f in einem M_i liegt, schon $f \in N\mathcal{O}_z$ für alle $z \in G$; somit ist also $f \in N$. Da $\{z_j^i\}_j^i$ abzählbar ist, erfüllt N also (II).

Wir können die abzählbare Menge $\{z_i\}$ nun so wählen, daß sie in $\mathring{G}$ diskret ist. Dies ergibt sich so: Wir setzen in 2.12.2 statt G die offene Hülle $\mathring{G}$ ein und definieren $U_i \equiv \mathring{G}$. Es gibt dann nach 2.12.2b) eine diskrete Menge von Punkten $z_i' \in G$ mit $\{f \in M \mid f \in N\mathcal{O}_{z_i'}\} = \{f \in M \mid f \in N\mathcal{O}_z\ \forall z \in \mathring{G}\}$. Wir nehmen dann von unserer abzählbaren Menge $\{z_i\}$ alle Punkte $z_i \in G - \mathring{G}$ und vereinigen diese mit der Menge $\{z_i'\}$ zu der Menge $\{z_i''\}$. Dann ist $\{z_i''\}$ eine gewünschte Menge.

Wenn M (I) erfüllt, gibt es also immer zu jedem endlichen Untermodul N eine abzählbare und in $\mathring{G}$ diskrete Menge z_i, so daß $\{f \in M \mid f \in N\mathcal{O}_{z_i}\} = N$ ist. Es gibt jedoch nicht immer eine solche diskrete Menge (s. [13]).

Satz 2.13.1. *Sei M ein analytischer Modul.*

a) *M erfüllt* I *genau dann, wenn M* II *erfüllt.*

b) *M erfüllt* $\bar{\text{I}}$ *genau dann, wenn M* $\overline{\text{II}}$ *erfüllt.*

Satz 2.13.2. *Sei G offen, M ein analytischer R-Modul. M erfüllt genau dann* I, *wenn es zu jedem endlich erzeugten Untermodul $N \subset M$ eine diskrete Menge $\{z_i\} \subset G$ gibt mit*

$$\{f \in M \mid f \in N\mathcal{O}_{z_i}\} = N.$$

Wir können nun auch für kompakte G Moduln mit I analog Satz 2.10 algebraisch charakterisieren.

Satz 2.14. *Sei G kompakt. Dann ist für einen analytischen R-Modul M äquivalent*

1) *M erfüllt* I.
2) *M erfüllt* $\bar{\text{I}}$.
3) *M erfüllt* II.
4) *M erfüllt* $\overline{\text{II}}$.
5) a) *M ist lokal noethersch.*
   b) *Für jedes $N \subsetneqq M$ ist* $\operatorname{Var} N \neq \emptyset$.
   c) *Die Untermoduln der Form $(z-z_0)^k M$ erfüllen* (I).

Beweis. 1)→2) M erfülle I. Sei $N \subset M$ ein beliebiger Untermodul und sei $f \in N\mathcal{O}_z\ \forall z \in G$. Aus $f \in N\mathcal{O}_{z_i}$ folgt sodann, daß es einen endlichen Untermodul $N_i \subset N$ gibt mit $f \in N_i\mathcal{O}_{z_i}$. Dann ist trivialerweise für eine gewisse Umgebung $U(z_i)$ auch $f \in N_i\mathcal{O}_z\ \forall z \in U(z_i)$. Da G kompakt ist, überdecken endlich viele $U(z_i)$ schon G; also $\bigcup\limits_1^m U(z_i) = G$. Dann sei $N' := \sum\limits_1^m N_i$. Es folgt, daß $f \in N'\mathcal{O}_z\ \forall z \in G$ ist. Da N' endlich ist, folgt $f \in N'$; somit erfüllt M $\bar{\text{I}}$.

4)→5) und 5)→2) s. 2.9.

Aus 2.14 folgt speziell, daß für kompakte Steinsche G stets $R(G)$ lokal noethersch ist. Der Beweis 2.14 kann auch auf den nichtarchimedischen Fall verallgemeinert werden, wobei jetzt G nur im Sinne von 2.11 kompakt ist:

Denn es gibt zu z_i eine Zariski-offene Umgebung $U(z_i)$ mit $f \in N_i \mathcal{O}_z \ \forall z \in U(z_i)$. Da das Maximalidealspektrum von $R(G)$ den Punkten aus G entspricht, ist G kompakt bezüglich Zariski-offener Überdeckungen.

Es sei noch bemerkt, daß für nichtkompakte G Satz 2.14 falsch ist: Die wichtigsten Moduln mit I erfüllen nie $\bar{\mathrm{I}}$ (s. [13]); hier ist nämlich nicht 2.14.5 b) erfüllt.

Wir müssen also nun untersuchen, wann jeder Untermodul eine nichtleere Varietät hat. Wir brauchen dazu einige Hilfssätze:

Hilfssatz 2.15. *Sei M ein endlicher R-Modul mit* I *bezüglich G. Dann ist für $N \subset M$ stets* $\operatorname{Var} N = \operatorname{Var}(\operatorname{Ann} M/N)$. *Insbesondere ist* $\operatorname{Tr} M := \{z \in G \mid M\mathcal{O}_z \neq 0\} = \operatorname{Var}(\operatorname{Ann} M)$.

Beweis. Sei $M = (m_1, \ldots m_q) R$. Ist $z \notin \operatorname{Var}(\operatorname{Ann} M/N)$, so gibt es $r \in R$, $r(z) \neq 0$ mit $rM \subset N$. Dann muß $M\mathcal{O}_z = N\mathcal{O}_z$ sein; somit ist $z \notin \operatorname{Var} N$. Ist $z \notin \operatorname{Var} N$, so ist $M\mathcal{O}_z = N\mathcal{O}_z$ und damit nach 2.7.1 auch $MR_z = NR_z$, d.h. zu jedem m_i gibt es ein s_i mit $s_i(z) \neq 0$ und $s_i m_i \in N$. Dann gilt für $s := \prod s_i$, daß $s(z) \neq 0$ und $sM \subset N$ ist; somit ist $z \notin \operatorname{Var}(\operatorname{Ann} M/N)$.

Hilfssatz 2.16. *Sei M ein analytischer Modul, so daß M_{z_0} noetherscher R_{z_0}-Modul für alle $z_0 \in G$ ist und so daß die Untermoduln $(z - z_0)^k M$ die Bedingung* (I) *erfüllen.*

Ist p ein primärer Untermodul mit dem assoziierten Primideal q, so ist $\operatorname{Var} p = \operatorname{Var} q$.

Beweis. Nach 2.2 ist $\operatorname{Var} q \supset \operatorname{Var} p$. Sei $z_0 \notin \operatorname{Var} p$. Dann ist für jedes $f \in M - p$ auch $f \in p\mathcal{O}_{z_0}$, also nach 2.7.1 schon $f \in pR_{z_0}$ und damit $fs \in p$ für ein $s \in R$ mit $s(z_0) \neq 0$. Es muß $s \in q$ sein und es ist dann $z_0 \notin \operatorname{Var} q$.

Folgerung 2.16.1. *Sei M ein Modul mit* I. *Dann gilt für jeden primären Untermodul p mit dem assoziierten Primideal q:*

$$\operatorname{Var} p = \operatorname{Var} q.$$

Es sei noch bemerkt, daß die Voraussetzungen in 2.16 nötig sind. Denn daß ein Modul M I erfüllt, bedeutet: Ist $N' \subset M$ ein fester endlicher Modul, so ist für jedes $N \supsetneq N'$ immer $\operatorname{Var}_N N' \neq \emptyset$. Mit dem vorigen folgt dann, daß die Voraussetzungen notwendig sind.

Folgerung 2.16.2. *Sei M ein analytischer Modul. Dann ist äquivalent*

1) *M erfüllt* I *(bzw.* II *bzw.* III*)*

2) a) *Jeder endlicher Untermodul $N \subset M$ hat eine beliebige (bzw. abzählbare bzw. endliche) Primärzerlegung $N = \bigcap p_i$ mit $\operatorname{Var} q_i \neq \emptyset$, wobei q_i die zu p_i assoziierten Primideale sind.*

b) *Die Untermoduln $(z - z_0)^k M$ erfüllen* (I).

Beweis. Nach Bemerkung 1.7.2 ist M_{z_0} noethersch für alle $z_0 \in G$: Denn aus der Eigenschaft (I) für $(z - z_0)M$ folgt die Beziehung $(z - z_0)M = m(z_0)M$ und jeder endliche Untermodul hat eine Primärzerlegung.

Nach Hilfssatz 2.16 folgt sodann, daß $\operatorname{Var} p_i \neq \emptyset$ ist. Die Behauptung folgt dann aus 2.7.

Insbesondere folgt noch:

Folgerung 2.16.3. *Sei M ein analytischer Modul über einem Ring $R \subset R_Z(G)$, in dem jedes Ideal $\supset \operatorname{Ann} M$ eine Nullstelle in G hat. Dann ist äquivalent:*

1) *M erfüllt* I.
2) *M erfüllt* $\bar{\mathrm{I}}$.
3) a) *M lokal noethersch.*
   b) $(z-z_0)^k M$ *erfüllt* (I).

Beweis folgt aus 2.8, 2.9 und 2.16.

Kriterien für Ringe, in denen jedes Ideal eine nichtleere Varietät hat, werden in [13] gegeben. Hier sollen mit ähnlichen Methoden etwas schwieriger zu formulierende Kriterien für Moduln gegeben werden.

Dazu betten wir wie zu Anfang den Raum Z in einen k^m ein und definieren dann für eine beliebige kompakte Teilmenge $G \subset Z$ (im Fall k algebraisch abgeschlossen nichtarchimedisch sei G im Sinne von 2.11 kompakt):
$\tilde{G} := \{z \in k^m |$. Jedes Polynom ohne Nullstelle in G hat auch in z keine Nullstelle$\}$.

Die Bedingung $G = \tilde{G}$ ist äquivalent damit, daß im Polynomring $R = k[Z_1, \ldots, Z_m]$ jedes maximale Ideal $(z - z^0)R$ für $z^0 \notin G$ ein Polynom P enthält mit $P(z) \neq 0$ für alle $z \in G$. Es folgt somit nach [13] mit der dort gebrauchten Terminologie: Erfüllt R_G I bezüglich G, so ist $G = \tilde{G}$.

Für $k = \mathbb{R}$ ist immer $G = \tilde{G}$: Denn $\sum (z_i - z_i^0)^2$ hat bei $z^0 \notin G$ in G keine Nullstelle.

Für $k = \mathbb{C}$ ist z.B. für jedes G mit $G = \hat{G}_{\mathbb{C}^m}$ (s. [11]) schon $G = \tilde{G}$: Denn zu $z^0 \in G$ gibt es ein $f \in R(\mathbb{C}^m)$ mit $|f(z^0)| > |f(z)|$ $\forall z \in G$. Wenn wir f genügend durch ein Polynom P approximieren, hat $P - P(z^0)$ keine Nullstelle in G. Weiter ist für jedes Polygebiet $G = \prod G_i$ mit $G_i \subset \mathbb{C}$ stets $G = \tilde{G}$.

Ist k nichtarchimedisch nicht algebraisch abgeschlossen, so ist für kompakte G stets $G = \tilde{G}$: Denn bei $z^0 \notin G$ hat das Ideal $(z - z^0) R_G$ im lokalisierten Polynomring $R_G = (k[z_1, \ldots, z_m])_G$ keine Nullstelle in G; nach [13], Hilfssatz 13 gibt es somit ein Polynom ohne Nullstelle in G, das eine Nullstelle in z^0 hat.

Bei nichtarchimedisch algebraisch abgeschlossenem k ist immer $G = \tilde{G}$, wenn G im Sinne von 2.11 kompakt ist: Denn zu $z^0 \notin G$ gibt es zunächst eine holomorphe Funktion f mit $|f(z)| \leqq 1$ $\forall z \in G$ und $|f(z^0)| > 1$ (oder umgekehrt). Wir approximieren f genügend durch ein Polynom P, so daß $P - P(z^0)$ keine Nullstelle in G hat.

Hilfssatz 2.17. *Sei M ein lokalendlicher Modul und sei $N \subsetneqq M$. Dann gibt es ein maximales Ideal $\mathfrak{m}$ und einen zu $\mathfrak{m}$ assoziierten primären Untermodul $p \subsetneqq M$ mit $p \supset N$ und $p \supset \mathfrak{m} M$.*

Beweis. Es gibt ein maximales Ideal $\mathfrak{m}$ mit $N_{\mathfrak{m}} \subsetneqq M_{\mathfrak{m}}$. Nach Zorn gibt es dann einen maximalen Untermodul $p_{\mathfrak{m}}$ in der Menge $\{Q_{\mathfrak{m}} | N_{\mathfrak{m}} \subset Q_{\mathfrak{m}} \subsetneqq M_{\mathfrak{m}}\}$. Nach Nakayama ist $\mathfrak{m} M_{\mathfrak{m}} \subset p_{\mathfrak{m}}$. Somit ist $p_{\mathfrak{m}}$ primär zu $\mathfrak{m} R_{\mathfrak{m}}$. Dann ist aber auch $p := p_{\mathfrak{m}} \cap M$ primär zu $\mathfrak{m}$ und $\mathfrak{m} M \subset p \subsetneqq M$.

Hilfssatz 2.18. *Sei M ein lokal endlicher analytischer Modul zu $G=\tilde{G}$. Die Untermoduln der Form $(z-z^0)^k M$ mögen* (I) *erfüllen, und es sei M_{z^0} noethersch für alle $z^0 \in G$. Für jedes $f \in R$ mit $f(z) \neq 0$ $\forall z \in G$ sei $fM=M$. Dann ist für jeden Untermodul $N \subsetneqq M$ schon* $\operatorname{Var} N \neq \emptyset$.

Beweis. Nach 2.17 genügt es, die Behauptung für alle $p \subsetneqq M$ mit $\mathfrak{m} M \subset p$ zu zeigen. $\mathfrak{m}$ kann keine Funktion f ohne Nullstelle in G enthalten: Denn sonst wäre nach Voraussetzung $fM=M$ und somit $M \subset fM \subset \mathfrak{m} M \subset p$.

1. $k=\mathbb{C}$. Wie in [13], Hilfssatz 12, folgt, da jedes Element aus $\mathfrak{m}$ eine Nullstelle in G hat: Es gibt $z^0 \in k^m$, so daß $(z-z^0) \in \mathfrak{m}$ ist. Für $z^0 \in k^m$, aber $z^0 \notin G$ folgte unter Benutzung von $G=\tilde{G}$, daß es ein Polynom $f \in \mathfrak{m}$ gibt, das keine Nullstelle in G hat. Dies ist ein Widerspruch zum obigen. Somit muß $z^0 \in G$ sein. Wenn $\mathfrak{m} \subset \mathfrak{m}(z^0)$ ist, sind wir mit Hilfssatz 2.16 fertig. Sonst gibt es $f \in \mathfrak{m}$ mit $f(z^0) \equiv b \neq 0$. Es ist dann, da $(z-z^0)M$ (I) erfüllt, $(f-b)M \subset (z-z^0)M$; somit folgte $bM \subset \mathfrak{m} M \subset p$, was ein Widerspruch ist.

2. k nicht algebraisch abgeschlossen. In [13] wurde gezeigt, daß jedes Ideal ohne Nullstelle in G schon ein f enthält mit $f(z) \neq 0$ für alle $z \in G$. Wie unter 1 folgt die Behauptung.

3. k algebraisch abgeschlossen nichtarchimedisch. Ist $J \subset R$ ein Ideal ohne Nullstelle in G, so gibt es $r_i \in R(G)$ und $j_i \in J$ mit $\sum_1^m j_i r_i = 1$. Wenn jedes r_i durch ein Polynom P_i genügend approximiert wird, hat $\sum_1^m j_i P_i$ keine Nullstelle in G. Wie oben folgt dann die Behauptung.

Es folgt nun:

Satz 2.19. *Sei $G=\tilde{G}$, M ein analytischer Modul zu G. Dann ist äquivalent:*

A. 1) *M erfüllt* I.

2) *M erfüllt* $\bar{\text{I}}$.

3) a) *M ist lokal noethersch.*

b) *Ist $f(z) \neq 0$ für alle $z \in G$, so ist $fM=M$.*

c) *Der Untermodul $(z-z^0)^k M$ erfüllt* (I).

B. 1) *M erfüllt* III.

2) *M erfüllt* $\overline{\text{III}}$.

3) a) *M ist noethersch.*

b) *Ist $f(z) \neq 0$ für alle $z \in G$, so ist $fM=M$.*

c) *Der Untermodul $(z-z^0)^k M$ erfüllt* (I).

Beweis. 1) $\to$ 2) folgt aus 2.14.

2) $\to$ 3) a) und c) folgen aus 2.14; b) folgt aus 2.14.5 b): Denn wenn $f(z) \neq 0$ für alle $z \in G$ ist, ist $\operatorname{Var} fM = \emptyset$.

3) $\to$ 1) Mittels 2.18 folgt, daß 2.14.5 b) erfüllt ist. Dann folgt aus 2.14 die Behauptung.

B folgt wie A unter Benutzung von 2.10.

Es sei noch bemerkt, daß die Voraussetzung $G=\tilde{G}$ wirklich nötig ist: Denn der lokalisierte Polynomring $(k[X_1, \ldots, X_n])_G$ (s. [13]) erfüllt alle Eigenschaften 3a)→3c), nach obigem erfüllt er jedoch nur für $G=\tilde{G}$ die Bedingung I.

Eng zusammen hängt damit folgender Satz:

2.20. *Sei R ein Ring mit* I *bezüglich G. Dann gibt es zu jedem $z^0 \notin G$ ein $f \in R$, das sich nicht dort holomorph fortsetzen läßt.*

Beweis. Es ist $(z-z^0)\,\mathcal{O}_z = \mathcal{O}_z\ \forall z \in G$, also $1 \in (z-z^0)\,R$.

Wir wollen nun speziell Kriterien dafür angeben, wenn ein *Ring* I erfüllt. Dazu benutzen wir:

Hilfssatz 2.21. *Sei $G \subset k^m$ beliebig, $R \subset R_m(G)$. Folgt aus $z^0 \in G$ und $f(z^0)=0$ schon $f \in (z-z^0)\,R$, so erfüllen die Untermoduln $(z-z^0)^k$ die Bedingung* I.

Beweis. Siehe Beweis von Satz 6 in [13].

Folgerung 2.21. *Sei $G \subset k^m$ beliebig, $R \subset R_m(G)$.*

A. *Dann ist äquivalent:*

1) *R erfüllt* $\bar{\mathrm{I}}$.

2) *R erfüllt* I, *jedes Ideal hat eine Nullstelle.*

3) a) *R ist lokal noethersch.*

b) *Jedes Ideal hat eine Nullstelle.*

c) *Aus $f \in R$, $f(z^0)=0$ folgt $f \in (z-z^0)\,R$.*

B. *Sei $G=\tilde{G}$. Dann ist äquivalent:*

1) *R erfüllt* I.

2) *R erfüllt* $\bar{\mathrm{I}}$.

3) a) *R ist lokal noethersch.*

b) *Aus $f \in R$, $f(z) \neq 0$ für alle $z \in G$ folgt f Einheit.*

c) *Aus $f \in R$, $f(z^0)=0$ folgt $f \in (z-z^0)\,R$.*

C. *Sei $G=\tilde{G}$. Dann ist äquivalent:*

1) *R erfüllt* III.

2) a) *R noethersch.*

b) *Aus $f \in R$, $f(z) \neq 0$ für alle $z \in G$ folgt f Einheit.*

c) *Aus $f \in R$, $f(z^0)=0$ folgt $f \in (z-z^0)\,R$.*

Beispiel 2.22. Aus C folgt, daß für algebraisch abgeschlossene k der Tatesche Ring T_n die Eigenschaft III bezüglich E_n: $\{z \in k^n \,|\, |z| \leqq 1\}$ erfüllt.

Aus A folgt, daß für algebraisch abgeschlossene k der Polynomring I und damit auch III bezüglich k^n erfüllt.

Die Konstruktion weiterer Ringe mit I folgt in [14] (andere Beispiele s. in [13]).

Betrachtet man anstatt Unterringe von $R_m(G)$ auch Unterringe von $R_Z(G)$, wobei Z Singularitäten hat, so ist dieser Satz falsch; ein entsprechender Satz ist *bei Moduln M* auch für $G \subset k^m$ falsch: Hier kann nicht von der Eigenschaft (I) für $(z-z^0)\,M$ auf die Eigenschaft (I) für $(z-z^0)^k M$ geschlossen werden.

Wir wollen nun noch kurz den Zusammenhang der Beziehungen I, II, III, $\bar{\mathrm{I}}$, $\overline{\mathrm{II}}$ und $\overline{\mathrm{III}}$ für einen R-Modul M tabellarisch festhalten:

2.23. A. G beliebig $\mathrm{III} \leftrightarrow \overline{\mathrm{III}} \rightarrow \overline{\mathrm{II}} \leftrightarrow \bar{\mathrm{I}} \rightarrow \mathrm{II} \leftrightarrow \mathrm{I}$
$\mathrm{III} \leftrightarrow \overline{\mathrm{III}} \leftrightarrow \mathrm{II} \wedge M$ noethersch $\leftrightarrow \mathrm{I} \wedge M$ noethersch.

2.23. B. G kompakt $\mathrm{III} \leftrightarrow \overline{\mathrm{III}} \rightarrow \overline{\mathrm{II}} \leftrightarrow \bar{\mathrm{I}} \leftrightarrow \mathrm{II} \leftrightarrow \mathrm{I}.^{*}$

2.23. C. G semianalytisch kompakt
$\mathrm{III} \leftrightarrow \overline{\mathrm{III}} \leftrightarrow \overline{\mathrm{II}} \leftrightarrow \bar{\mathrm{I}} \leftrightarrow \mathrm{II} \leftrightarrow \mathrm{I}.$

Literatur

1. Cartan, H.: Seminaire E.N.S. 1953/54 (Seminarausarbeitung).
2. — Ideaux et Moduls de fonctions analytiques de variables complexes. Bull. Soc. Math. France **78**, 28—64 (1950).
3. Forster, O.: Primärzerlegungen in Steinschen Algebren. Math. Ann. **154**, 307—329 (1964).
4. — Zur Theorie der Steinschen Algebren und Moduln. Math. Z. **97**, 376—405 (1967).
5. Frisch, J.: Points des platitude d'un morphisme d'espaces analytiques complexes. Inventiones math. **4**, 118—138 (1967).
6. Grauert, H.: Analytische Faserung über holomorph vollständigen Räumen. Math. Ann. **133**, 139—159 (1957).
7. Remmert, R.: Nichtarchimedische Funktionentheorie. Weierstraß-Festband. Opladen; Westdeutscher Verlag 1966.
8. — — Analytische Stellenalgebren. Grundlehren der Mathematik, Bd. 176. Berlin-Heidelberg-New York: Springer 1971.
9. Gunning, R.C., Rossi, H.: Analytic functions of several complex variables. Englewood Cliffs, N.J.: Prentice Hall 1965.
10. Hervé, M.: Several Complex Variables, Local Theory. Oxford University Press 1963.
11. Hörmander, L.: An Introduction to complex analysis in several variables. Princeton: Van Nostrand 1966.
12. Kiehl, R.: Theorem A und Theorem B in der nichtarchimedischen Funktionentheorie. Inventiones math. **2**, 256—273 (1967).
13. Langmann, K.: Ringe holomorpher Funktionen und endliche Idealverteilungen. Schriftenreihe des Math. Inst. der Univ. Münster, Serie 2, Heft 3 (1971).
14. — Konstruktionen globaler Moduln und Anwendungen. (Manuskript.)
15. Malgrange, B.: Ideals of Differentiable Functions. Oxford University Press, 1966.
16. Nagata, M.: Local Rings. New York: Interscience 1962.
17. Narasimhan, R.: Introduction to the Theory of Analytic Spaces. Lecture Notes **25**, Berlin-Heidelberg-New York: Springer 1966.
18. Nastold, H.-J.: Neuere Methoden in der lokalen Algebra. Manuskript, Münster, 1966.
19. — Nonarchimedian Function Theory. Acta des Coloquio Int. Sobre Geometria Algebraica, Madrid, 1965.
20. Remmert, R.: Algebraische Aspekte in der nichtarchimedischen Analysis. Proceedings of a Conference on local fields. Berlin-Heidelberg-New York: Springer 1967.
21. — Stein, K.: Über die wesentlichen Singularitäten analytischer Mengen. Math. Ann. **126**, 263—306 (1953).
22. Schilling, O.F.G.: Ideal theory on open Riemann surfaces. Bull. Amer. Math. Soc. **52**, 945—963 (1946).
23. Serre, J.P.: Algèbre locale, Multiplicités. Lecture Notes **11**, Berlin-Heidelberg-New York: Springer 1965.

Dr. Klaus Langmann
Mathematisches Institut der Universität
D-4400 Münster, Roxeler Straße 64
Deutschland

(Eingegangen am 14. Juni 1971)

Math. Z. 124, 169–185 (1972)

A Characterization of the Finite Groups $PSL(n, q)$

KOK-WEE PHAN*

The object of this paper is to present a characterization of the simple group $L_n(q)\,(=PSL(n,q))$ where $n\geqq 5$ and q is odd, in terms of the centralizer of an element of order 2. The cases $n\leqq 4$ and q is odd have already been treated ([4], $n=2$; [1, 2, 3], $n=3$; [10], $n=4$). Hence these works together with our result in this paper give characterizations of all $L_n(q)$ when q is odd. Similar characterizations exist for all $L_n(q)$ with even q ([8, 9], $n=4, 5$; $q=2$, [14], $n=4$, $q=2$; [12], all other cases).

To state our Main Theorem, let $SL(n, q)$ denote the special linear group of an n-dimensional vector space V over the finite field F_q of order q where q is odd. Let t' be the image in $L_n(q)$ of an involution in $SL(n, q)$ whose fixed elements in V form a subspace of dimension $n-2$. Denote by $H(n, q)$ the centralizer of t' in $L_n(q)$. The following result is proved.

Main Theorem. *Let G be a finite group with an involution t whose centralizer $C(t)$ in G is isomorphic to $H(n,q)$ where $n\geqq 5$ and q is odd. Then $G=O(G)\,C(t)$ or G is isomorphic to one of the following groups.*

(i) $PGL(3,q)\times SL(n-2,q)$;

(ii) $PSL(3,q)\times SL(n-2,q)$;

(iii) $SL(3,q)\times SL(n-2,q)$;

(iv) $PSL(3,q)\times SL(n-2,q)\times Z_3$;

(v) *a non-central extension of* $PSL(3,q)\times SL(n-2,q)$ *by* Z_3;

(vi) *an extension of the central product of* $SL(3,q)$ *and* $SL(n-2,q)$ *by* Z_3;

(vii) $M_{11}\times SL(n-2,3)$;

(viii) $SL(3,q)\wr Z_2$

(ix) *an extension of the central product of two* $SL(3,q)$ *by a group of order* 6;

(x) $L_n(q)$.

As usual $O(G)$ denotes the largest normal subgroup of odd order of G; Z_i denotes the cyclic group of order i and M_{11} is the Mathieu group of order 7920.

Exactly which of the possible structures of G occurs depends on the fusion pattern of involutions in $C(t)$. This in turn is determined by q and the greatest common divisors $d=(n,q-1)$ and $e=(n-2,q-1)$. There are four types of fusion of involutions in $C(t)$ and they are given in Theorems A, B, C and D of § 3, 4 and 5. Thus our Main Theorem follows from the fact that the group G fulfills the hypothesis of one of these theorems.

* This work is partially supported by the National Science Foundation Grant GP11342.

The major portion of the proof is concerned with the analysis of the possible fusion of involutions in $C(t)$. Fortunately it can often be accomplished by rather elementary methods. However in some cases the analysis is quite complex. Once the fusion pattern of involutions is determined, this leads rapidly to a construction of a subgroup G_0 of G with known structure using the theorems of Brauer [1, 2, 3] in the nonsimple cases and our result [11] in the simple case. Then it remains to show that G_0 is in fact G. This is done by proving in essense that G_0 is a strongly embedded subgroup.

It is probable that other involutions in $L_n(q)$ which are images of involutions in $SL(n, q)$ may be used in place of t' for this characterization. Our choice of t' has the virtue, among other things, that proofs of most lemmas are almost identical for both odd and even n, thus considerably shorten our paper. It is of interest to remark that some part of this work is similar to Wong's characterization of $PS_p(2n, q)$ [15] and it is also the case here that the involution t' needs not lie in the center of a Sylow 2-subgroup of $L_n(q)$.

The notation used is standard. Thus if x, y are elements of a group, x^y, $[x, y]$ denote $y^{-1}xy$ and $x^{-1}y^{-1}xy$ respectively and X^y denotes $y^{-1}Xy$ if X is a subset. If U is a vector space of dimension m over F_q where q is odd, an involution z of $SL(m, q)$ will be called an extremal involution if the fixed elements of z in U form a subspace of dimension $m-2$.

§ 1. The Structure of the Centralizer

We shall identify $C(t)$ with $H(n, q)$ (hence $t=t'$ since $Z(H(n, q))$ has the unique involution t') and introduce notation for the elements of $C(t)$. The group $C(t)$ will also be denoted by the letter H.

First we have a decomposition of the vector space V

$$V=V^- \oplus V^+$$

where V^- is the subspace of elements of V mapped by t into its negative and V^+ is the subspace of fixed elements of t in V. Choose bases $\{v_1, v_2\}$ and $\{v_3, v_4, \ldots, v_n\}$ of V^- and V^+ respectively. Let N, R, T be the sets $\{1, 2, \ldots, n\}$; $\{3, 4, \ldots, n\}$ and $\{1, 2\}$ respectively. If X is any subset of N, let V_X be the subspace spanned by $\{v_i \mid i \in X\}$. For any subset A of R, let A' be the complementary set $R-A$.

Denote by L_A where $A \subseteq R$ the image in $L_n(q)$ of the subgroup $\tilde{L}_A$ generated by the V_A-invariant linear transformations which act as identities on V^- and $V_{A'}$. Similarly if $B \subseteq T$, let L_B denote the image in $L_n(q)$ of the V_B-invariant subgroup $\tilde{L}_B$ generated by linear transformations which act as identities on V_{T-B} and V_R. For brevity we also use the letters L, M for the groups $L_{\{1,2\}}$ and L_R respectively.

Clearly we have

(1) $L_\phi = 1$;

(2) $L_A \cong SL(|A|, q)$ where $A \subseteq T$ or $A \subseteq R$; and if A, B are non-empty subsets of R:

(3) $L_A \cong L_B$ if and only if $|A| = |B|$;

(4) $A \subseteq B$ implies $L_A \subseteq L_B$;

(5) $L_A L_B = L_A \times L_B$ if $A \cap B = \emptyset$;

(6) $\langle L_A, L_B \rangle \cong SL(|A \cup B|, q)$ if $A \cap B \neq \emptyset$.

If $i, j \in N$ and $i < j$, let h_{ij} denote the image in $L_n(q)$ of the linear transformation $\tilde{h}_{ij}$ such that

$$\tilde{h}_{ij}(v_k) = v_k \quad \text{if } i \neq k \neq j; \quad \tilde{h}_{ij}(v_i) = \lambda v_i$$

and $\tilde{h}_{ij}(v_j) = \lambda^{-1} v_j$ where λ is a primitive element of F_q. Thus h_{ij} has order $q-1$ and $\langle h_{ij} \rangle$ contains a unique involution denoted by t_{ij}.

For subsets X, Y of N, let $X + Y$ be the symmetric difference $X \cup Y - X \cap Y$. If A is any subset of even order of N, let

$$
\begin{aligned}
t_A &= 1 \text{ if } A = \emptyset; \\
&= \text{the unique involution in } Z(L_A) \text{ if } A \subseteq R; \\
&= t_{12}\, t_{A+T} \text{ if } A \supseteq T; \\
&= t_{12}\, t_{23}\, t_{A+\{1,3\}} \text{ if } A \cap T = \{1\}; \\
&= t_{23}\, t_{A+\{2,3\}} \text{ if } A \cap T = \{2\}.
\end{aligned} \tag{7}
$$

With the above definition of t_A, we observe that if X, $Y \subseteq N$, then

$$t_X\, t_Y = t_{X+Y} \tag{8}$$

and $t_X = t_{N+X}$ if n is even.

If $A = \{i, j, k, \ldots\} \subseteq N$, we also write

$$
\begin{aligned}
&t_{ijk\ldots} \text{ for } t_A \text{ if } |A| \text{ is even} \\
&L_{ijk\ldots} \text{ for } L_A \text{ if } L_A \text{ is defined.}
\end{aligned}
$$

Notice that this notation is consistent with the definition of t_{ij} introduced earlier. We emphasize t_A has no meaning unless $|A|$ is even and L_A is only defined if $A \subseteq R$ or $A \subseteq T$.

Let the center $Z(SL(n, q))$ of $SL(n, q)$ be $\langle cI \rangle$ where I is the identity transformation and $c \in F_q$. Denote the 2-part of c by c'. If c' is a square in F_q, let $\sqrt{c'}$ be a fixed element of F_q different from 1 such that $(\sqrt{c'})^2 = c'$. Let u be the image in $L_n(q)$ of the linear transformation θ with $\theta(v_1) = -\sqrt{c'}\, v_1$ and $\theta(v_i) = \sqrt{c'}\, v_i$ for all $i \neq 1$. Note that u is an involution. If c' is a nonsquare and n is even, let w be the image in $L_n(q)$ of the linear transformation θ with $\theta(v_i) = v_{i+1}$ and $\theta(v_{i+1}) = c' v_i$ for all $i = 1, 3, \ldots, n-1$. Of course the existence of u and w depends on n, c' and $d = (n, q-1)$ and is mutually exclusive. Here we shall adopt the convention that whenever a statement is made involving u (or w), it is void if u (or w) does not exist.

It is now a simple matter to compute H which is $LM\langle h_{23}\rangle$. Clearly we have

$$
\begin{aligned}
&\langle L, h_{23}\rangle \cong GL(2,q);\\
&\langle M, h_{23}\rangle \cong GL(n-2,q);\\
&[L, M]=1\\
&L\cap M=\begin{cases}1 & \text{if } n \text{ is odd}\\ \langle t_{12}\rangle & \text{if } n \text{ is even}\end{cases}\\
&\text{and}\qquad t=t_{12}.
\end{aligned}
\tag{9}
$$

We proceed next to determine the conjugate classes of involutions in H.

(1.1) **Lemma.** (i) *Suppose that n is odd. Then an involution in H is conjugate to t_A for some $A\subseteq N$. The involutions t_X and t_Y are conjugate in H if and only if $|X\cap T|=|Y\cap T|$ and $|X|=|Y|$.*

(ii) *Suppose that n is even and c' is a square in F_q. Then an involution of H is conjugate in H to t_A or to $u\,t_A$ for some $A\subseteq N$ with $|A|\leqq n/2$. The involutions t_X and $u\,t_Y$ are not conjugate in H for all X, Y in N. If A, B are subsets of N with $|A|$, $|B|\leqq n/2$, t_A and t_B are conjugate if and only if either $|A\cap T|=|B\cap T|$ and $|A|=|B|$ or $|A\cap T|\equiv|B\cap T|\equiv 0$ (modulo 2) and $|A|=|B|=n/2$.*

(iii) *Suppose n is even and c' is a nonsquare in F_q. Then an involution of H is conjugate in H to t_A if $(-c')^{n/2}\neq 1$ and to t_A or to w if $(-c')^{n/2}=1$ where $A\subseteq N$ with $|A|\leqq n/2$. The involutions w and t_X are not conjugate in H for all $X\subseteq N$. If A, B are subsets of N with $|A|, |B|\leqq n/2$, then t_A is conjugate to t_B if and only if either $|A\cap T|=|B\cap T|$ and $|A|=|B|$ or $|A\cap T|\equiv|B\cap T|\equiv 0$ (modulo 2) and $|A|=|B|=n/2$.*

Proof. To prove this lemma, it is more convenient to work with the inverse image $\tilde{H}$ in $GL(n,q)$ of $H(n,q)$. Let $\tilde{L}_{12}$, $\tilde{L}_R$ and $\tilde{h}_{ij}$ have the same meaning as before. Suppose a, b are linear transformations such that $a(v_2)=\lambda v_2$; $a(v_i)=v_i$ $(i\neq 2)$ and $b(v_3)=\lambda^{-1}v_3$; $b(v_j)=v_j$ $(j\neq 3)$. Set $X_1=\langle\tilde{L}_{12}, a\rangle$ and $X_2=\langle\tilde{L}_R, b\rangle$. Note that $X_1\cong GL(2,q)$; $X_2\cong GL(n-2,q)$; $[X_1, X_2]=1$ and $\tilde{H}\subseteq X_1X_2$.

Let z be an involution in $H(n,q)$. We may choose a preimage $s=s_1s_2$ of z in $\tilde{H}$ where $s_i\in X_i$ and $s^2=xI$ with $x\in\langle c'\rangle$ and I is the identity mapping. If v is a non-zero vector of V^+, the subspace $U=\langle v, s_2(v)\rangle$ is $\langle s_2\rangle$-invariant and we distinguish two cases:

(1) dimension of $U=1$. Then x is necessarily a square and $s_2(v)=\pm\sqrt{x}\,v$ where $\sqrt{x}\in F_q$ such that $(\sqrt{x})^2=x$ and $\sqrt{x}\neq x$.

(2) dimension of U is 2. If x is a square, then U is $\langle s_2\rangle$-reducible and we can choose a basis of U such that relative to this basis, s_2 restricted to U corresponds to

$$\begin{pmatrix}\sqrt{x} & \\ & -\sqrt{x}\end{pmatrix}.$$

On the other hand if x is a nonsquare, then $x \in \langle c' \rangle - \langle c'^2 \rangle$ and U is $\langle s_2 \rangle$-irreducible. Relative to the basis $\{v, s_2(v)\}$, s_2 restricted to U corresponds to

$$\begin{pmatrix} 0 & x \\ 1 & 0 \end{pmatrix}.$$

Since s_2 has 2 power-order, it follows from Maschke's theorem that relative to a suitable basis of V^+, s_2 restricted to V^+ corresponds to one of the following matrices

$$\begin{pmatrix} -\sqrt{x}\, I_m & \\ & \sqrt{x}\, I_{n-2-m} \end{pmatrix}; \quad \begin{pmatrix} 0 & x & & & & & \\ 1 & 0 & & & & & \\ & & 0 & x & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & x \\ & & & & & 1 & 0 \end{pmatrix}$$

where I_i denotes the identity matrix of size i. Because $\langle \tilde{L}_R, ab \rangle$ is transitive on the bases of V^+, there is an $g \in \langle \tilde{L}_R, ab \rangle$ such that $g^{-1} s_2 g$ restricted to V^+ and relative to $\{v_3, v_4, \ldots, v_n\}$ corresponds to one of the above matrices. A similar argument shows that s_1 restricted to V^- and relative to a suitable basis of V^- corresponds to

$$\begin{pmatrix} \sqrt{x} & \\ & \pm\sqrt{x} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & x \\ 1 & 0 \end{pmatrix}.$$

Moreover 2-elements θ in $GL(2, q)$ such that $\theta^2 = x I_2$ are all conjugate in $SL(2, q)$. Hence there is $g' \in \tilde{L}_{12}$ such that $s_1^{gg'}$ restricted to V^- and relative to $\{v_1, v_2\}$ corresponds to one of the above matrices. Thus $(s_1 s_2)^{gg'}$ relative to $\{v_1, v_2, \ldots, v_n\}$ is represented by

$$\begin{pmatrix} \sqrt{x} & & & \\ & \pm\sqrt{x} & & \\ & & -\sqrt{x}\, I_m & \\ & & & \sqrt{x}\, I_{n-2-m} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & x & & & & & \\ 1 & 0 & & & & & \\ & & 0 & x & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & x \\ & & & & & 1 & 0 \end{pmatrix}.$$

Finally it is easy to check that

$$\begin{pmatrix} 0 & (c')^{1+i} \\ (c')^i & 0 \end{pmatrix}$$

is conjugate in $SL(2, q)$ to

$$\begin{pmatrix} 0 & (c')^{1+2i} \\ 1 & 0 \end{pmatrix}$$

for all i and so

$$\begin{pmatrix} 0 & x & & & & & \\ 1 & 0 & & & & & \\ & & 0 & x & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & x \\ & & & & & 1 & 0 \end{pmatrix}$$

is conjugate in $\tilde{L}_{12} \times \tilde{L}_R$ to

$$\begin{pmatrix} 0 & y & & & & & \\ 1 & 0 & & & & & \\ & & 0 & y & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & y \\ & & & & & 1 & 0 \end{pmatrix}$$

modulo $\langle c I \rangle$ for all x, y in $\langle c' \rangle - \langle c'^2 \rangle$.

The assertions of the lemma follow by a straight-forward computation.

(1.2) **Lemma.** *Let A be a proper subset of even order in N different from R. Denote by C_A the subgroup $C(t, t_A)''$ if $|A \cap T| = 1$; $|A| = n/2$; $q > 3$ and the subgroup $C(t, t_A)'$ in all other cases. Denote by D_A the subgroup $C(t, u\,t_A)''$ if either $|A \cap T| = 0$; $|A| + 1 = n/2$, $q > 3$ or $|A \cap T| = 2$; $|A| - 1 = n/2$; $q > 3$ and the subgroup $C(t, u\,t_A)'$ in all other cases. Set $B = A + A \cap T$ and $B' = R - B$.*

Then

i) $C_A = L L_B L_{B'}$ *if* $|A \cap T| = 0$ *or* 2;

ii) $C_A = L_B L_{B'}$ *if* $|A \cap T| = 1$;

iii) $D_A = L L_B L_{B'}$ *if* $|A \cap T| = 1$;

iv) $D_A = L_B L_{B'}$ *if* $|A \cap T| = 0$ *or* 2.

Proof. Let $i \in B$ and $j \in B'$ if $|B'| \neq \emptyset$. A simple computation shows that $C(t, t_A) = L L_B L_{B'} \langle h_{23}, h_{ij} \rangle$ if $|A \cap T| = 0$ or 2 and $C(t, t_A) = L_B L_{B'} \langle h_{12}, h_{23}, h_{ij} \rangle$ or $L_B \langle h_{12}, h_{23} \rangle$ (according to whether $B \neq R$ or $B = R$) in all cases where $|A \cap T| = 1$ except when $|A| = n/2$. In the exceptional case $C(t, t_A) = L_B L_{B'} \langle h_{12}, h_{23}, h_{ij}, z \rangle$ where z is the image of the linear transformation θ such that $\theta(v_1) = v_2$, $\theta(v_2) = -v_1$; $\theta(v_m) = v_{n-m+3}$; $\theta(v_{n-m+3}) = -v_m$, $m = 3, 4, \ldots, n+2/2$. The results for C_A follows immediately by computation.

A similar argument holds for D_A.

(1.3) **Lemma.** *Suppose n is even and w exists. Then $C(t)' \cap C(w) = C_L(w)\, C_M(w) \langle w' \rangle$ where $C_L(w)$ is cyclic of order $q+1$; $C_M(w)$ is isomorphic to* *an extension of* $SL(n-2/2, q^2)$ *by a cyclic group of order* $q+1$ *and* w' *is an involution normalizing* $C_L(w)$ *and* $C_M(w)$.

Proof. Let U be a vector space of even dimension m over F_q; g an element in $GL(m, q)$ such that $g^2 = c' I$ where c' is a non-square. We may regard U as a vector space of dimension $m/2$ over F_{q^2} by defining

$$(\alpha+\beta\eta)\, v = \alpha v + \beta g(v)$$

where $\alpha, \beta \in F_q$; $\eta \in F_{q^2}$ with $\eta^2 = c'$ and $v \in U$. An F_q-linear transformation commutes with g if and only if it is F_{q^2}-linear. Hence it follows that $C = C_{GL(m,q)}(g)$ is isomorphic to $GL(m/2, q^2)$ and hence $C \cap SL(m, q)$ is isomorphic to an extension of $SL(m/2, q^2)$ by a cyclic group of order $q+1$.

The result now follows immediately, bearing in mind that when we work with the inverse image $\tilde{H}$ of $H(n, q)$ in $GL(n, q)$, there exists $\tilde{w}'$ in $(\tilde{H})'$ mapping by conjugation

$$\begin{pmatrix} 0 & c' & & & & & \\ 1 & 0 & & & & & \\ & & 0 & c' & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & c' \\ & & & & & 1 & 0 \end{pmatrix} \quad \text{into} \quad \begin{pmatrix} 0 & -c' & & & & & \\ -1 & 0 & & & & & \\ & & 0 & -c' & & & \\ & & -1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & -c' \\ & & & & & -1 & 0 \end{pmatrix}$$

§ 2. Non Fusion of Involutions

First we prove a simple proposition due essentially to Wong [15].

(2.1) **Proposition.** *Let* X *be a group with subgroups* $X_1, X_2, \ldots, X_m$ *such that* $X = X_1 X_2, \ldots, X_m$; $[X_i, X_j] = 1$ *if* $i \neq j$ *and* $X_i \cong SL(n_i, q)$ *where* $n_i \geqq 2$ *and* q *is odd. Then* X_i *is uniquely determined by these conditions.*

Proof. Let $Z(X)$ be the center of X. Then

$$X/Z(X) = X_1 Z(X)/Z(X) \times X_2 Z(X)/Z(X) \times \cdots \times X_m Z(X)/Z(X)$$

and $X_i\, Z(X)/Z(X) \cong X_i/Z(X_i)$ is isomorphic to $L_{n_i}(q)$ and so is an indecomposable group with trivial center. By the Krull-Schmidt theorem, the subgroups $X_i\, Z(X)$ are uniquely determined. Hence $X_i = (X_i\, Z(X))'$ are uniquely determined if $q \neq 3$. For $q = 3$, X_i is the least normal subgroup of $X_i\, Z(X)$ having an index a power of 2, therefore X_i are also uniquely determined. This completes the proof.

If Y is a group and $Y = Y_1 \ldots Y_n$, a decomposition also satisfying the conditions of (2.1) and θ is an isomorphism of X on Y, then it follows $n = m$ and θ maps X_i on the Y_i in some order. In particular every automorphism of X permutes the X_i in some order.

(2.2) **Lemma.** *The involution t is not conjugate to w nor to $u\,t_A$ for any subset A of N.*

Proof. We note first that $J=C(t)'\cap C(w)$ is normal in $C(t,w)$. By (1.3), we see easily that J'' is isomorphic to $SL(n-2/2, q^2)$ and that t is the unique involution in $Z(J'')$.

Suppose t is conjugate to w in G. Then there is an inner automorphism θ mapping $C(t,w)$ into $C(t)$ and w on t. By (1.1)(iii), we may assume that either $t^\theta=w$ or $t^\theta=t_A$ for some $A\subseteq N$. Now the center of a Sylow p-subgroup P of $C(t,t_A)$ has order q, q^2 or q^3 where p is the prime dividing q. If this order is q^2, its nonidentity elements lie in at least three conjugate classes in $C(t,t_A)$. On the other hand, the center of a Sylow p-subgroup of $C(t,w)$ has order q^2 and all of its nonidentity elements lie in one conjugate class in $C(t,w)$. Hence we must have $t^\theta=w$ and so $C(t,w)^\theta=C(t,w)$. Since J'' is normal in $C(t,w)$, either $(J'')^\theta=J''$ or $J''\cap(J'')^\theta\subseteq Z(J'')$. The first case gives rise to a contradiction since it would imply that $t^\theta=t$. The second case is impossible since $C(t,w)$ has only one composition factor isomorphic to $L_{(n-2)/2}(q^2)$.

Next suppose that $u\,t_A$ is conjugate to t for some subset A of N. Then n is necessary even and there is an inner automorphism θ mapping $C(t,u\,t_A)$ into $C(t)$ and $u\,t_A$ on t. By (1.1)(ii), we may assume $t^\theta=t_X$ or $t^\theta=u\,t_X$. Set $B=A+A\cap T$, $B'=R-B$; $Y=X+X\cap T$; $Y'=R-Y$. By (1.2), the fact that n is even $\geqq 6$ and the remark following (2.1), we conclude that $t^\theta\neq t_X$ for any X. Thus $t^\theta=u\,t_X$ and so $(D_A)^\theta=D_X$. It implies that $|A\cap T|\equiv|X\cap T| \pmod 2$ and so either $(L\,L_B\,L_{B'})^\theta=L\,L_Y\,L_{Y'}$ or $(L_B\,L_{B'})^\theta=L_Y\,L_{Y'}$. The first possibility occurs only if $|B|$, $|B'|$, $|Y|$, $|Y'|$ are all odd. This means $L^\theta=L$ and $t^\theta=t$, a contradiction. The other possibility implies that $|B|$, $|B'|$, $|Y|$, $|Y'|$ are all even. So $t^\theta=(t_B\,t_{B'})^\theta=t_Y\,t_{Y'}=t$ by (2.1), again a contradiction. This completes the proof.

(2.3) **Lemma.** *Let A be a subset of N of even order greater than 2 and $|A\cap T|=0$ or 2.*

i) *if n is odd, then t_A is not conjugate to t;*

ii) *if n is even and $|N-A|\geqq 4$, then t_A is not conjugate to t.*

Proof. Assume the contrary. Then there exists an inner automorphism mapping $C(t,t_A)$ into $C(t)$ and t_A on t. We may assume that $t^\theta=t_X$ by (1.1) and (2.2). By (2.1) and the assumption that $|A|\geqq 4$, it follows $|X\cap T|=0$ or 2. Let $B=A+A\cap T$; $B'=R-B$; $Y=X+X\cap T$; $Y'=R-Y$. Thus we must have $(C_A)^\theta=C_X$ i.e. $(L\,L_B\,L_{B'})^\theta=L\,L_Y\,L_{Y'}$.

Suppose n is odd. Then $|B'|$ and $|Y'|$ are odd. Since θ is induced by element not in H; we get $L_B^\theta=L$ and $t_B^\theta=t$. This means that $A=B$. But $L_B\cong SL(|B|,q)\cong L$ by (2). So $|A|=2$, a contradiction to our assumption $|A|\geqq 4$.

Next suppose that n is even. By our assumption on $|A|$ and $|N-A|$, $n\geqq 8$. By (2.1) either $L_B^\theta=L$ or $(L_{B'})^\theta=L$. If $L_B^\theta=L$, then $t_B^\theta=t$ and $|B|=2$. It follows that either $A=B$ or $B=N-A$. Both cases are clearly impossible. If $(L_{B'})^\theta=L$, then $(t_{B'})^\theta=(t\,t_B)^\theta=t$ and $|B'|=2$. A cannot be equal to B, otherwise we would get $(t\, t_B)^\theta = t^\theta\, t = t$, an impossibility. Therefore $A = B \cup T$ and $B' = N - A$, which is a contradiction to the fact $|N - A| \geqq 4$.

The next result will prove useful later.

(2.4) **Proposition.** *Let be a subset of R of order 2. If x is an element in G with $t^x = t_A$, then $L^x = L_A$.*

Proof. The result is a consequence of (2.1), (2.2), (2.3) and the normality of L in H.

§ 3. Non Simple Cases of Types A and B

An application of Glauberman theorem [6] due to Wong [15] produces our first nonsimple group.

(3.1) **Theorem A.** *Let G be a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n, q)$. Suppose that t is not conjugate to another involution in $C(t)$. Then $G = C(t)\, O(G)$.*

Proof. See Wong [15].

In view of the above result, we shall assume from now on that the involution t of G satisfies the following condition.

(*) *t is conjugate in G to another involution of $C(t)$.*

For the rest of this section, we also assume that

(**) *t is not conjugate to another involution of $C(t)'$.*

(3.2) **Lemma.** *Let A be an even subset of N such that $|A \cap T| = 1$ and $|A| \leqq n-2$. Then t is not conjugate to t_A.*

Proof. Suppose there is an $x \in G$ with $t^x = t_A$. Let $i \in A - T$ and $j \in N - A \cup T$. If $t_{ij} \in C(t_A)'$, then $t_A\, t_{ij} \in C(t_A)'$. Since t_A is conjugate to $t_A\, t_{ij}$ in $\langle M, h_{23} \rangle$, this would imply that t is conjugate to another involution in $C(t)'$, a contradiction to (**).

Hence t_{ij} belongs to $C(t_A) - C(t_A)'$. From the structure of H and (2.1), we easily see that an involution a in $H - H'$, not of the form $u\, t_X$ where $|X \cap T| = 1$, is conjugate in H to $a\, t$. We claim that t_{ij} is not conjugate in $C(t_A)$ to such $u\, t_X$. Assume the contrary then $t_{ij} = v^x$ where v is in H and is conjugate to $u\, t_X$ in H. By (2.2), t is not conjugate to $t\, v$ in G. Thus $t_A = t^x$ is not conjugate to $(t\, v)^x = t_A\, t_{ij}$ in G. But this is a contradiction since t_A is already conjugate to $t_A\, t_{ij}$ in H. Thus we conclude that t_{ij} is conjugate to $t_A\, t_{ij}$ in $C(t_A)$ and so t_{ij} is conjugate to t_A and t. This contradicts (**) since $t_{ij} \in C(t)'$ and $t_{ij} \neq t$.

An immediate consequence of the above lemma is the following.

(3.3) **Corollary.** *Let G be a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n, q)$ where $n \geqq 5$ and q is odd. Assume that t satisfies (*) and (**), then n is odd. Furthermore, the only conjugate class in H fused with that of t is the one with representative $t_{23\ldots n}$.*

For the rest of this section, n is an odd integer.

(3.4) **Lemma.** *The centralizer $C(M)$ of M is isomorphic to* (i) $PGL(3,q)$; (ii) $PSL(3,q)$; (iii) $SL(3,q)$; (iv) $PSL(3,q)\times Z_3$ *where* $q\equiv 4 \pmod 9$ *or* $q\equiv 7 \pmod 9$ *or* $q\equiv 1 \pmod 3$ *but* $\not\equiv 1 \pmod 9$; *or* (v) M_{11}.

The number $d=(n,q-1)$ is 1 except in the case $q\equiv 1 \pmod 3$ and $C(M)\cong PSL(3,q)$ in which case $d=3$.

Proof. A simple computation shows that $X=C(M)\cap H$ is equal to $\langle L,h\rangle$ where $h=h_{23}^{n-2}h_{34}^{n-3}\dots h_{n-1,n}$. This group X is isomorphic to $GL(2,q)/Z$ where $Z\subseteq Z(GL(2,q))$ (note that n is odd). By (3.3), there exists $x\in G$ such that $t^x=t_A$ where $A=\{2,3,\dots,n\}$. Now $C(t_A)$ contains M and M^x. By the structure of H, $M=M^x$. In particular it follows that $L^x\subseteq C(M)$.

Since $C(t_A)\cap C(M)$ contains a unique subgroup isomorphic to $SL(2,q)$, this must be L^x. Hence it follows h_{12} normalizes L^x and $|\langle h_{12}\rangle\cap L^x|$ is odd. Thus $\langle h_{12},L^x\rangle$ contains a 2-subgroup isomorphic to a Sylow 2-subgroup of $GL(2,q)/Z$. On the other hand, a Sylow 2-subgroup of $C(t)\cap C(M)$ contains the unique involution t in its center; it is thus also a Sylow 2-subgroup of $C(M)$. From the fact $C(M)'\supseteq\langle L,L^x\rangle$, it follows $C(M)$ has no subgroup of index 2. We may then apply Brauer theorem [2, 3] to get our result.

(3.5) **Lemma.** *Let* $G_0=MC(M)\langle h_{23}\rangle$; $d=(n,q-1)$; $e=(n-2,q-1)$. *Then one of the following holds.*

i) $G_0\cong PGL(3,q)\times SL(n-2,q)$; $d=1=e$;

ii) $G_0\cong PSL(3,q)\times SL(n-2,q)$; $d=1$ *if* $q\not\equiv 1 \pmod 3$ *and* $d=3$, *if* $q\equiv 1 \pmod 3$; $e=1$;

iii) $G_0\cong SL(3,q)\times SL(n-2,q)$; $d=1=e$;

iv) $G_0\cong$ *an extension of the central product of* $SL(3,q)$ *and* $SL(n-2,q)$ *by* Z_3; $d=1$; $e=3$ *and* $q\equiv 1 \pmod 3$;

v) $G_0\cong PSL(3,q)\times Z_3\times SL(n-2,q)$; $d=1=e$; $q\equiv 4$ *or* $7 \pmod 9$ *or* $q\equiv 1 \pmod 3$ *but* $q\not\equiv 1 \pmod 9$;

vi) $G_0\cong$ *a non-central extension of* $PSL(3,q)\times SL(n-2,q)$ *by* Z_3; $d=1$; $e=3$ *and* q *same as in* (v);

vii) $G_0\cong M_{11}\times SL(n-2,3)$; $d=1=e$.

Proof. Since h_{23} normalizes M, it also normalizes $C(M)$ and so G_0 is a group. Clearly $[G_0: MC(M)]=e$. Further we observe that $C(M)\cap M\subseteq Z(C(M))$ and $\langle h^{q-1/e}\rangle=C(M)\cap M$. The lemma follows from enumerating all possible cases of (3.4) satisfying above conditions.

(3.6) **Theorem B.** *Let G be a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n,q)$ where $n\geqq 5$ and q is odd. Suppose t is conjugate to another involution in $C(t)$ but not conjugate to other involutions in $C(t)'$. Then $G=G_0$ of* (3.5).

Before proving Theorem B, we state the following well-known facts on dihedral groups.

(3.7) *Let t_1, t_2 be nonconjugate involutions of a group X. Denote by Y the dihedral group $\langle t_1, t_2\rangle$. Then* (i) *$\langle t_1 t_2\rangle$ contains a unique involution v which lies in $Z(Y)$;*

(ii) *there are three conjugate classes of involutions in Y with representatives t_1, t_2 and v;*

(iii) *if 4 is the maximum 2-power divisor of $|Y|$, then $t_1 v$ is conjugate in Y to t_2 and $t_2 v$ is conjugate to t_1 in Y;*

(iv) *if 8 divides $|Y|$, then $t_1 v$ is conjugate to t_1 and $t_2 v$ is conjugate to t_2 in Y.*

Proof of (3.6). We shall show first that $C(t\,t_A) \subseteq G_0$ for all $A \subseteq R$. We have $C_{G_0}(t\,t_A) = C_M(t_A)\, C_{C(M)}(t)\, \langle h_{23}\rangle$. Let $h = h_{23}^{n-2} h_{34}^{n-3} \ldots h_{n-1,n}$ and S_1, S_2 be Sylow 2-subgroups of $\langle L, h\rangle$ and $C_M(t_A)$ respectively. Thus S_1 is either quasi-dihedral $(q \equiv -1 \pmod 4)$ or wreathed $(q \equiv 1 \pmod 4)$. In any case the center of S_1 contains the unique involution t. Hence it follows that $S_1 S_2$ is a Sylow 2-subgroup of $C(t\,t_A)$ by assumption (**) and a usual argument.

Suppose that t is conjugate in $C(t\,t_A)$ to another involution in S_1, say x. Then there exists 2-element $g \in C(t\,t_A) \cap C(x) - H$ normalizing $C_{S_1}(x)\, S_2$. Since $\Omega_1(C_{S_1}(x))$ is a four-group and no involution in $C_{S_1}(x)\, S_2 - \Omega_1(C_{S_1}(x))$ is conjugate to t by (3.3), $t^g = t\,x$. On the other hand $(C_{S_1}(x)\, S_2)' = S_2'$ and $t_A \in S_2'$. Therefore $t_A^g \in S_2$. It follows that $(t\,t_A)^g = t\,x\,t_A^g = t\,t_A$. This implies that $x = t_A\, t_A^g \in S_2$, a contradiction to (**).

Hence we have shown that t is not conjugate in $C(t\,t_A)$ to another involution in $S_1 S_2$. By the Glauberman theorem [6], $C(t\,t_A) = (C(t) \cap C(t\,t_A))\, O(C(t\,t_A))$. Set $K = O(C(t\,t_A))$. Then $\langle t, t_B\rangle$ acts on K where $B = \{2, 3, \ldots, n\}$. Therefore $K = C_K(t)\, C_K(t_B)\, C_K(t\,t_B)$. The factors all lie in G_0. So $K \subseteq G_0$ and also $C(t\,t_A) \subseteq G_0$.

Incidentally we have also shown that $t\,t_A$ are never conjugate to t_X for $A, X \subseteq R$ since $C(t_X) \supseteq C(M)$ whereas $C(t\,t_A) \not\supseteq C(M)$.

Let $x \in G$ and z an involution in M. By (3.7), $\langle t\,z^x\rangle$ contains a unique involution v and $t\,v$ is conjugate to t or to z. Since $v \in H$, v is conjugate in H to t, $t\,t_A$ or t_A for some $A \subseteq R$. The first two cases mean that $z^x \in C(v)$ and $C(v)$ is in G_0. The last case is impossible because it would mean that $t\,t_A$ is conjugate to t or to z. This has already been ruled out. So we have $z^x \in G_0$ and thus z^x must lie in M. It means that x normalizes M and so also normalizes $C(M)$. In particular $t^x \in C(M)$. Since $C(M)$ has only one class of involutions, there exists $h \in C(M)$ such that $x\,h \in C(t)$ which is in G_0. Therefore $x \in G_0$. The proof is now complete.

§ 4. Non Simple Cases of Type C

In view of § 3, we can now assume that t is conjugate to another involution of H'.

(4.1) **Lemma.** *When $n = 5$, either t_{23} or t_{2345} (but not both) is conjugate to t. When $n \geqq 6$, t_{23} is conjugate to t.*

Proof. Because of (2.3), we may choose a suitable $x\in G$ such that $t^x=t_{n-1,n}$ and $t^x_{n-1,n}=t$. By (2.4) and (2.1), $L^x=L_{n-1,n}$; $L^x_{n-1,n}=L$ and x normalizes $L_{34\ldots n-2}$.

We claim that t^x_{23} is not conjugate to w nor to $u\,t_A$ for any $A\subseteq N$. Set $X=C(t,t_{n-1,n},t_{23})$ and $Y=X^x=C(t,t_{n-1,n},t^x_{23})$. Replacing x with a suitable element in $x\,C(t,t_{n-1,n})$, we can suppose that $t^x_{23}=w$ or $t^x_{23}=u\,t_A$. The first case is not possible since it implies that $L\not\subseteq Y$ whereas we know that $L=L^x_{n-1,n}\subseteq X^x=Y$. If $t^x_{23}=u\,t_A$, then $|A\cap T|=1$ and $|A\cap\{n-1,n\}|=1$ because $L\subseteq Y$ and $L_{n-1,n}\not\subseteq Y$. On the one hand, we have $X'=L_{n-1,n}L_{4\ldots n-2}$ and on the other, $Y'=L L_B L_{B'}$ where $B=A-\{1,2,n-1,n\}$ and $B'=N-(A\cup\{1,2,n-1,n\})$. However $|B|$ and $|B'|$ are both even (since n is even) and yet $|\{4,\ldots,n-2\}|$ is odd. So $X'\not\cong Y'$, a contradiction.

Hence we may suppose that $t^x_{23}=t_A$ for some $A\subseteq N$. Again from $L\subseteq Y$ and $L_{n-1,n}\not\subseteq Y$, we get that $|A\cap\{1,2\}|=0$ or 2 and $|A\cap\{n-1,n\}|=1$. Since X' is isomorphic to a direct product of $SL(2,q)$ and $SL(n-5,q)$ and $|A|$ is even, we must have $|A\cap(N-\{1,2,n-1,n\})|$ equal to 1 when n is odd and to 1 or $n-5$ when n is even. Since $t_A=t_{N-A}$ when n is even, thus we may assume that $|A\cap(N-\{1,2,n-1,n\})|=1$ in all cases and still have $|A\cap\{1,2\}|=0$ or 2 and $|A\cap\{n-1,n\}|=1$. Replacing x with a suitable element in $x\,L_{34\ldots n-2}$ we may suppose that $t^x_{23}=t_{n-2,n-1}$ or $t\,t_{n-2,n-1}$. The second case means that $(t_{23\,n-1,n})^x=t_{n-2,n-1}$.

When $n=5$, the assertion is clear by (2.3). If $n=6$, there is nothing to prove since t_{23} and $t_{23\,n-1,n}$ are conjugate. Assume then $n\geqq 7$. Since $X'=L_{45\ldots n-2}L_{n-1,n}$ and $Y'=L_{12}L_{34\ldots n-3}$. Therefore $(L_{45\ldots n-2})^x=L_{34\ldots n-3}$, $t^x_{45}\in L_{34\ldots n-3}$ and by (2.3), t^x_{45} is conjugate in $L_{34\ldots n-3}$ to t_{45} when $n\geqq 8$ and $t^x_{45}=t_{34}$ when $n=7$. Clearly we may suppose $t^x_{45}=t_{45}$ when $n\geqq 8$ and $t^x_{45}=t_{34}$ when $n=7$. If $t^x_{23}=t\,t_{n-2,n-1}$, then $(t_{2345})^x=t\,t_{45}\,t_{n-2,n-1}$ when $n\geqq 8$ and $(t_{2345})^x=t\,t_{3456}$ when $n=7$. Hence by (2.3), t_{2345} is not conjugate to t if $n\neq 8$. This is a contradiction, since $t_{2345}\sim t_{23\,n-2,n-1}$ in H and $(t_{23\,n-2,n-1})^x=t_{n-2,n-1}$ which is conjugate to t.

Let $n=8$ and suppose that $(t_{2378})^x=t_{67}$. Choose an $z\in C_M(t_{78})$ such that $(t_{2345})^{xz}=(t\,t_{67}\,t_{45})^z=t_{67}$. Set $x_1=x\,z$. Thus x_1 maps $C(t,t_{78},t_{2345})'=L_{345}L_{78}$ onto $C(t,t_{78},t_{67})'=L_{12}L_{345}$. Therefore $(L_{345})^{x_1}=L_{345}$. In particular there exists $g\in L_{345}$ such that $x_1\,g$ centralizes t_{34} and hence normalizes $C(t_{34})\cap L_{345}=\langle L_{34},h_{45}\rangle$. By the structure of $C(t)$, we see that there is an element $h\in\langle L_{34},h_{45}\rangle$ such that $x_1\,g\,h=x_2$ centralizes $\langle L_{34}.h_{45}\rangle$.

Let θ be the image in $H(n,q)$ of the linear transformation which interchanges v_1 and v_2; v_3 and v_8; v_4 and v_7; v_5 and v_6. Since $\langle L,t_{2345}\rangle^{x_2\theta}=\langle L_{34},t_{45}\rangle$, $C(L,t_{2345})\cong C(L_{34},t_{45})$. We observe that L_{345} is normal in $C(L,t_{2345})$ but $(L_{345})^{x_2\theta}=L_{678}$ is not normal in $C(L_{34},t_{45})$ since we have $x_2\in C(L_{34},t_{45})$ which maps L_{78} into L_{12}. This completes the proof.

For the rest of this section, we shall assume that $n=5$ and t is conjugate in G to t_{2345}.

(4.2) **Lemma.** *The centralizer $C(M)$ of M in G is isomorphic to $SL(3,q)$. Either $C(M)\,M\,\langle h_{23}\rangle\cong SL(3,q)\times SL(3,q)$; $q\not\equiv 1\ (\mathrm{mod}\ 3)$ and $q\not\equiv 1\ (\mathrm{mod}\ 5)$ or*

$C(M) M \langle h_{23} \rangle$ is isomorphic to an extension of the central product of two $SL(3, q)$ by Z_3; $q \equiv 1 \pmod 3$ and $q \not\equiv 1 \pmod 5$.

Proof. Let x be an element of G which interchanges t and t_{45} and maps t_{2345} to t_{34} by conjugation. (See (4.1).) Let $\tilde{L} = x L_{34} x^{-1}$ and $C = \langle L, \tilde{L} \rangle$. Then $C = x L_{345} x^{-1}$ and hence $C \cong SL(3, q)$. Let y be an element of C which interchanges t and t_{2345} by conjugation, therefore y normalizes $C(t, t_{2345})'$ which is M. Hence $C \subseteq C(M)$. The result then follows by Brauer's theorem [1, 2, 3] as in (3.4).

(4.3) **Lemma.** *The group G contains an involution z normalizing $G_0 = MC(M) \langle h_{23} \rangle$ and interchanges M and $C(M)$ by conjugation.*

Proof. Let x be the element introduced in (4.2). We have already shown that $M^x = C(M)$. Now $C(M)^x = M^{x^2} = M$ since $x^2 \in H$ and $M \lhd H$. Since x interchanges t and t_{45} by conjugation, $x \in N(C(t, t_{45}))$ which is $L_{12} L_{45} \langle h_{23}, h_{34} \rangle$. In particular, $h_{23}^x \in G_0$ and so x normalizes G_0.

Since $x^2 \in H$ and $H \subseteq G_0$, by taking a suitable odd power of x, we may suppose that x is a 2-element. Let S be a Sylow 2-subgroup of $G_0 \langle x \rangle$ containing x. Then $S = S_1 S_2 \langle x \rangle$ where S_1 is a Sylow 2-subgroup of M containing t and $S_2 = S_1^x$. Therefore $x^2 \in S_1 \times S_2$. Let $x^2 = s_1 s_2$ where $s_i \in S_i$. Since $x^2 = (x^{-1} s_1 x)(x^{-1} s_2 x) = s_1 s_2$, it follows $x^{-1} s_1 x = s_2$ and $x^{-1} s_2 x = s_1$ since $S_1 \cap S_2 = 1$. Let $z = x s_1^{-1}$. Then z is the involution with the desired properties.

(4.4) **Lemma.** *G contains a subgroup $\tilde{G}$ of index 2 such that $G = \tilde{G} \langle z \rangle$.*

Proof. As in (4.3), let $S = \langle z \rangle S_1 S_2$ be a Sylow 2-subgroup of $G_0 \langle z \rangle$ where S_1, S_2 are the Sylow 2-subgroups of M, $C(M)$ respectively. S_i is either quasi-dihedral or wreathed. In any case $\Omega_1(Z(S)) = \langle t t^z \rangle$.

If S were not a Sylow 2-subgroup of G, then there would be a 2-element $g \notin G_0$ normalizing S and $g^2 \in S$. Let $K = S_1 S_2$. Clearly $[K : K \cap K^g] \leqq 2$. In particular $K \cap K^g \supseteq K' = S_1' \times S_2'$ and therefore $K \cap K^g$ contains t. Both t and t^g belongs to $K \cap K^g$. By the structure of $G_0 \langle z \rangle$, t and t^g are already conjugate in $G_0 \langle z \rangle$. There is an element h in $G_0 \langle z \rangle$ such that $t^{gh} = t$ and hence $g \in G_0 \langle z \rangle$ since $C(t) \subseteq G_0$, a contradiction.

Suppose that G has no subgroup of index 2. Then by a result of Thompson [7], z is conjugate to an involution in $S_1 S_2$ i.e. to t or to $t t^z$. We have $C(z) \cap G_0 \langle z \rangle = \langle z \rangle J$ where $J = \{x x^z \mid x \in M\}$. If z is conjugate to t, then J is conjugate to a factor of M. This would then imply that t is conjugate to $t t^z$, a contradiction to (2.3). Suppose next that z is conjugate to $t t^z$. Then there is a 2-element $g \in C(z) - C(z) \cap G_0 \langle z \rangle$ normalizing $\langle z \rangle \tilde{J}$ where $\tilde{J} = \langle s_1 s_1^z \mid s_1 \in S_1 \rangle$ and $\tilde{J}$ is quasi-dihedral or wreathed. Clearly $(\langle z \rangle \tilde{J})' = (\tilde{J})'$ and it follows $g \in C(t t^z)$, in contradiction to our choice of g.

By the structure of $MC(M) \langle z \rangle$, $C(M) M \subseteq \tilde{G}$. Since $S_1 S_2 \subseteq C(M) M$ and $S_1 S_2 \langle z \rangle$ is a Sylow 2-subgroup of G, $z \notin \tilde{G}$. This completes the proof.

(4.4) **Theorem C.** *Let G be a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n, q)$ where q is odd. Suppose t is conjugate to another involution in $C(t)'$ and to t_{2345}. Then $n = 5$, and either G is isomorphic to*

$SL(3, q) \wr Z_2$, $(5, q-1)=1=(3, q-1)$, or G is isomorphic to an extension of the central product of two $SL(3, q)$ by Z_3, $(5, q-1)=1$ and $(3, q-1)=3$.

Proof. Let $\tilde{G}$ be a subgroup of index 2 in G. Since $z \notin \tilde{G}$, $\tilde{G}$ satisfies the hypothesis of Theorem B. Hence the result follows from Theorem B since we already know that $C(M) \cong SL(3, q)$.

§ 5. The Simple Case

To complete the proof of the Main Theorem, it remains to consider the following case:

G is a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n, q)$ where $n \geqq 5$ and q is odd and t_{23}, t_{34} are conjugate to t in G.

(5.1) **Lemma.** *Let g be an element in G such that $t^g = t_{23}$ and let $L_{23} = L^g$. Then the subgroup G_0 generated by L_{23} and $C(t)$ is isomorphic to $L_n(q)$.*

Proof. As usual we may choose a suitable g such that $t^g = t_{23}$; $t_{23}^g = t$ and still $L_{23} = L^g$ by (2.4). It follows that g normalizes $C(t, t_{23})'$ and thus both L and L_{23} centralizes $L_{45 \ldots n}$.

Next we note that there exists an $x \in G$ which interchanges t and $t_{n-1,n}$ and maps $t_{n-2,n-1}$ to t_{23} by conjugation. (See (4.1).) Hence $(L_{n-2,n-1,n})^x = \langle L, L_{23} \rangle$ since $(L_{n-1,n})^x = L$ and $(L_{n-2,n-1})^x = L_{23}$ by (2.4). Therefore $\langle L, L_{23} \rangle \cong SL(3, q)$.

By the structure of $C(t)$, we easily check that if v is an involution in $C(t)$ such that all involutions in $\langle t, v \rangle$ are conjugate to t, then $\langle t, v \rangle$ is conjugate in $C(t)$ to $\langle t, t_{23} \rangle$. It follows that $\langle t_{23}, t_{34} \rangle$ is conjugate to $\langle t, t_{23} \rangle$ in G. Therefore $\langle L_{23}, L_{34} \rangle \cong SL(3, q)$ by (2.4) and the fact $\langle L, L_{23} \rangle \cong SL(3, q)$.

By the structure of $SL(3, q)$, $C(h_{34}, t_{23}) \cap \langle L_{23}, L_{34} \rangle = J$ is a direct product of two cyclic groups of order $q-1$ and $J \subseteq L_{23} \langle h_{34} \rangle$. Let $h'_{23} = J \cap L_{23}$. Then clearly $J = \langle h'_{23}, h_{34} \rangle$. Similarly there is an element h'_{12} in L_{12} such that $\langle h'_{12}, h'_{23} \rangle$ is abelian of order $(q-1)^2$. We check that $L_{12}, L_{23}, \ldots, L_{n-1,n}$ and $h'_{12}, h'_{23}, h_{34}, \ldots, h_{n-1,n}$ satisfy the condition of Theorem 2 of [11]. A comparison of the orders of the centralizers show that $G_0 \cong L_n(q)$ for all $q \neq 3$.

For $q=3$, the proof of the quoted result can be slightly modified to show that $G_0/Z(G_0)$ is isomorphic to $L_n(3)$. Again it follows that $Z(G_0)=1$ by a comparison of the orders of the centralizers. This completes the proof.

We proceed next to prove several results with the aim of showing that the centralizers of a number of involutions all lie in G_0.

(5.2) **Lemma.** *Let S be a 2-subgroup of G_0 containing a conjugate of t and S^* be a 2-subgroup of G containing S. Then $S^* \subseteq G_0$.*

Proof. We may assume that $t \in S$ and $S^* \cap G_0 = S$. If $S^* \supset S$, there is an $x \in S^* - S$ normalizing S and so $t^x \in G_0$. Since involutions in G_0 conjugate to t in G are already conjugate in G_0. This leads to a contradiction since $C(t) \subseteq G_0$.

(5.3) **Corollary.** (i) *If v is an involution in G_0, then $C(v) \cap G_0$ contains a Sylow 2-subgroup of $C(v)$.*

(ii) *Two involutions in G_0 are conjugate in G if and only if they are conjugate in G_0.*

Proof. The first assertion is a consequence of the fact that all conjugate classes of involutions in G_0 have representatives in H and the above lemma. The other assertion follows immediately from above and a comparison of centralizers of involutions in G_0. (See [5].)

(5.4) **Lemma.** *The centralizer $C(t_A)$ is contained in G_0 where $|A|$ is 4.*

Proof. When $n=6$, there is nothing to prove. Assume that $n\neq 6$ and $A=\{1,2,3,4\}$. Let $x\in C(t_A)$. t is not conjugate to t_{45} in $C(t_A)$, for otherwise $t_{12}\,t_A=t_{34}$ would be conjugate to $t_{45}\,t_A=t_{1235}$ in contradiction to (2.3). Hence $\langle t\,t_{45}^x\rangle$ contains a unique involution v in $C(t, t_{45}^x)$ and $t\,v$ is conjugate to t in G by (3.7). Replacing x with a suitable element in $x\,C(t, t_A)$, we may assume that $v=t_B$ for some $B\subseteq N$. We distinguish three cases.

(i) $B\supseteq\{1,2\}$.

Then $|B+\{1,2\}|=2$ or $n-2$, since $t\,t_B$ is conjugate to t. Hence $|B|=4$ since $t_B\neq 1$. We may suppose that $B=\{1,2,3,4\}$ or $\{1,2,4,5\}$ or $\{1,2,5,6\}$. If $t_B=t_{1234}$, then $t\,t_B$ cannot be conjugate to t_{45} in $C(t_A)$; for otherwise we would get t_{12} conjugate to t_{45} in $C(t_A)$ which we just ruled out. By (3.7), $t_{45}^x\,t_B=(t_{1235})^x$ is conjugate to t_{45}^x, a contradiction to (2.3). If $t_B=t_{1245}$, then $t_{45}^x\in C(t_{1234}, t_{1245})$ $\subseteq C(t_{35})$ which lies in G_0. Then x is in G_0 and we are done. Finally if $t_B=t_{1256}$, first we note that t is not conjugate to $t_{12}\,t_B$ in $C(t_A)$ if $n\neq 8$, since otherwise it would imply that $t_{12}\,t_A$ is conjugate to $t\,t_B\,t_A$ which is not possible by (2.3). So by (3.7), $t\,t_B=t_{56}$ is conjugate to t_{45}^x in $C(t_A)$ if $n\neq 8$. This leads to a contradiction since $t_{1234}\in C(L_{56})$ whereas $t_{1234}\notin C(L_{45})$.

Therefore it remains to settle the case $n=8$ and $t_B=t_{1256}$. We have $C(t_A)\cap G_0=K_1\,K_2\langle h_{45}\rangle\langle z\rangle$ where $K_1=\langle L, L_{23}, L_{34}\rangle$; $K_2=L_{5678}$ and z is an involution which interchanges L and L_{78}; L_{23} and L_{67}; L_{34} and L_{56} and inverts h_{45} by conjugation. Furthermore involutions in $\langle t_A, z\rangle$ are all conjugate in G. We claim that $C(t_A)$ contains a subgroup C of index 2 with $z\notin C$. If not, z is conjugate in $C(t_A)$ to an involution v' in $K_1\,K_2\langle h_{45}\rangle$ by Thompson's result [6; 265] and (5.2). We may assume that $v'=t_{1235}$ or t_{3456}. The first possibility can be dismissed immediately since $\langle t_A, t_{1235}\rangle$ contains nonconjugate involutions. The other possibility cannot occur because $C(t_A, z)\cap G_0$ and $C(t_A, t_{3456})\cap G_0$ both contain involutions conjugate to t. Hence G_0 contains Sylow 2-subgroups of $C(t_A, z)$ and $C(t_A, t_{3456})$ by (5.2) and they are non-isomorphic. This proves our claim.

For the same reason as in previous case, $t\,t_B$ is not conjugate to t_{45}^x in $C(t_A)$. So $t\,t_B$ is conjugate in $C(t_A)$ to t and by (3.7), there exists an element a of order 4 in $\langle t\,t_{45}^x\rangle$. Thus $t^a=t_{56}$ and hence $t^{az}=t_{34}$. There is an $g\in K_1$ such that $a\,z\,g$ centralizes t. It follows $a\in C(t_A)-C$. But this is a contradiction since $a\in\langle t\,t_{45}^x\rangle$ and $t\,x^{-1}\,t_{45}\,x$ already lies in C.

(ii) $|B\cap\{1,2\}|=1$.

Then $|B|$ is necessarily 2 or $n-2$. In either case we have t_B is conjugate to t and $t_{45}^x \in C(t_B)$ which is in G_0. Therefore $x \in G_0$.

(iii) $|B \cap \{1,2\}|=0$.

Then n must be even and $|B|=n-4$. Since $t_B=t_{N-B}$, we are back to (i).

The proof is now complete.

(5.5) **Lemma.** *The centralizer $C(t_A)$ is contained in G_0 for any $A \subseteq N$.*

Proof. In view of (5.4), we may assume that $n \geqq 7$; $n \neq 8$ and $|A| \geqq 6$. Let $A=\{1,2,\ldots,m\}$. Let $B=\{m,m+1\}$ if $m \neq (n+2)/2$ and $B=\{m+1,m+2\}$ if $m=(n+2)/2$. Since $t\,t_A$ is not conjugate to $t_B\,t_A$ by (5.2), t is not conjugate to t_B in $C(t_A)$. Therefore $\langle t\,t_B^x\rangle$ contains a unique involution v in G_0 conjugate to t_X in G_0. As in (5.4), $|X|=2$, 4 or $n-2$. In all cases, $t_B^x \in C(v)$ which is in G_0. Then it follows $x \in G_0$.

(5.6) **Lemma.** *Let $n=6$. Then*

(i) *$C(u\,t_A) \subseteq G_0$ for all $A \subseteq N$.*

(ii) *$C(w) \subseteq G_0$.*

Proof. (i) A comparison of the structures of the Sylow 2-subgroups of $C(u,t)$ and $C(u,t_{23})$ shows that t and t_{23} are not conjugate in $C(u)$. Let $x \in C(u)$. Then $\langle t\,t_{23}^x\rangle$ contains an involution v in G_0 and v is conjugate in G to t. Hence $t_{23}^x \in C(v)$ which is in G_0. So $x \in G_0$. Similarly $C(u\,t_A) \subseteq G_0$ for all $A \subseteq N$.

(ii) From the structure of $L_6(q)$, $C(w) \cap G_0 = K\langle z\rangle$ where K is isomorphic to a subgroup of order $(q+1)/d \cdot |SL(3,q^2)|$ in $GL(3,q^2)/Z$ for some $Z \subseteq Z(GL(3,q^2))$ and z is an involution conjugate to w in G_0 normalizing K. As in (5.4), we conclude that $C(w)$ contains a subgroup C of index 2 with $z \notin C$. Let C_1, C_2 be members of a chief series of C such that C_1/C_2 involves $L_3(q^2)$. Since a Sylow 2-subgroup of C is a direct product of a quasi-dihedral group with a cyclic group. (Here $q \equiv -1 \pmod 4$ for the existence of w), it follows that $C_1/C_2 \cong L_3(q^2)$ and a Sylow 2-subgroup of C_2 must be cyclic. By Burnside's theorem, C_2 has a normal 2-complement $O(C_2)$. Apply Brauer Welandt result [13] to the action of $\langle t, t_{23}\rangle$ on $O(C_2)$. We see easily that $O(C_2) \subseteq G_0$. Since C/C_1 cannot involve $L_3(q^2)$ because it has cyclic Sylow 2-subgroup, $K \subseteq C_1$ and in particular we have shown that $C_1 \subseteq G_0$. If $x \in C$, then $t^x \in C_1$. Since t and t^x are already conjugate in C_1, it follows as usual that $x \in G_0$ and also $C(w) \subseteq G_0$.

Finally we are in a position to prove the following theorem and thus to complete the proof of the Main Theorem.

(5.7) **Theorem D.** *Let G be a finite group with an involution t whose centralizer $C(t)$ is isomorphic to $H(n,q)$ where $n \geqq 5$ and q is odd. Suppose that t is conjugate to another involution in $C(t)'$ and t_{23} is conjugate to t. Then G is isomorphic to $L_n(q)$.*

Proof. It suffices to show that $G=G_0$. Let $z=t_{1234}$ if $n \neq 6$ and $z=u$ or w if $n=6$. Since t is not conjugate to z, then there exists an involution $v \in G_0$ in $\langle t\,z^x\rangle$ for any $x \in G$. By (3.7), v is conjugate to t_B for some $B \subseteq N$ if $n \neq 6$ and v is conjugate to t; $u\,t_A$ or w if $n=6$. In all cases $C(v) \subseteq G_0$ by (5.5) and (5.6). Thus $x \in G_0$. This completes the proof.

References

1. Brauer, R.: On finite Desarguesian planes I. Math. Z. **90**, 117–123 (1966).
2. — On finite Desarguesian planes II. Math. Z. **91**, 124–151 (1966).
3. On finite Desarguesian planes III. Math. Z. **117**, 76–82 (1970).
4. — Suzuki, M., Wall, G.E.: A characterization of the one-dimensional unimodular groups over finite fields. Illinois J. Math. **2**, 718–745 (1958).
5. Carter, R., Fong, P.: The Sylow 2-subgroups of the finite classical groups. J. Algebra **1**, 139–151 (1964).
6. Glauberman, G.: Central elements in core-free groups. J. Algebra **4**, 403–420 (1966).
7. Gorenstein, D.: Finite groups. New York: Harper & Row 1968.
8. Held, D.: A characterization of the alternating groups of degree eight and nine. J. Algebra **7**, 218–237 (1967).
9. — The simple group related to M_{24}. J. Algebra **13**, 253–296 (1969).
10. Phan, K.W.: A characterization of four-dimensional unimodular groups. J. Algebra **15**, 252–279 (1970).
11. — A theorem on special linear groups. J. Algebra **16**, 509–518 (1970).
12. Suzuki, M.: Characterization of linear groups. Bull. Amer. Math. Soc. **75**, 1043–1091 (1969).
13. Wielandt, H.: Beziehungen zwischen den Fixpunktzahlen von Automorphismengruppen einer endlichen Gruppe. Math. Z. **73**, 146–158 (1960).
14. Wong, W.J.: A characterization of the alternating group of degree 8. Proc. London Math. Soc. **50**, 359–383 (1963).
15. — A characterization of the finite simple groups $PSp_{2n}(q)$. J. Algebra **14**, 531–551 (1970).

Dr. K.-W. Phan
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556
USA

(Received June 1, 1971)

Math Z. 124, 186 – 190 (1972)

Maximale stabile Punktmengen und minimale Überdeckungen gleichgewichtiger Hypergraphen

GEORG SCHRAGE

Für eine Klasse von Hypergraphen wird gezeigt, daß die Ordnung einer maximalen Menge nichtbenachbarter Knotenpunkte gleich der Ordnung einer minimalen Überdeckung ist. Dieses Ergebnis wurde unabhängig voneinander von Berge und Las Vergnas [2] und Schrage [4] gefunden. In der vorliegenden Arbeit werden ein neuer Beweis und einige Folgerungen aus diesem Satz angegeben. Zahlreiche Beispiele, bei denen die Frage nach maximalen Mengen nichtbenachbarter Punkte oder minimalen Überdeckungen von Hypergraphen auftritt, finden sich in [1].

Sei $X=\{x_1, x_2 \dots x_n\}$ eine endliche Menge und $\mathfrak{P}=(P_i, i\in I)$ eine Familie von Teilmengen von X. Das Paar $H=(X; \mathfrak{P})$ heißt *Hypergraph*[1] der Ordnung n, wenn

1. $P_i \neq \emptyset$ für alle $i\in I$ und
2. $\bigcup_{i\in I} P_i = X$.

Die Elemente $x_1, x_2 \dots x_n$ heißen *Knotenpunkte* des Hypergraphen, die Elemente von $\mathfrak{P}$ werden *Kanten* genannt.

Ein Hypergraph besteht also ebenso wie ein Graph aus Objekten, die als Knotenpunkte und Kanten bezeichnet werden. Der Unterschied besteht darin, daß beim Graphen zu jeder Kante höchstens zwei Knotenpunkte gehören, während eine solche Einschränkung für die Kanten eines Hypergraphen nicht gilt. Mit anderen Worten: ein Graph ist ein Hypergraph, bei dem die Ordnung jeder Kante kleiner oder gleich 2 ist.

Wir wollen zunächst einige aus der Theorie der Graphen geläufige Begriffe verallgemeinern.

Definition. In $H=(X; \mathfrak{P})$ heißen zwei Knotenpunkte x und y genau dann *benachbart*, wenn ein $P\in\mathfrak{P}$ existiert, so daß $x, y\in P$. Zwei Kanten von H heißen *benachbart*, wenn ihr Durchschnitt nicht leer ist.

Wir schreiben $x \circ y$, falls x und y benachbart sind und $x \not\circ y$ anderenfalls.

Definition. In $H=(X; \mathfrak{P})$ heißt eine Folge der Form $(x_1, P_1, x_2, P_2, \dots, P_q, x_{q+1})$ *Weg der Länge* q, falls

1. $x_i \neq x_j$ für $i\neq j$ und $i, j=1, 2 \dots q$,

[1] Synonym wird gelegentlich die Bezeichnung „abstrakter Komplex“ gebraucht.

2. $P_i \neq P_j$ für $i \neq j$,
3. $x_k, x_{k+1} \in P_k$ für $k = 1, 2 \dots q$.

Der Weg heißt *elementar*, wenn keine Kante des Weges drei Knotenpunkte des Weges als Elemente enthält. Ist $q > 1$ und gilt $x_{q+1} = x_1$, so heißt der Weg *Kreis der Länge q*.

Definition. Sei $H = (X; \mathfrak{P})$ und $A \subset X$. Der Hypergraph $H_A = (A; \mathfrak{P}_A)$ mit $\mathfrak{P}_A = \{P_i \cap A / P_i \in \mathfrak{P}$ und $P_i \cap A \neq \emptyset\}$ heißt der durch A erzeugte *Unterhypergraph* von H.

Ist $B = X - A$, so schreiben wir auch $H - B = (X - B; \mathfrak{P} - B)$ statt $H_A = (A; \mathfrak{P}_A)$.

Definition. Sei $H = (X; \mathfrak{P})$. Eine Menge $S \subset X$ heißt *stabil*, wenn für zwei verschiedene Knotenpunkte $x, y \in S$ stets gilt $x \nmid y$. Als *Stabilitätszahl* $\alpha(H)$ bezeichnen wir die Kardinalzahl einer stabilen Menge von maximaler Ordnung, d.h.

$$\alpha(H) = \max\{\# S / S \text{ ist stabil}\}.$$

Definition. Sei $H = (X; \mathfrak{P})$. Eine Kantenmenge $\mathfrak{U} \subset \mathfrak{P}$ heißt *Überdeckung* von H, wenn $X \subset \bigcup_{P_i \in \mathfrak{U}} P_i$. Als *Überdeckungszahl* $\beta(H)$ von H bezeichnen wir die Kardinalzahl einer Überdeckung von minimaler Ordnung, d.h.

$$\beta(H) = \min\{\# \mathfrak{U} / \mathfrak{U} \text{ ist Überdeckung}\}.$$

Ist $\mathfrak{U} = \{U_1, U_2 \dots U_k\}$ eine Überdeckung von $H = (X; \mathfrak{P})$ und ist $\overline{\mathfrak{U}} = \{\overline{U}_1, \overline{U}_2 \dots \overline{U}_k\}$ mit $\overline{U}_i \subset U_i$ und $\bigcup_{i=1}^{k} \overline{U}_i = X$, so werden wir gelegentlich auch $\overline{\mathfrak{U}}$ als Überdeckung von H bezeichnen.

Als nächstes wollen wir eine Klasse von Hypergraphen charakterisieren, die sich als besonders interessant erwiesen hat.

Definition. Ein Hypergraph $H = (X; \mathfrak{P})$ heißt *gleichgewichtig*, wenn in H kein elementarer Kreis ungerader Länge existiert.

Lemma. *Ist H ein gleichgewichtiger Hypergraph, so ist auch jeder Unterhypergraph H_A gleichgewichtig.*

Beweis. Gibt es in H_A einen elementaren Kreis $\mu = (x_1, P_1 \cap A, x_2, P_2 \cap A, \dots, x_1)$, von ungerader Länge, so ist $\nu = (x_1, P_1, x_2, P_2, \dots, x_1)$ ein ebensolcher Kreis in H.

Satz 1 (Berge, Las Vergnas, Schrage). *Ist $H = (X; \mathfrak{P})$ ein gleichgewichtiger Hypergraph, so ist die Stabilitätszahl $\alpha(H)$ gleich der Überdeckungszahl $\beta(H)$.*

Beweis. $M(H)$ bezeichne stets eine maximale stabile Menge von H und $\mathfrak{A}(H)$ eine minimale Überdeckung.

Aus der Definition der Zahlen $\alpha(H)$ und $\beta(H)$ folgt unmittelbar, daß für beliebige Hypergraphen gilt

$$\alpha(H) = \# M(H) \leqq \# \mathfrak{A}(H) = \beta(H).$$

Zu zeigen bleibt also: ist H gleichgewichtig, so gilt $\alpha(H) \geqq \beta(H)$. Den Beweis führen wir durch vollständige Induktion nach der Ordnung von X.

Trivialerweise ist die Behauptung richtig für $\# X = 1$. Wir nehmen an, sie sei richtig für $\# X \leqq n$.

Sei $H_{n+1} = (X_{n+1}; \mathfrak{P}_{n+1})$ ein gleichgewichtiger Hypergraph mit $X_{n+1} = \{x_1, x_2 \ldots x_{n+1}\}$ und $H_n = (X_n; \mathfrak{P}_n)$ der von $X_n = X_{n+1} - \{x_{n+1}\}$ erzeugte Unterhypergraph.

a) Sei $\alpha(H_{n+1}) = K + 1$ und $\alpha(H_n) = K$.

Wegen des Lemmas können wir die Induktionsvoraussetzung anwenden und es gilt $\beta(H_n) = K$. Sei $P \in \mathfrak{P}_{n+1}$ mit $x_{n+1} \in P$. Dann ist $\mathfrak{A}(H_{n+1}) = \mathfrak{A}(H_n) \cup P$ eine Überdeckung von H_{n+1} der Ordnung $K+1$.

b) Sei $\alpha(H_{n+1}) = \alpha(H_n) = K$.

(i) Es gibt ein $P \in \mathfrak{P}_{n+1}$, so daß $\alpha(H_{n+1} - P) = K - 1$. Nach Induktionsvoraussetzung gilt $\beta(H_{n+1} - P) = K - 1$. $\mathfrak{A}(H_{n+1}) = \mathfrak{A}(H_n) \cup P$ hat somit die Ordnung K.

(ii) Für alle $P \in \mathfrak{P}_{n+1}$ gilt $\alpha(H_{n+1} - P) = \alpha(H_{n+1}) = K$.

Wir zeigen, daß in diesem Fall H_{n+1} nicht gleichgewichtig ist.

Unter obiger Voraussetzung ist $\beta(H_{n+1}) = \alpha(H_{n+1}) + 1 = K + 1$. Sei $V = \{x / x \circ x_{n+1}$ und $x \neq x_{n+1}\}$. Wir wählen $v \in V$ und $C_1 \in \mathfrak{P}_{n+1}$, so daß $x_{n+1}, v \in C_1$. $H_{n+1} - \{v\}$ hat nach Induktionsvoraussetzung die Überdeckungszahl K. Sei $\mathfrak{A}(H_{n+1} - \{v\}) = \{Q_1, Q_2 \ldots Q_K\}$, und es sei $x_{n+1} \in Q_K$. Wir setzen nun $C_2 = Q_K$. Es ist $\alpha(H_{n+1} - C_1) = \alpha(H_{n+1} - C_2) = K$ und $\alpha(H_{n+1} - C_1 - C_2) = K - 1$. Daraus folgt:

A 1. Es gibt ein $z_1 \in C_1$ und eine maximale stabile Knotenpunktmenge $M_1(H_{n+1})$, so daß $z_1 \in M_1(H_{n+1})$ und $C_2 \cap M_1(H_{n+1}) = \emptyset$.

A 2. Es gibt ein $z_2 \in C_2$ und eine maximale stabile Knotenpunktmenge $M_2(H_{n+1})$, so daß $z_2 \in M_2(H_{n+1})$ und $C_1 \cap M_2(H_{n+1}) = \emptyset$.

Sei $\mathfrak{A}(H_n) = \{R_1, R_2 \ldots R_K\}$. Für alle $i = 1, 2 \ldots K$ gilt $R_i \cup \{x_{n+1}\} \notin \mathfrak{P}_{n+1}$, da sonst die Voraussetzung von (i) erfüllt wäre. Wir setzen $D = C_1 \cup C_2 - \{z_1, z_2\}$ und betrachten den Hypergraphen $H_{n+1} - D = (X_{n+1} - D;\ \mathfrak{P}_{n+1} - D)$. Wegen A 1 bzw. A 2 gilt $\# M(H_{n+1} - D) = K$ und somit $\beta(H_{n+1} - D) = K$. $\mathfrak{A}(H_{n+1} - D) = \{R_1 - D, R_2 - D, \ldots, R_K - D\}$ ist also eine Überdeckung minimaler Ordnung von $H_{n+1} - D$.

Wir konstruieren nun einen Hypergraphen $\bar{H} = (\bar{X}; \bar{\mathfrak{P}})$, indem wir zur Knotenpunktmenge von $H_{n+1} - D$ zwei Punkte y_1 und y_2 hinzunehmen und die Kantenmenge um $\{y_1, z_1\}$ und $\{y_2, z_2\}$ erweitern. Es ist also

$$\bar{X} = X_{n+1} - D \cup \{y_1, y_2\} \quad \text{und} \quad \bar{\mathfrak{P}} = (\mathfrak{P}_{n+1} - D) \cup \{y_1, z_1\} \cup \{y_2, z_2\}.$$

$\mathfrak{B}(\bar{H}) = \mathfrak{A}(H_{n+1} - D) \cup \{y_1, z_1\} \cup \{y_2, z_2\}$ liefert eine (nichtminimale) Überdeckung der Ordnung $K+2$ von H. Eine minimale Überdeckung $\mathfrak{A}(\bar{H})$ hat die Ordnung $K+1$, da $\mathfrak{A}(H_{n+1} - C_1 - C_2) \cup \{y_1, z_1\} \cup \{y_2, z_2\}$ eine Überdeckung der Ordnung $K+1$ ist, und weil $\# M(\bar{H}) = K + 1$.

Sei $\mathfrak{A}(\bar{H})=\{A_1, A_2 \ldots A_{K+1}\}$ und $\mathfrak{B}(\bar{H})=\{B_1, B_2 \ldots B_{K+2}\}$ mit $B_i=R_i-D$ für $i=1,2\ldots K$, $B_{K+1}=\{y_1, z_1\}$ und $B_{K+2}=\{y_2, z_2\}$. Wir ordnen diesen beiden Überdeckungen von H den Graphen $G=G(\mathfrak{A}, \mathfrak{B})=(E; \mathfrak{K})$ zu, für den gilt

1. $E=\{a_1, a_2 \ldots a_{K+1}, b_1, b_2 \ldots b_{K+2}\}$,
2. $a_i \circ b_j \Leftrightarrow A_i \cap B_j \neq \emptyset$,
3. $a_i \not\circ a_j$ und $b_i \not\circ b_j$ für $i, j=1,2\ldots K+2$.

Die Knotenpunkte von G entsprechen also umkehrbar eindeutig den Elementen von $\mathfrak{A}(\bar{H})$ und $\mathfrak{B}(\bar{H})$, und zwei Knotenpunkte von G sind genau dann benachbart, wenn die beiden entsprechenden Kanten in $\bar{H}$ benachbart sind, und eine der Kanten $\mathfrak{A}(\bar{H})$ angehört und die andere $\mathfrak{B}(\bar{H})$.

Sowohl $\mathfrak{A}(\bar{H})$ als auch $\mathfrak{B}(\bar{H})$ enthält die Kanten $\{y_1, z_1\}$ und $\{y_2, z_2\}$. Sagen wir, es sei $A_1=B_{K+1}=\{y_1, z_1\}$ und $A_2=B_{K+2}=\{y_2, z_2\}$. In G sind dann a_1 und b_{K+1} benachbart und ebenfalls a_2 und b_{K+2}. Wir wollen nun zeigen, daß a_1 und a_2 zu einer Zusammenhangskomponente von G gehören, und damit auch a_1, a_2, b_{K+1} und b_{K+2}.

Wegen $\#\mathfrak{A}(\bar{H})<\#\mathfrak{B}(\bar{H})$ gibt es eine Zusammenhangskomponente $\tilde{G}$ von G mit den Knotenpunkten $(a_i, i\in J)$ und $(b_i, i\in L)$, derart, daß $\# J<\# L$. Wegen der Konstruktion des Graphen gilt für die entsprechenden Kanten von $\bar{H}$: $\bigcup_{i\in L} B_i \subset \bigcup_{i\in J} A_i$.

$\tilde{\mathfrak{B}}(\bar{H}-\{y_1\})=\mathfrak{B}(\bar{H})-\{y_1, z_1\}$ ist eine Überdeckung von $\bar{H}-\{y_1\}$ der Ordnung $K+1$. Wegen A1 ist aber $M_1(H_{n+1})\cup\{y_2\}$ eine stabile Menge und somit $\alpha(\bar{H}-\{y_1\})=K+1$. $\mathfrak{B}(\bar{H}-\{y_1\})$ ist also eine Überdeckung von minimaler Ordnung, d.h. $\beta(\bar{H}-\{y_1\})=K+1$. Nehmen wir nun an, a_1 und b_{K+1} gehörten nicht zu $\tilde{G}$, dann wäre $\tilde{\mathfrak{B}}(\bar{H})-\{y_1\})-(\bigcup_{i\in L} B_i)\cup(\bigcup_{i\in J} A_i)$ eine Überdeckung von $\bar{H}-\{y_1\}$ von kleinerer Ordnung als $K+1$ in Widerspruch dazu, daß $\beta(\bar{H}-\{y_1\})=K+1$. Ebenso zeigt man, daß a_2 und b_{K-2} zu G gehören.

Wir betrachten nun in G einen kürzesten Weg von a_1 nach a_2. Es mögen dabei die Knotenpunkte $a_1=a_{i1}, b_{i1}, a_{i2}, b_{i2}\ldots a_{i_r}=a_2$ in dieser Reihenfolge durchlaufen werden. Jeder dieser Punkte ist in G nur zu den Punkten dieses Weges benachbart, die ihm auch in der Folge benachbart sind. Daher entspricht diesem Weg in $\bar{H}$ eine Kantenfolge $A_1=A_{i1}, B_{i1}, A_{i2}, B_{i2}\ldots A_{i_r}=A_2$ derart, daß zwei Kanten der Folgen genau dann einen nichtleeren Durchschnitt haben, wenn sie in der Folge nebeneinander stehen. Daher gibt es in $\bar{H}$ einen elementaren Weg der Form:

$$z_1=x_1, B_{i1}, x_2, A_{i2}, x_3, B_{i2}, \ldots, B_{i(r-1)}, x_{2(r-1)}=z_2.$$

In dieser Folge kann weder B_{K+1} noch B_{K+2} auftreten. A_1 ist nämlich die einzige Kante von $\mathfrak{A}(\bar{H})$, die mit B_{K+1} einen nichtleeren Durchschnitt hat. Jedes B_{ik} der Folge ist aber zwei Elementen von $\mathfrak{A}(\bar{H})$ benachbart. Ebenso zeigt man, daß B_{K+2} nicht zu der Folge gehört.

Sei wieder $D=C_1\cup C_2-\{z_1, z_2\}$ und $\mathfrak{A}(H_n)=\{R_1, R_2\ldots R_K\}$ wie oben definiert. Betrachten wir nun einen Weg

$$\mu=(z_1=x_1, R_{i1}, x_2, P_1, x_3, R_{i2}, x_4, P_2\ldots R_{i(r-1)}, x_{2(r-1)}=z_2),$$

wobei $P_k \in \mathfrak{P}_{n+1}$ und $A_{ik} = P_k - D$. Die Kanten R_j, $j = 1, 2 \ldots K$ sind ebenfalls Elemente von $\mathfrak{P}_{n+1}$ und wegen der Konstruktion der B_j, $j = 1, 2 \ldots K$ gilt $B_{ik} = R_{ik} - D$. Daher ist μ ein elementarer Weg in H_{n+1}. Für alle R_{ik}, $k = 1, 2 \ldots r-1$ gilt $x_{n+1} \notin R_{ik}$. Falls auch für alle Kanten P_i des Weges gilt $x_{n+1} \notin P_i$, so ist $(x_{n+1}, C_1, x_1, R_{i1}, x_2, P_1, \ldots, R_{i(r-1)}, x_{2(r-1)}, C_2, x_{n+1})$ ein elementarer Kreis von ungerader Länge. Nehmen wir also an, die Bedingung $x_{n+1} \notin P_i$, $i = 1, 2 \ldots r-1$ sei nicht erfüllt und l sei der kleinste Index, so daß $x_{n+1} \in P_l$. Dann ist aber $(x_{n+1}, C_1, x_1, R_{i1}, x_2, P_1, \ldots, P_l, x_{n+1})$ ein elementarer Weg ungerader Länge.

Wir wollen noch eine naheliegende Verallgemeinerung des Satzes 1 angeben.

Definition. $H = (X; \mathfrak{P})$ heißt *im wesentlichen gleichgewichtig*, wenn eine Kantenmenge $\tilde{\mathfrak{P}} \subset \mathfrak{P}$ existiert, so daß für alle $P \in \mathfrak{P}$ ein $\tilde{P} \in \tilde{\mathfrak{P}}$ existiert mit $P \subset \tilde{P}$ und wenn $\tilde{H} = (X; \tilde{\mathfrak{P}})$ gleichgewichtig ist.

Falls H im wesentlichen gleichgewichtig ist, gilt offenbar $\alpha(H) = \alpha(\tilde{H})$ und $\beta(H) = \beta(\tilde{H})$ und wir erhalten

Satz 2. *Ist $H = (X; \mathfrak{P})$ im wesentlichen gleichgewichtig, so gilt $\alpha(H) = \beta(H)$.*

Folgende bekannte Sätze ergeben sich in einfacher Weise als Spezialfälle aus Satz 1 bzw. Satz 2.

Satz 3. *Sei G ein paarer Graph. Dann ist die maximale Zahl nichtbenachbarter Knotenpunkte gleich der Kardinalzahl einer G überdeckenden Kantenmenge von minimaler Ordnung.*

Satz 4 (König [3]). *Sei G ein paarer Graph. Dann ist die maximale Zahl nichtbenachbarter Kanten gleich der Kardinalzahl einer Knotenpunktmenge minimaler Ordnung, die mit jeder Kante einen nichtleeren Durchschnitt hat.*

Satz 5 (van der Waerden [5]). *In einer Menge X der Ordnung $n \cdot k$ seien zwei Partitionen $\mathfrak{K} = \{K_1 \ldots K_n\}$ und $\mathfrak{L} = \{L_1 \ldots L_n\}$ gegeben mit $\# K_i = \# L_i = k$ für $i = 1, 2 \ldots n$. Dann gibt es eine Menge $Y \subset X$, die zugleich Repräsentantenmenge für $\mathfrak{K}$ und für $\mathfrak{L}$ ist, d.h. für die gilt $\#(Y \cap K_i) = \#(Y \cap L_i) = 1$ für $i = 1, 2 \ldots n$.*

Literatur

1. Berge, C.: Graphes et hypergraphes. Paris: Dunod 1970.
2. – Las Vergnas, M.: Sur un théorèm du type König pour Hypergraphes. Int. Conf. on Combinatorial Mathematics. Annals New York Ac. of Sc. **175**, 32–40 (1970).
3. König, D.: Theorie der endlichen und unendlichen Graphen. Leipzig: Akad. Verlagsgesellschaft 1936.
4. Schrage, G.: Der Zusammenhang zwischen der Ordnung maximaler stabiler Punktmengen und minimaler Überdeckungen bei paaren Komplexen. Tagungsbericht „Combinatorial Aspects of Finite Geometries", Mathematisches Forschungsinstitut Oberwolfach, 1970, S. 10.
5. Waerden, B. L. van der: Ein Satz über Klasseneinteilung von endlichen Mengen. Abh. math. Sem. Univ. Hamburg **5**, 185–188 (1927).

Dr. G. Schrage
D-5930 Hüttental-Setzen, Sonnenstraße 34
Deutschland

(Eingegangen am 5. August 1971)

Math. Z. 124, 191–198 (1972)

Tensorproduktmethoden bei mehrdimensionaler Interpolation

WERNER HAUSSMANN

In [5] werden Existenz-, Eindeutigkeits- und Darstellungsfragen bei Tensorprodukten von Interpolationsproblemen behandelt. Als Anwendungen der dort entwickelten Theorie ergeben sich Aussagen über mehrdimensionale Hermite-Interpolation für Polynome aus

$$\Pi_{MN} = \left\{ P\colon\ P(x,y) = \sum_{i=0}^{M} \sum_{j=0}^{N} a_{ij}\, x^i y^j,\ a_{ij} \in \mathbb{R} \right\}, \tag{1}$$

wenn zur Interpolation Knotengitter der Gestalt

$$W = \{x_0, \ldots, x_m\} \times \{y_0, \ldots, y_n\} \tag{2}$$

verwendet werden. Die Interpolationsbedingungen, die dabei an jedem Punkt von W vorgeschrieben werden, ergeben sich aus den Daten zweier eindimensionaler Hermite-Interpolationsprobleme.

Im Anschluß an diese Untersuchungen stellen sich für die mehrdimensionale Hermite-Interpolation zwei Fragen:

(A) Lassen sich die Interpolationsbedingungen – also partielle Ableitungen, die an den Knoten auszuwerten sind – „unregelmäßiger" auf die Interpolationsknoten $(x_\mu, y_\nu) \in W$ verteilen als beim Tensorprodukt zweier Hermite-Interpolationsprobleme, ohne daß Existenz- und Eindeutigkeitsaussagen verlorengehen?

(B) Kann von der regelmäßigen Verteilung der Knoten (2) abgegangen und dennoch Existenz und Eindeutigkeit eines Interpolationspolynoms aus Π_{MN} zu gegebenen Interpolationsdaten gesichert werden?

Beide Fragen lassen sich positiv beantworten. Eine „zu unregelmäßige" Verteilung der Ableitungsbedingungen und der Interpolationsknoten darf jedoch nicht vorliegen. Auf einer Geraden $G_c = \{(x,y)\colon y=c,\ x\in\mathbb{R}\} \subset \mathbb{R}^2$ können nämlich nicht mehr als $M+1$ von insgesamt $(M+1)\cdot(N+1)$ Interpolationsbedingungen für Polynome aus Π_{MN} vorgeschrieben werden. Aber auch dann, wenn höchstens $M+1$ Bedingungen an Punkten einer Geraden G_c $(c\in\mathbb{R})$ vorgeschrieben werden, und eine entsprechende Einschränkung für jede Gerade $H_c = \{(x,y)\colon x=c,\ y\in\mathbb{R}\}$ vorausgesetzt wird, sind Existenz und Eindeutigkeit bei unregelmäßiger Verteilung der Interpolationsknoten nicht gesichert. Soll nämlich beispielsweise ein Lagrange-Interpolationspolynom aus $\Pi_{1,1}$ durch Angabe von Interpolationsdaten an den Punkten $(-1,0)$, $(0,-1)$, $(1,1)$ sowie

$(u, v) \in \mathbb{R}^2$ bestimmt werden, so ist dieses genau dann eindeutig bestimmt, wenn u und v nicht die Relation $v = \frac{u+1}{3u-1}$ erfüllen.

Daraus ergibt sich, daß im Fall der Fragen (A) und (B) nicht „zu unregelmäßige" Forderungen gestellt werden können. Das Ziel der vorliegenden Untersuchungen ist es, einen Schritt zur Beantwortung der Fragen (A) und (B) zu tun, und zwar nicht speziell für die Hermite-Interpolation, sondern im Rahmen einer allgemeinen Theorie, bei der direkte Summen von Tensorprodukten linearer Räume benutzt werden. Dementsprechend führen wir *direkte Summen gewisser Interpolationsprobleme* bezüglich eines weiteren Interpolationsproblems ein und zeigen, daß sich auf diese Weise aus eindeutig lösbaren Interpolationsproblemen neue solche Probleme ergeben. Als Spezialfall erhalten wir das Tensorprodukt zweier Interpolationsprobleme. Bei der Anwendung auf mehrdimensionale Hermite-Interpolation ergibt sich eine unregelmäßigere Verteilung der Interpolationsknoten und der Interpolationsbedingungen als in [5]. Dies ist insbesondere für Anwendungen von Bedeutung.

In [2] betrachten Ehlich und Haußmann Konvergenz von Interpolationsprozessen in $C[a, b] \otimes \bar{C}[\bar{a}, \bar{b}]$. Wir führen diese Untersuchungen weiter und erhalten Ergebnisse über Konvergenz von Tensorprodukten gewisser Interpolationsprozesse in Tensorprodukten beliebiger linearer normierter Räume.

Unsere Aussagen, die wir für den zweidimensionalen Fall beweisen, lassen sich unschwer auf höhere Dimensionen verallgemeinern.

1. Direkte Summen von Interpolationsproblemen bezüglich eines Interpolationsproblems

Es sei X ein linearer Raum über $\mathbb{R}$ und F ein $(m+1)$-dimensionaler Unterraum von X. $\varphi_0, \ldots, \varphi_m$ seien lineare Funktionale auf F und $\Phi_0, \ldots, \Phi_m \in X^*$ (algebraischer Dualraum von X) Erweiterungen der $\varphi_0, \ldots, \varphi_m$. Diese Daten legen das Interpolationsproblem $\mathscr{P} = (X, F; \Phi_0, \ldots, \Phi_m)$ fest. Nach Davis [1] gilt:

Genau dann gibt es zu jedem $x \in X$ ein eindeutig bestimmtes $u \in F$ mit

$$\varphi_\mu(u) = \Phi_\mu(x) \qquad 0 \leqq \mu \leqq m,$$

wenn $\{\varphi_0, \ldots, \varphi_m\}$ eine Basis von F^* ist.

Wir nennen dann $\mathscr{P}$ eindeutig lösbar. Da bei Existenz-, Eindeutigkeits- und Darstellungsfragen der Grundraum X nicht bekannt zu sein braucht, schreiben wir bisweilen abgekürzt $(F, \varphi_0, \ldots, \varphi_m)$ statt $(X, F; \Phi_0, \ldots, \Phi_m)$ (vgl. [5]).

Wie in [5] benutzen wir das Symbol $[z_0, \ldots, z_k]$ für die lineare Hülle von Elementen $z_0, \ldots, z_k$ eines linearen Raumes Z über $\mathbb{R}$.

Gegeben sei jetzt ein $(r+1)$-dimensionaler linearer Raum F über $\mathbb{R}$ sowie die Funktionale $\varphi_{\sigma\rho} \in F^*$ $(0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma$ mit nichtnegativen ganzen Zahlen $r_\sigma \leqq r)$. Wir bilden die Interpolationsprobleme $\mathscr{J}_\sigma = (F_\sigma, \tilde{\varphi}_{\sigma 0}, \ldots, \tilde{\varphi}_{\sigma r_\sigma})$ mit Teilräumen $F_\sigma \subset F$, $\dim F_\sigma = r_\sigma + 1$, und den Einschränkungen $\tilde{\varphi}_{\sigma\rho} = \varphi_{\sigma\rho} | F_\sigma$ $(0 \leqq \sigma \leqq s,$

$0 \leqq \rho \leqq r_\sigma$). Weiter sei $\mathcal{J}=(G, \psi_0, \ldots, \psi_s)$ ein eindeutig lösbares Interpolationsproblem. Dann existiert die Dualbasis $\{g_0, \ldots, g_s\}$ von $\{\psi_0, \ldots, \psi_s\}$. Schließlich sei $G_\sigma = [g_\sigma]$ $(0 \leqq \sigma \leqq s)$.

Definition 1. Unter der *direkten Summe der Interpolationsprobleme* $\mathcal{J}_\sigma$ *bezüglich* $\mathcal{J}$ verstehen wir das Problem

$$\mathcal{L}=\left(H, (\varphi_{\sigma\rho} \otimes \psi_\sigma) | H\colon 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma\right) \quad \text{mit} \quad H = \bigoplus_{\sigma=0}^{s} F_\sigma \otimes G_\sigma .$$

Wir erhalten damit den folgenden Existenz- und Eindeutigkeitssatz:

Satz 1. *Das Interpolationsproblem* $\mathcal{L}$ *ist bei eindeutig lösbarem* $\mathcal{J}$ *genau dann eindeutig lösbar, wenn die Interpolationsprobleme* $\mathcal{J}_\sigma$ *eindeutig lösbar sind* $(0 \leqq \sigma \leqq s)$.

Beweis. Die Probleme $\mathcal{J}_\sigma = (F_\sigma, \tilde{\varphi}_{\sigma 0}, \ldots, \tilde{\varphi}_{\sigma r_\sigma})$ seien sämtlich eindeutig lösbar $(0 \leqq \sigma \leqq s)$. Dann sind nach Davis [1] die linearen Funktionale $\tilde{\varphi}_{\sigma 0}, \ldots, \tilde{\varphi}_{\sigma r_\sigma} \in F_\sigma^*$ linear unabhängig $(0 \leqq \sigma \leqq s)$. $\{f_{\sigma 0}, \ldots, f_{\sigma r_\sigma}\} \subset F_\sigma$ seien die Dualbasen von $\{\tilde{\varphi}_{\sigma 0}, \ldots, \tilde{\varphi}_{\sigma r_\sigma}\}$ für $0 \leqq \sigma \leqq s$. Nun gilt: $f_{\sigma\rho} \otimes g_\sigma \in H$ $(0 \leqq \sigma \leqq s,\ 0 \leqq \rho \leqq r_\sigma)$ und es ist (vgl. Greub [4])

$$\dim H^* = \dim H = \sum_{\sigma=0}^{s} (r_\sigma + 1).$$

Wir zeigen jetzt die lineare Unabhängigkeit der Funktionale $\varphi_{\sigma\rho} \otimes \psi_\sigma$ $(0 \leqq \sigma \leqq s$, $0 \leqq \rho \leqq r_\sigma)$ auf H. Es sei dazu mit $\beta_{\sigma\rho} \in \mathbb{R}$

$$\omega = \sum_{\sigma=0}^{s} \sum_{\rho=0}^{r_\sigma} \beta_{\sigma\rho} \cdot \varphi_{\sigma\rho} \otimes \psi_\sigma .$$

Gilt $\omega | H = 0$, so verschwindet ω insbesondere auf den Elementen der Menge $\{f_{\sigma\rho} \otimes g_\sigma\colon 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma\}$. Also erhalten wir für $l \in \{0, \ldots, s\}$ und $k \in \{0, \ldots, r_l\}$:

$$\begin{aligned} 0 = \omega(f_{lk} \otimes g_l) &= \sum_{\sigma=0}^{s} \sum_{\rho=0}^{r_\sigma} \beta_{\sigma\rho} \cdot \varphi_{\sigma\rho}(f_{lk}) \cdot \psi_\sigma(g_l) \\ &= \sum_{\sigma=0}^{s} \sum_{\rho=0}^{r_\sigma} \beta_{\sigma\rho} \cdot \delta_{\rho k} \cdot \delta_{\sigma l} = \beta_{lk} . \end{aligned}$$

Damit ist die lineare Unabhängigkeit gezeigt. Aus Anzahlgründen ist $\{(\varphi_{\sigma\rho} \otimes \psi_\sigma) | H\colon 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma\}$ eine Basis von H^*. Damit ist das Interpolationsproblem

$$\mathcal{L} = \left(\bigoplus_{\sigma=0}^{s} F_\sigma \otimes G_\sigma, (\varphi_{\sigma\rho} \otimes \psi_\sigma) | H\colon 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma \right)$$

eindeutig lösbar.

Es sei jetzt umgekehrt $\mathcal{L}$ eindeutig lösbar. Wir nehmen an, das Problem $(F_l, \tilde{\varphi}_{l0}, \ldots, \tilde{\varphi}_{l r_l})$ sei nicht eindeutig lösbar für ein $l \in \{0, \ldots, s\}$. Dann gilt nach [5, Lemma 1 und 2]:

$$\dim [(\varphi_{l\rho} \otimes \psi_l) | H\colon 0 \leqq \rho \leqq r_l] = \dim [\tilde{\varphi}_{l0}, \ldots, \tilde{\varphi}_{l r_l}] = r'_l < r_l + 1.$$

Für $\sigma \neq l$ gilt weiter

$$\dim[(\varphi_{\sigma\rho} \otimes \psi_\sigma) | H: 0 \leqq \rho \leqq r_\sigma] = \dim[\tilde{\varphi}_{\sigma 0}, \ldots, \tilde{\varphi}_{\sigma r_\sigma}] = r'_\sigma \leqq r_\sigma + 1.$$

Damit ergibt sich aber

$$\dim[(\varphi_{\sigma\rho} \otimes \psi_\sigma) | H: 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma] = \sum_{\sigma=0}^{s} r'_\sigma < \sum_{\sigma=0}^{s} (r_\sigma + 1).$$

Also bilden die $(\varphi_{\sigma\rho} \otimes \psi_\sigma) | H$ keine Basis von H^*, d.h. das Problem $\mathscr{L} = (H, (\varphi_{\sigma\rho} \otimes \psi_\sigma) | H: 0 \leqq \sigma \leqq s, 0 \leqq \rho \leqq r_\sigma)$ ist nicht eindeutig lösbar – im Gegensatz zur Voraussetzung.

Durch die Spezialisierung $\mathscr{J}_\sigma = (F, \varphi_0, \ldots, \varphi_r)$ für alle $0 \leqq \sigma \leqq s$ erhalten wir die Existenz- und Eindeutigkeitsaussage für Tensorprodukte von Interpolationsproblemen (vgl. [5, Satz 1]).

2. Verallgemeinerte Hermite-Interpolation im $\mathbb{R}^2$

Neben dem zweidimensionalen Polynomraum Π_{MN} aus (1) benötigen wir die eindimensionalen Polynomräume

$$\Pi_M = [1, x, \ldots, x^M] \quad \text{und} \quad \overline{\Pi}_N = [1, y, \ldots, y^N].$$

Weiter sei $\{y_0, \ldots, y_n\}$ eine Menge von paarweise verschiedenen reellen Zahlen. Für jedes $\nu \in \{0, \ldots, n\}$ sei eine nichtnegative ganze Zahl N_ν gegeben, und es sei

$$N := \sum_{\nu=0}^{n} (N_\nu + 1) - 1.$$

Mit Hilfe der Funktionale $\psi_\nu^\omega \in \overline{\Pi}_N^*$, die durch

$$\psi_\nu^\omega: \overline{\Pi}_N \ni P \mapsto P^{(\omega)}(y_\nu) = \frac{d^\omega}{dy^\omega} P(y_\nu) \in \mathbb{R}$$

für $0 \leqq \nu \leqq n$, $0 \leqq \omega \leqq N_\nu$ festgelegt sind, bilden wir das Interpolationsproblem

$$\mathscr{H} = (\overline{\Pi}_N, \psi_\nu^\omega: 0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu).$$

Zu $\mathscr{H}$ seien die folgenden Mengen

$$X_\nu^\omega = \{x_{\nu\mu}^\omega \in \mathbb{R}: 0 \leqq \mu \leqq m_\nu^\omega\} \qquad (0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu)$$

mit $x_{\nu\mu}^\omega \neq x_{\nu\rho}^\omega$ $(\mu \neq \rho)$ und nichtnegativen Zahlen $m_\nu^\omega \in \mathbb{Z}$ gegeben. Weiter sei $X_\nu = \bigcup_{\omega=0}^{N_\nu} X_\nu^\omega$. Wir bilden das Knotengitter

$$V = \bigcup_{\nu=0}^{n} X_\nu \times \{y_\nu\}.$$

Zu jedem $x_{\nu\mu}^\omega \in X_\nu^\omega$ sei eine nichtnegative ganze Zahl $M_{\nu\mu}^\omega$ $(0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu, 0 \leqq \mu \leqq m_\nu^\omega)$ gegeben, und es gelte:

$$\sum_{\mu=0}^{m} (M_{\nu\mu}^\omega + 1) - 1 =: M \qquad (\text{unabhängig von } \nu, \omega).$$

Wir definieren die linearen Funktionale $\varphi_{\nu\mu}^{\omega\tau} \in \Pi_M^*$:

$$\varphi_{\nu\mu}^{\omega\tau}: \Pi_M \ni P \mapsto P^{(\tau)}(x_{\nu\mu}^{\omega}) = \frac{d^\tau}{dx^\tau} P(x_{\nu\mu}^{\omega}) \in \mathbb{R}$$

für $0 \leqq \mu \leqq m_\nu^\omega$, $0 \leqq \tau \leqq M_{\nu\mu}^\omega$ sowie $0 \leqq \nu \leqq n$ und $0 \leqq \omega \leqq N_\nu$. Auf diese Weise erhalten wir die N Interpolationsprobleme

$$\mathscr{H}_\nu^\omega = (\Pi_M, \varphi_{\nu\mu}^{\omega\tau}: 0 \leqq \mu \leqq m_\nu^\omega, 0 \leqq \tau \leqq M_{\nu\mu}^\omega).$$

Aus den Problemen $\mathscr{H}$ und $\mathscr{H}_\nu^\omega$ $(0 \leqq \nu \leqq n,\ 0 \leqq \omega \leqq N_\nu)$ läßt sich jetzt ein zweidimensionales Interpolationsproblem bilden. Wie in [5] verwenden wir das Symbol

$$D^{(\tau,\omega)} Q = \frac{\partial^{\tau+\omega}}{\partial x^\tau \partial y^\omega} Q$$

für partielle Ableitungen eines Polynoms $Q \in \Pi_{MN}$; alle diese Ableitungen sind von der Differentiationsreihenfolge unabhängig. Für jeden Punkt

$$(x_{\nu\mu}^\omega, y_\nu) \in X_\nu^\omega \times \{y_\nu\}$$

definieren wir die linearen Funktionale $\Lambda_{\nu\mu}^{\omega\tau} \in \Pi_{MN}^*$:

$$\Lambda_{\nu\mu}^{\omega\tau}: \Pi_{MN} \ni Q \mapsto D^{(\tau,\omega)} Q(x_{\nu\mu}^\omega, y_\nu) \in \mathbb{R}$$

für $0 \leqq \nu \leqq n$, $0 \leqq \omega \leqq N_\nu$, $0 \leqq \mu \leqq m_\nu^\omega$, $0 \leqq \tau \leqq M_{\nu\mu}^\omega$. Wir erhalten damit

Definition 2. Unter einem *verallgemeinerten Hermite-Interpolationsproblem* in zwei Veränderlichen verstehen wir das Problem

$$\mathscr{K} = (\Pi_{MN}, \Lambda_{\nu\mu}^{\omega\tau}: 0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu, 0 \leqq \mu \leqq m_\nu^\omega, 0 \leqq \tau \leqq M_{\nu\mu}^\omega).$$

Numerieren wir die Interpolationsprobleme

$$\mathscr{H}_0^0, \ldots, \mathscr{H}_0^{N_0}, \mathscr{H}_1^0, \ldots, \mathscr{H}_1^{N_1}, \ldots, \mathscr{H}_n^0, \ldots, \mathscr{H}_n^{N_n} \tag{3}$$

durchlaufend von $0, 1, \ldots, N$, und bezeichnen wir mit $\mathscr{P}_\sigma$ das σ-te dieser Probleme $(0 \leqq \sigma \leqq N)$, so ergibt sich

Satz 2. *Das zweidimensionale verallgemeinerte Hermite-Interpolationsproblem $\mathscr{K}$ ist die direkte Summe der Interpolationsprobleme $\mathscr{P}_\sigma$ $(0 \leqq \sigma \leqq N)$ bezüglich $\mathscr{H}$.*

Beweis. Es sei $\{g_\nu^\omega: 0 \leqq \nu \leqq n,\ 0 \leqq \omega \leqq N_\nu\} = \{g_\sigma: 0 \leqq \sigma \leqq N\}$ (gemäß der Numerierung (3)) die Dualbasis von $\{\psi_\nu^\omega: 0 \leqq \nu \leqq n,\ 0 \leqq \omega \leqq N_\nu\}$. Dann gilt mit $[g_\nu^\omega] = G_\nu^\omega$ bzw. $[g_\sigma] = G_\sigma$:

$$\bigoplus_{\sigma=0}^{N} \Pi_M \otimes G_\sigma = \bigoplus_{\nu=0}^{n} \left(\bigoplus_{\omega=0}^{N_\nu} \Pi_M \otimes G_\nu^\omega \right) = \bigoplus_{\nu=0}^{n} \Pi_M \otimes \hat{G}_\nu,$$

wo $\hat{G}_\nu = \bigoplus_{\omega=0}^{N_\nu} G_\nu^\omega$ gesetzt wird, und damit weiter

$$= \Pi_M \otimes \left(\bigoplus_{\nu=0}^{n} \hat{G}_\nu \right) = \Pi_M \otimes \bar{\Pi}_N = \Pi_{MN}.$$

Wie in [5] ergibt sich die Beziehung

$$\Lambda_{\nu\mu}^{\omega\tau} = \varphi_{\nu\mu}^{\omega\tau} \otimes \psi_{\nu}^{\omega}.$$

Damit folgt die Behauptung.

Aus Satz 1 und der eindeutigen Lösbarkeit der Interpolationsprobleme $\mathscr{H}_\nu^\omega$ $(0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu)$ und $\mathscr{H}$ ergibt sich mit dem obigen Ergebnis die

Folgerung. *Jedes zweidimensionale verallgemeinerte Hermite-Interpolationsproblem ist eindeutig lösbar.*

Wir weisen auf einige Spezialfälle hin:

(i) Wählen wir $m_\nu^\omega = m$ (für alle $0 \leqq \nu \leqq n$, $0 \leqq \omega \leqq N_\nu$), so erhalten wir als Knotenmenge ein $(m+1, n+1)$-*Gitter* im $\mathbb{R}^2$ (vgl. Ehlich und Haußmann [3]). Für $M=m$ und $N=n$ ergibt sich aus der Folgerung die Eindeutigkeit der Lagrange-Interpolation auf einem $(m+1, n+1)$-Gitter.

(ii) Ist $m_\nu^\omega = m$ $(0 \leqq \nu \leqq n$, $0 \leqq \omega \leqq N_\nu)$, und gilt für jedes $\mu \in \{0, \ldots, m\}$ $x_{\nu\mu}^\omega = x_\mu$ (unabhängig von $\nu \in \{0, \ldots, n\}$, $\omega \in \{0, \ldots, N_\nu\}$), so erhalten wir das *Rechteckgitter* W (vgl. (2)) als Menge der Interpolationsknoten im $\mathbb{R}^2$. Es ist dann $X_\nu = X_\nu^\omega = \{x_0, \ldots, x_m\}$ für alle $0 \leqq \nu \leqq n, 0 \leqq \omega \leqq N_\nu$. Mit Hilfe der Interpolationsfunktionale $\varphi_{\nu\mu}^{\omega\tau} \otimes \psi_\nu^\omega = \Lambda_{\nu\mu}^{\omega\tau}$ schreiben wir am Punkt $(x_\mu, y_\nu) \in W$ für ein Polynom $Q \in \Pi_{MN}$ die folgenden partiellen Ableitungen vor:

$$\begin{array}{cccc}
D^{(0,0)}Q, & D^{(0,1)}Q, & \ldots, & D^{(0,N_\nu)}Q \\
D^{(1,0)}Q, & D^{(1,1)}Q, & \ldots, & D^{(1,N_\nu)}Q \\
\vdots & \vdots & & \vdots \\
D^{(M_{\nu\mu}^0, 0)}Q, & D^{(M_{\nu\mu}^1, 1)}Q, & \ldots, & D^{(M_{\nu\mu}^{N_\nu}, N_\nu)}Q.
\end{array}$$

In diesem Fall kann also eine wesentlich allgemeinere Verteilung der Interpolationsbedingungen vorliegen als beim

(iii) Tensorprodukt zweier Hermite-Interpolationsprobleme (vgl. [5]). Diesen Fall erhalten wir für $m_\nu^\omega = m$ sowie $x_{\nu\mu}^\omega = x_\mu$, $M_{\nu\mu}^\omega = M_\mu$ und $\varphi_{\nu\mu}^{\omega\tau} = \varphi_\mu^\tau$ (jeweils unabhängig von $\nu \in \{0, \ldots, n\}$ und $\omega \in \{0, \ldots, N_\nu\}$).

3. Konvergenz bei Tensorprodukten von Interpolationsproblemen

Es seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ zwei reelle normierte lineare Räume sowie $(m_r)_{r \geqq 0}$ und $(n_r)_{r \geqq 0}$ zwei Folgen nichtnegativer ganzer Zahlen. $(F_r)_{r \geqq 0}$ und $(G_r)_{r \geqq 0}$ seien zwei Folgen von Unterräumen von X bzw. Y mit

$$\dim F_r = m_r + 1 \quad \text{bzw.} \quad \dim G_r = n_r + 1 \qquad (r \geqq 0).$$

Durch die linearen Funktionale $\Phi_0^{(r)}, \ldots, \Phi_{m_r}^{(r)} \in X^*$ und $\Psi_0^{(r)}, \ldots, \Psi_{n_r}^{(r)} \in Y^*$ $(r = 0, 1, 2, \ldots)$ seien zwei Folgen von eindeutig lösbaren Interpolationsproblemen

$$\mathscr{K}_r = (X, F_r; \Phi_0^{(r)}, \ldots, \Phi_{m_r}^{(r)}), \qquad \mathscr{L}_r = (Y, G_r; \Psi_0^{(r)}, \ldots, \Psi_{n_r}^{(r)})$$

gegeben, die jeweils eine Folge von Interpolationsoperatoren

$$K_r \colon X \to X, \quad L_r \colon Y \to Y \quad (r \geqq 0)$$

definieren. Sind $\varphi_\mu^{(r)} = \Phi_\mu^{(r)} | F_r$ und $\psi_\nu^{(r)} = \Psi_\nu^{(r)} | G_r$ $(0 \leqq \mu \leqq m_r,\ 0 \leqq \nu \leqq n_r,\ r \geqq 0)$ die Einschränkungen der betreffenden Funktionale auf F_r bzw. G_r, und sind $\{f_0^{(r)}, \ldots, f_{m_r}^{(r)}\}$ und $\{g_0^{(r)}, \ldots, g_{n_r}^{(r)}\}$ die Dualbasen von $\{\varphi_0^{(r)}, \ldots, \varphi_{m_r}^{(r)}\}$ bzw. $\{\psi_0^{(r)}, \ldots, \psi_{n_r}^{(r)}\}$, so gilt: Die Operatoren K_r bzw. L_r lassen sich in der Form

$$K_r\colon X \ni x \mapsto \sum_{\mu=0}^{m_r} \Phi_\mu^{(r)}(x) \cdot f_\mu^{(r)} \in F_r,$$

$$L_r\colon Y \ni y \mapsto \sum_{\nu=0}^{n_r} \Psi_\nu^{(r)}(y) \cdot g_\nu^{(r)} \in G_r$$

darstellen. Durch das Tensorprodukt

$$\mathcal{K}_r \otimes \mathcal{L}_r := (X \otimes Y, F_r \otimes G_r;\ \Phi_\mu^{(r)} \otimes \Psi_\nu^{(r)}\colon 0 \leqq \mu \leqq m_r,\ 0 \leqq \nu \leqq n_r),$$

bei dem $X \otimes Y$ durch eine Tensorproduktnorm $\|.\|_{X \otimes Y}$ (cross-norm, vgl. Schatten [7] sowie Robertson und Robertson [6]) normiert ist, wird eine Folge T_r von Interpolationsoperatoren gegeben:

$$T_r\colon X \otimes Y \ni z \mapsto \sum_{\mu=0}^{m_r} \sum_{\nu=0}^{n_r} \Phi_\mu^{(r)} \otimes \Psi_\nu^{(r)}(z) \cdot f_\mu^{(r)} \otimes g_\nu^{(r)} \in F_r \otimes G_r.$$

Damit erhalten wir

Satz 3. *Gegeben seien die Folgen $\mathcal{K}_r$ und $\mathcal{L}_r$ von eindeutig lösbaren Interpolationsproblemen mit den Interpolationsoperatoren K_r bzw. L_r. $X_m \subset X$ und $Y_n \subset Y$ seien m- bzw. n-dimensionale Teilräume von X bzw. Y $(m, n \in \mathbb{N})$. Gilt dann*

$$\lim_{r \to \infty} \|x - K_r x\|_X = 0 \quad \textit{für alle } x \in X_m,$$

$$\lim_{r \to \infty} \|y - L_r y\|_Y = 0 \quad \textit{für alle } y \in Y_n,$$

so folgt

$$\lim_{r \to \infty} \|z - T_r z\|_{X \otimes Y} = 0 \quad \textit{für alle } z \in X_m \otimes Y_n.$$

Beweis. Der zu $\mathcal{K}_r \otimes \mathcal{L}_r$ gehörige Interpolationsoperator ist

$$T_r = K_r \otimes L_r \qquad (r \geqq 0).$$

Jedes $z \in X_m \otimes Y_n$ besitzt die Darstellung $z = \sum_{i=1}^{R} x_i \otimes y_i$ $(x_i \in X_m,\ y_i \in Y_n)$. Aufgrund der Linearität von $K_r \otimes L_r$ gilt:

$$T_r(z) = T_r\left(\sum_{i=1}^{R} x_i \otimes y_i\right) = K_r \otimes L_r\left(\sum_{i=1}^{R} x_i \otimes y_i\right) = \sum_{i=1}^{R} K_r(x_i) \otimes L_r(y_i).$$

Hieraus ergibt sich

$$\begin{aligned}\|z - T_r z\|_{X \otimes Y} &= \left\| \sum_{i=1}^{R} (x_i \otimes y_i - K_r(x_i) \otimes L_r(y_i)) \right\|_{X \otimes Y} \\ &\leqq \sum_{i=1}^{R} \left\{ \|x_i \otimes (y_i - L_r(y_i))\|_{X \otimes Y} + \|(x_i - K_r(x_i)) \otimes L_r(y_i)\|_{X \otimes Y} \right\}.\end{aligned}$$

Da $\|\cdot\|_{X\otimes Y}$ eine Tensorproduktnorm ist, folgt weiter:

$$\leqq \sum_{i=1}^{R} \{\|x_i\|_X \cdot \|y_i - L_r(y_i)\|_Y + \|x_i - K_r(x_i)\|_X \cdot \|L_r(y_i)\|_Y\}.$$

Nach Voraussetzung gilt für $1 \leqq i \leqq R$

$$\lim_{r\to\infty} \|x_i - K_r(x_i)\|_X = \lim_{r\to\infty} \|y_i - L_r(y_i)\|_Y = 0,$$

und damit

$$\|L_r(y_i)\|_Y \leqq \|y_i\|_Y + 1 \qquad \text{für } r \geqq r_0$$

mit geeignetem $r_0 \in \mathbb{N}$. Damit ergibt sich aber

$$\lim_{r\to\infty} \|z - T_r z\|_{X\otimes Y} = \lim_{r\to\infty} \left\| \sum_{i=1}^{R} x_i \otimes y_i - T_r \left(\sum_{i=1}^{R} x_i \otimes y_i \right) \right\|_{X\otimes Y} = 0.$$

Dies war zu zeigen.

Bemerkung. Setzen wir anstelle der Konvergenz von K_r bzw. L_r auf X_m bzw. Y_n voraus, daß die Interpolationsoperatoren von einer *gewissen Ordnung* konvergieren (vgl. Ehlich und Haußmann [2]), so konvergiert T_r auf $X_m \otimes Y_n$ von derselben Ordnung. Diese Aussage läßt sich entsprechend wie Satz 3 beweisen.

Literatur

1. Davis, P.J.: Interpolation and approximation. New York-Toronto-London: Blaisdell 1965.
2. Ehlich, H., Haußmann, W.: Konvergenz mehrdimensionaler Interpolation. Numerische Math. **15**, 165–174 (1970).
3. Ehlich, H., Haußmann, W.: Čebyšev-Approximation stetiger Funktionen in zwei Veränderlichen. Math. Z. **117**, 21–34 (1970).
4. Greub, W.H.: Multilinear algebra. Berlin-Heidelberg-New York: Springer 1967.
5. Haußmann, W.: Tensorprodukte und mehrdimensionale Interpolation. Math. Z. **113**, 17–23 (1970).
6. Robertson, A.P., Robertson, W.J.: Topologische Vektorräume. Mannheim: Bibliographisches Institut 1967.
7. Schatten, R.: A theory of cross-spaces. Annals of mathematics studies 26. Princeton: University Press 1950.

Dozent Dr. Werner Haußmann
Institut für Mathematik
der Ruhr-Universität
D-4630 Bochum
Buscheystraße Geb. NA
Deutschland

(Eingegangen am 5. Juli 1971)

Math. Z. 124, 199–202 (1972)

On a Class of Normaloid Operators

VASILE I. ISTRĂŢESCU

1. Let H be a Hilbert space and T be a bounded linear operator on H. We use the following notations and terminology: $\sigma(T)$ for the spectrum of T, $W(T)$ for the numerical range of T, and $\mathrm{cl}\, W(T)$ for its closure, $w(T)$ for the numerical radius, i.e.,

$$w(T) = \sup\{|\lambda|, \lambda \in W(T)\}$$

$\|T\|$ for the usual norm of T.

An operator T is called normaloid [6] if $\|T\| = w(T)$ and T is called transaloid if for all complex numbers $T_\lambda = T - \lambda I$ is normaloid.

The purpose of the present Note is to record some results concerning the class of transaloid operators.

2. In [1] Berberian raised the following question: If $T = UR$ (the standard polar decomposition) is an invertible operator such that U is cramped does it follows that $0 \notin \mathrm{cl}\, W(T)$? A unitary operator U is cramped if

$$\sigma(U) \subset \{e^{i\theta}, \ \theta_0 < \theta < \theta_0 + \pi\}.$$

We need also the notion of bare point. Let F be a compact set in the complex plane and $\lambda \in F$. The point λ is called bare point if there is a circle through such that no points of F lie outside this circle.

Our result is

Theorem 1. *If $T = UR$ is an invertible transaloid operator such that U is cramped then $0 \notin \mathrm{cl}\, W(T)$.*

Proof. For the proof we use the properties of bare points given in [5]. Suppose now that $0 \in \mathrm{cl}\, W(T)$ and thus 0 is in the convex hull of $\sigma(T)$. We find a sequence $\{s_n\}$ of points in the convex hull of $\sigma(T)$ and even more, in the convex hull of bare points of $\mathrm{cl}\, W(T)$,

$$1. \quad s_n = \sum_{i=1}^{n} t_i \alpha_i, \quad s_n \to 0,$$

$$2. \quad \alpha_i \text{ are bare points}, \ \sum_{i=1}^{n} t_i = 1.$$

It is easy to see that every bare point of $\mathrm{cl}\, W(T)$ is in $\sigma(T)$ and if B_p denotes the set all bare points of $\mathrm{cl}\, W(T)$, our hypothesis implies

$$\inf_{\alpha \in \mathbf{B}_p} \{|\alpha|\} > 0.$$

Since
$$\sum_{i=1}^{n} t_i\,\alpha_i=\sum_{i=1}^{n} t_i\,|\alpha_i|\,\frac{\alpha_i}{|\alpha_i|}=\sum_{i=1}^{n} t_i'\,\frac{\alpha_i}{|\alpha_i|}$$
and if
$$t_i''=\frac{t_i'}{\sum\limits_{i=1}^{n} t_i'}$$
we have the following relations
$$\begin{aligned} &1.\quad \sum_{i=1}^{n} t_i''=1, \qquad t_i''\geqq 0,\\ &2.\quad \sum_{i=1}^{n} t_i''\,\frac{\alpha_i}{|\alpha_i|}\to 0. \end{aligned} \tag{$*$}$$

But α_i's are bare points and thus we find a sequence of vectors such that
$$\begin{aligned} &1.\quad \|x_n^i\|=1,\\ &2.\quad \lim_{n\to\infty}(T-\alpha_i)\,x_n^i=0. \end{aligned}$$
It is clear that we have the following relations:
$$\begin{aligned} &1^\circ.\quad \lim \|T x_n^i\|=|\alpha_i|=\lim \|R x_n^i\|,\\ &2^\circ.\quad \lim\langle T x_n^i, x_n^i\rangle=\alpha_i,\\ &3^\circ.\quad T-\alpha_i I=UR-\alpha_i I=U(R-|\alpha_i|)-|\alpha_i|\left(-U+\frac{\alpha_i}{|\alpha_i|}I\right) \end{aligned}$$
and thus $\frac{\alpha_i}{|\alpha_i|}\in\sigma(U)$ for all i.

This with ($*$) implies that $0\in \mathrm{cl}\, W(U)$. Thus [1] Lemma 3, U is not cramped, which is a contradiction.

The theorem is proved.

Remarks. 1. In the above relation 1° the fact that α_i is bare point is essential for its validity.

2. Our proof furnishes a new proof for the case of normal operators since the proof given in [1] uses normality in an essential way.

3. If u, v denotes elements in H, $\sin^2\langle u, v\rangle$ is defined as follows
$$\sin^2\langle u, v\rangle=1-\frac{|\langle u, v\rangle|^2}{\|u\|^2\,\|v\|^2}$$
and $\sin\langle u\,v\rangle=0$ if $\langle u, v\rangle=0$.

It is known [4] that if N is a normal operator then
$$\sup_{\|u\|\neq 0}\frac{\|N u\|}{\|u\|}\sin\langle u, N u\rangle=R_N$$

where R_N is the radius of the smallest circular disc Γ_N containing $\sigma(N)$. In this section we generalize the result to the case of transaloid operators.

Theorem 2. *If T is transaloid operator then*

$$\sup_{\|u\| \neq 0} \frac{\|Tu\|}{\|u\|} \sin \langle u, Tu \rangle = R_T$$

where R_T is the radius of the smallest circular disc Γ_T *countaining* $\sigma(T)$.

We begin with a simple remark. If α, β are distinct bare points of cl $W(T) = \operatorname{conv} \sigma(T)$ (for transaloid operators) then there exist two sequences of unit vectors $\{x_n\}$, $\{y_n\}$ such that

1. $(T-\alpha) x_n \to 0, \quad (T^* - \bar{\alpha}) x_n \to 0,$
2. $(T-\beta) y_n \to 0, \quad (T^* - \bar{\beta}) y_n \to 0.$

From these relations it is clear that

$$\lim \langle x_n, y_n \rangle = 0$$

which easily follows from the identity

$$(\alpha - \beta) \langle x_n, y_n \rangle = \langle x_n, (T^* - \bar{\beta}) y_n \rangle + \langle (\alpha - T) x_n, y_n \rangle.$$

If z_T denotes the center of the above disc Γ_T it is easy to see that

$$z_T \in \operatorname{conv}(\sigma(T) \cap \partial \Gamma_T)$$

(∂ denotes the boundary) and thus we find $\lambda_1, \ldots, \lambda_r \in \sigma(T)$ such that

1°. $|\lambda_i - z_T| = R_T,$

2°. $z_T = \sum_{i=1}^{r} m_i \lambda_i, \quad \sum_{1}^{r} m_i = 1, \quad m_i > 0.$

From 1° it is clear that λ_i are bare points and thus we find $\{x_n^i\}$, $\|x_n^i\| = 1$ such that

1. $\|T x_n^i - \lambda_i x_n^i\| \to 0,$
2. $\|T^* x_n^i - \bar{\lambda}_i x_n^i\| \to 0.$

For each $i \neq j$

$$\lim \langle x_n^i, x_n^j \rangle = 0.$$

Now, it is clear that

$$\begin{aligned} &1. \quad x_n = \sum_{i=1}^{r} m_i x_n^i, \quad \|x_n\| = 1, \\ &2. \quad \|T x_n - z_T x_n\| \to R_T, \qquad\qquad (*) \\ &3. \quad \langle T x_n m x_n \rangle \to z_T. \end{aligned}$$

But for all complex numbers λ, $T - \lambda I = T_\lambda$ is normaloid and since for $x \in H$, $\|x\| = 1$

$$\|Tx\|^2 - |\langle Tx, x \rangle|^2 = \|Tx - \lambda x\|^2 - |\langle Tx, x \rangle - \lambda|^2$$

and taking $\lambda = z_T$ we obtain

$$\sup_{\|x\| \neq 0} \left\{ \frac{\|Tx\|^2}{\|x\|^2} - \frac{|\langle Tx, x\rangle|^2}{\|x\|^4} \right\} = \sup_{\|x\|=1} \{\|Tx - z_T\|^2 - |\langle Tx, x\rangle - z_T|\}$$
$$\leqq \|T - z_T\| = R_T^2.$$

From the relations (*) it is clear that we have in fact equality. The theorem is proved.

Remark. It is interesting to note that our proof combined with results in [4] gives the following result (with the above notations) if T is convexoid operator then

$$\sup_{\|x\| \neq 0} \left\{ \frac{\|Tx\|}{\|x\|^2} - \frac{|\langle Tx, x\rangle|^2}{\|x\|^4} \right\} \geqq R_T.$$

(An operator T is called convexoid if $\operatorname{cl} W(T) = \operatorname{conv}(\sigma(T))$.) We conjecture that the equality holds also for this class or for the subclass of operators with G_1 property, [5].

References

1. Berberian, S. K.: The numerical range of an operator. Duke math. J. **33**, 479–483 (1964).
2. Björck, G., Thomée, V.: A property of bounded normal operators in Hilbert space. Ark. Mat. **4**, N. 43, 551–555 (1963).
3. Durszt, E.: Remark on a paper of S. K. Berberian. Duke math. J. **35**, 795–796 (1966).
4. Hildebrandt, S.: Über den Numerischen Wertebereich eines Operators. Math. Ann. **164**, 230–247 (1966).
5. Orland, G. H.: On a class of operators. Proc. Amer. Math. Soc. **15**, 75–79 (1964).
6. Wintner, A.: Zur Theorie der beschränkten Bilinearformen. Math. Z. **30**, 228–282 (1929).

Dr. V. I. Istrăţescu
Institut de Mathématique
Calea Grivitei, 21
Bucarest 12
Rumänien

(Received August 30, 1971)

Math. Z. 124, 203-207 (1972)

On Boolean Extensions of Primal Algebras*

AWAD A. ISKANDER

Boolean extensions of universal algebras were introduced and intensively studied by Foster ([1, 2, 3]) and others (see e.g. [4, 5] and connected litterature cited there). Let $\mathbf{A}$ be a Boolean algebra and $\mathbf{P}=\langle P; F\rangle$ a finite universal algebra. Then the Boolean extension of $\mathbf{P}$ with $\mathbf{A}$ is the set $P[A]$ of all functions α from P to A such that $\alpha(x)\wedge\alpha(y)=0$ if $x, y\in P$ and $x\neq y$ and $\vee\{\alpha(x): x\in P\}=1$, together with the family F of operations extended to $P[A]$ in a way similar to convolution; i.e. let $f\in F$ be an n-ary operation $\alpha_1,\dots,\alpha_n\in A[P]$ then $f(\alpha_1,\dots,\alpha_n)=\alpha$ where

$$\alpha(x)=\vee\{\alpha_1(x_1)\wedge\cdots\wedge\alpha_n(x_n): f(x_1,\dots,x_n)=x,\ x_1,\dots,x_n\in P\}$$

for all $x\in P$. It is well known (see e.g. [1, 2, 3, 5]) that the Boolean extension is always an algebra of the same similarity type and satisfies the same identities as $\mathbf{P}$.

The aim of the present note is to show that, in case $\mathbf{P}$ is primal, $\mathbf{P}[\mathbf{A}]$ can be considered as a commutative ring with a closure operator and also with an induced distributive lattice structure so that many of the properties of $\mathbf{P}[\mathbf{A}]$ can be viewed in a natural way.

Proposition 1. *The commutative ring $\mathbf{Z}_m$ $(m>1)$ (of residue classes modulo m) with 1 together with the unary operation c such that $c(0)=0$ and $c(x)=1$ if $x\neq 0$, is a primal algebra.*

Denote this algebra by $\mathbf{Z}_m^*$.

Proof. The algebra is finite and contains $m(>1)$ elements. It remains to show that every function is a polynomial. By [6, 7] the algebra is functionally complete if there is a ternary polynomial $t(x, y, z)$ such that $t(x, y, z)=x$ if $x=y$ and $t(x, y, z)=z$ if $x\neq y$. This polynomial is

$$t(x, y, z)=(z-x)\,c(x-y)+x.$$

Since every element of $\mathbf{Z}_m$ is a multiple of 1 the algebra is primal.

Corollary 1. *An algebra $\langle P, F\rangle$ is primal iff one of the equivalent conditions* a), b) *holds.*

a) *$\langle P, F\rangle$ is equivalent to $\mathbf{Z}_m^*$, where $m=|P|$.*

b) *there is a bijection of Z_m onto P which maps the operations of $\mathbf{Z}_m^*$ onto polynomials of $\langle P, F\rangle$.*

* This research was supported in part by an N.S.F. Grant through Vanderbilt University.

This is immediate since any two primal algebras are equivalent iff they are composed of the same number of elements.

Thus every primal algebra **P** can be thought of as $\mathbf{Z}_m^*$ with $m=|P|$.

The ordering $1>2>3>\cdots>m-1>0$ turns $\mathbf{Z}_m^*$ into a chain and hence a distributive lattice with 0 and 1. The operations of join and meet are polynomials in the algebra $\mathbf{Z}_m^*$. Thus, if **A** is a Boolean algebra, then $\mathbf{Z}_m^*[\mathbf{A}]$, the Boolean extension of $\mathbf{Z}_m^*$ by **A**, inherits the distributive lattice structure as well as the ring structure and it satisfies all the identitites satisfied by $\mathbf{Z}_m^*$. So $\mathbf{Z}_m^*[\mathbf{A}]$ is a commutative ring with 1 of characteristic m and with operation c. It is also a distributive lattice with 0 and 1; the least element of the lattice is also the zero of the ring, and the largest element of the lattice is the unit of the ring. We also have

Proposition 2. $\mathbf{Z}_m^*[\mathbf{A}]$ *satisfies the following identities:*

(i) $c(c(x))=c(x)$

(ii) $(c(x))^2=c(x)$

(iii) $x\cdot(c(x))=x$

(iv) $c(x\vee y)=c(x)\vee c(y)=c(x)+c(y)-c(x)\cdot c(y)$

(v) $c(x\wedge y)=c(x)\wedge c(y)=c(x)\cdot c(y)$

(vi) $c(0)=0$

(vii) $c(1)=1$

(viii) $(x\vee y)+(x\wedge y)=x+y$

for all $x, y\in\mathbf{Z}_m^*[\mathbf{A}]$.

It is immediate that all these identities are satisfied in $\mathbf{Z}_m^*$ and hence in $\mathbf{Z}_m^*[\mathbf{A}]$.

Identities (i), (iv), (vi) and (vii) indicate that c is a closure operator on $\mathbf{P}[\mathbf{A}]$, viewed as a distributive lattice. $a\in\mathbf{Z}_m^*[\mathbf{A}]$ is said to be closed if $c(a)=a$.

Lemma. *The following conditions on an element a of* $\mathbf{Z}_m^*[\mathbf{A}]$ *are equivalent:*

1) *a is closed* $(c(a)=a)$,

2) *a is idempotent* $(a^2=a)$,

3) *a has a complement (i.e. there is* $a'\in Z_m^*[A]$, $a\vee a'=1$ $a\wedge a'=0$).

Proof. Consider the representation of $\mathbf{Z}_m^*[\mathbf{A}]$ as a subdirect I-th power of $\mathbf{Z}_m^*$. Each of the conditions 1), 2), 3) is equivalent to the condition that for every $i\in I$, the i-th projection of a is either 0 or 1. This condition is also equivalent to the condition that $a(k)=0$ for all $k\in Z_m$, $k\neq 0$, $k\neq 1$.

The complement a' of a, if it exists, is $(1-a)$.

Proposition 3. *The set of all closed elements of* $Z_m^*[A]$ *is a Boolean sublattice of the lattice* $\mathbf{Z}_m^*[\mathbf{A}]$; *moreover* $\phi(a)=a(1)$ *is an isomorphism of the Boolean algebra of all closed elements onto* **A** *and c is a lattice retraction of* $\phi^{-1}(A)$.

Proof. That the set of all closed and hence complemented elements of a distributive lattice is a Boolean sublattice is well known. Let a, b be closed

$$\begin{aligned}\phi(a \vee b) &= (a \vee b)(1)\\ &= \bigvee \{a(i) \wedge b(j)\colon i \vee j = 1, i, j \in Z_m\}\\ &= \left(\bigvee \{a(1) \wedge b(j)\colon j \in Z_m\}\right) \vee \left(\bigvee \{a(i) \wedge b(1)\colon i \in Z_m\}\right)\\ &= \left[a(1) \wedge \left(\bigvee \{b(j)\colon j \in Z_m\}\right)\right] \vee \left[b(1) \wedge \left(\bigvee \{a(i)\colon i \in Z_m\}\right)\right]\\ &= a(1) \vee b(1)\end{aligned}$$

$$\phi(a') = a'(1) = a(0) = (a(1))' = (\phi(a))'$$

(since a is closed iff $a(k) = 0$ for all $k \neq 0$, $k \neq 1$ and hence $a(0) = (a(1))'$).

The identitites (iv), (v), (vi), (vii) of Proposition 2 imply that c is an endomorphism of the lattice $\mathbf{Z}_m^*[\mathbf{A}]$; identity (i) shows that it is a retraction.

We will identify A with $\phi^{-1}(A)$ by the isomorphism ϕ. The following result is essentially due to Foster.

Proposition 4. *Let* $\mathbf{A}, \mathbf{B}$ *be Boolean algebras and let* h *be a homomorphism of* $\mathbf{Z}_m^*[\mathbf{A}]$ *into* $\mathbf{Z}_m^*[\mathbf{B}]$. *Then the restriction* h_1 *of* h *to* $\mathbf{A}$ *is a Boolean homomorphism of* $\mathbf{A}$ *into* $\mathbf{B}$; *moreover* $(h(a))(k) = h_1(a(k))$, $k \in \mathbf{Z}_m$. h *is one-to-one iff* h_1 *is one-to-one and* h *is onto iff* h_1 *is onto.*

This is clear since h is necessarily a lattice homomorphism which maps $0, 1$ of $\mathbf{Z}_m^*[\mathbf{A}]$ onto $0, 1$ of $\mathbf{Z}_m^*[\mathbf{B}]$ and hence complemented elements to complemented elements.

Proposition 4 has the categorical consequence that the functors $\mathbf{A} \to \mathbf{Z}_m^*[\mathbf{A}]$, $h_1 \to h$ and $\mathbf{Z}_m^*[\mathbf{A}] \to \mathbf{A}$, $h \to h_1$ provide the equivalence of the category of all Boolean algebras and the category of all Boolean extensions of $\mathbf{Z}_m^*$. These functors preserve monomorphisms, epimorphisms and cartesian products and hence they preserve injectives, projectives and free products.

Corollary 2. $\mathbf{Z}_m^*[\mathbf{A}]$ *is injective iff the lattice* $\mathbf{Z}_m^*[\mathbf{A}]$ *is complete.*

This is immediate since $\mathbf{Z}_m^*[\mathbf{A}]$ is injective iff $\mathbf{A}$ is complete and hence iff $\mathbf{Z}_m^*[\mathbf{A}]$ (as a lattice) is complete.

Proposition 5. *If* m *is a power of* 2 *and* $\mathbf{Z}_m^*[\mathbf{A}]$ *is free then* $\mathbf{A}$ *is free.*

Proof. If $\mathbf{Z}_m^*[\mathbf{A}]$ has one generator then $|Z_m^*[A]| = m^m$ and $\mathbf{Z}_m^*[\mathbf{A}]$ is the m-th direct power of $\mathbf{Z}_m^*$. $\mathbf{A}$ is then the m-th direct power of $\mathbf{2}$, i.e. $|A| = 2^m$ and since m is a power of 2 then $\mathbf{A}$ is free.

Since $\mathbf{Z}_m^*[\mathbf{A}] \to \mathbf{A}$ preserves free products, Proposition 5 follows for any set of free generators.

Remark. If $m = 2^k$ and $\mathbf{Z}_m^*[\mathbf{A}]$ is freely generated by n generators, then A is freely generated by $k\,n$ generators.

The converse of Proposition 5 is not always true. In fact $\mathbf{Z}_m^*[\mathbf{A}]$ is free whenever $\mathbf{A}$ is free iff $m = 2$.

Since $\mathbf{Z}_m^*[\mathbf{A}]$ is a ring, every congruence relation is completely determined by any of its classes. If θ is a congruence relation on $\mathbf{Z}_m^*[\mathbf{A}]$, then $[0]\,\theta = \{x\colon x \in Z_m^*[A],\ 0\,\theta\,x\}$ is an ideal of $\mathbf{Z}_m^*[A]$ (as a ring) which is closed under c, i.e. $c([0]\,\theta) \subseteq [0]\,\theta$ such ideals, i.e. ideals of the ring $\mathbf{Z}_m^*[\mathbf{A}]$, that are closed under c, will be called c-ideals. Every congruence on $\mathbf{Z}_m^*[\mathbf{A}]$ is determined completely by a c-ideal and conversely.

Proposition 6. *$M \to M \cap A = c(M)$ is an isomorphism of the lattice of all c-ideals of $\mathbf{Z}_m^*[\mathbf{A}]$ onto the lattice of all ideals of* $\mathbf{A}$.

Proof. Since

$$x \vee y = (x \vee y)\cdot c(x \vee y) = (x \vee y)\big(c(x) \vee c(y)\big) = (x \vee y)\big(c(x) + c(y) - c(x)\cdot c(y)\big)$$

and

$$x \wedge y = (x \wedge y)\cdot c(x \wedge y) = (x \wedge y)\,\big(c(x) \wedge c(y)\big) = (x \wedge y)\cdot\big(c(x)\cdot c(y)\big),$$

every c-ideal of $\mathbf{Z}_m^*[\mathbf{A}]$ is also an ideal of the lattice $\mathbf{Z}_m^*[\mathbf{A}]$. Also since c is a lattice homomorphism of $\mathbf{Z}_m^*[\mathbf{A}]$ onto $\mathbf{A}$, it maps ideals of $\mathbf{Z}_m^*[\mathbf{A}]$ onto ideals of $\mathbf{A}$. Thus $c(M)$ is an ideal of $\mathbf{A}$. But $c(M) \subseteq M$ and so $M \cap \mathbf{A} \subseteq c(M)$. To show that $c(M) \subseteq M \cap A$, let $a \in M \cap A$. Then $a = c(a) \in c(M)$.

That $M \to M \cap A$ is an isomorphism will be obvious once we show that $M = c^{-1}\big(c(M)\big)$. Let $x \in c^{-1}\big(c(M)\big)$. Then there is some $y \in M$ such that $c(x) = c(y) \in M$. Now $x = x\cdot c(x) = x \cdot c(y) \in M$.

If M is a maximal c-ideal of $\mathbf{Z}_m^*[\mathbf{A}]$ and $a \in Z_m^*[A]$, then M intersects the set $\{a, a+1, a+2, \ldots, a+m-1\}$ in precisely one element.

Proposition 7. *Let $M_0, M_1, \ldots, M_q$ $(q<m)$ be $q+1$ distinct maximal c-ideals of $\mathbf{Z}_m^*[\mathbf{A}]$. Then there is $a \in Z_m^*[A]$ with the property that $a+k \in M_k$, $k=0, 1, \ldots, q$.*

Proof. Since M_i and M_j are distinct for distinct i and j and both M_i and M_j are maximal, there is $b_{ij} \in M_i$ $b_{ij} \notin M_j$; the b_{ij} may be chosen closed since $x \in M_i$ iff $c(x) \in M_i$. Put

$$b_j = \pi\{b_{ij}\colon \text{all } i \text{ distinct from } j\}.$$

$c(M_j)$ is maximal and hence prime in $\mathbf{A}$. For every i distinct from j b_{ij} does not belong to $c(M_j)$. The product of idempotents (closed) is again idempotent (closed) and hence $b_j \notin c(M_j)$. Hence $b_j \notin M_j$. But, as $b_{ij} \in M_i$, then $b_j \in M_i$ for every $i \neq j$. Put

$$d_j = (m-j)\,b_j \quad \text{and} \quad a = d_1 + d_2 + \cdots + d_q.$$

$d_1, d_2, \ldots, d_q \in M_0$, hence $a \in M_0$.

Let π_k denote the natural homomorphism of $\mathbf{Z}_m^*[\mathbf{A}]$ onto its quotient by M_k, which is isomorphic to $\mathbf{Z}_m^*$

$$\begin{aligned}\pi_k(a+k) &= \textstyle\sum\{\pi_k(d_i)\colon i \neq k\} + \pi_k(d_k+k)\\ &= \pi_k(d_k+k) \qquad (\text{since } d_i \in M_k \text{ if } i \neq k)\\ &= \pi_k\big((m-k)\,b_k + k\big)\\ &= (m-k)\,\pi_k(b_k) + k = (m-k)\cdot 1 + k = 0\end{aligned}$$

i.e. $a+k \in M_k$. $k=0, 1, 2, \ldots, q$.

References

1. Foster, A. L.: Generalized "Boolean" theory of universal algebras, part I: Subdirect sums and normal representation theorem. Math. Z. **58**, 306–336 (1953).
2. – Generalized "Boolean" theory of universal algebras, part II: Identities and subdirect sums of functionally complete algebras. Math. Z. **59**, 191–199 (1953).
3. – The identities of – and unique subdirect factorization within – classes of universal algebras. Math. Z. **62**, 171–188 (1955).
4. Gould, M. I., Grätzer, G.: Boolean extensions and normal subdirect powers of finite universal algebras. Math. Z. **99**, 16–25 (1967).
5. Grätzer, G.: Universal algebra. Princeton: Van Nostrand 1968.
6. Pixley, A. F.: Distributivity and permutability of congruence relations in equational classes of algebras. Proc. Amer. Math. Soc. **14**, 105–109 (1963).
7. Werner, H.: A characterization of functionally complete algebras. Notices Amer. Math. Soc. **17**, No. 2, 70T–A51 (1970).

Prof. A. A. Iskander
Department of Mathematics
University of Southwestern Louisiana
Lafayette, Louisiana 70501
USA

(Received January 25, 1971)

Math. Z. 124, 208–212 (1972)

On Gamma-Type Approximation Operators

DANY LEVIATAN

1. Introduction

Let A denote the space of all complex-valued functions which are measurable in $(0, \infty)$, bounded in any interval $[r, R]$, $0<r<R<\infty$, and such that $f(t)=O(e^{a/t})$ as $t\to 0$ and $f(t)=O(t^b)$ as $t\to\infty$ for some positive constants a and b. The Gamma operators were introduced by Müller [6] in the following way. With $f\in A$ we associate the functions $G_n f$ defined by

$$G_n(f; x)=\frac{x^{n+1}}{n!}\int_0^\infty u^n e^{-ux} f\left(\frac{n}{u}\right) du, \qquad x>0, \tag{1.1}$$

and we observe that for a fixed x, $G_n(f; x)$ exists at least for

$$n\geqq \max\{[a/x]+1, [b]\}$$

($[y]$ denotes the largest integer not exceeding y). It was proved by Müller [6] (see also [5]) that for every point of continuity, x, of $f(t)$

$$\lim_{n\to\infty} G_n(f; x)=f(x). \tag{1.2}$$

Furthermore if $f(t)$ is continuous in $[c, d]$, $0<c<d<\infty$, then (1.2) holds uniformly in $[c, d]$.

The kernel $u^n e^{-ux}$ of G_n is obtained from the Laplace kernel e^{-ux} by application n times of the differential operator $(-d/dx)$. It is well known that for a fixed u the function e^{-ux} is a completely monotonic function therefore one may ask if it is possible to generalize the Gamma approximation operators and to obtain some of their properties using other completely monotonic or generalized completely monotonic functions instead of the Laplace kernel.

We shall discuss here two types of Gamma-operators. One will be defined by means of some convolution kernels and the other will be defined by means of Studden's kernels [7]. In both cases approximation properties of the operators are derived.

The author wishes to thank Professor W. Meyer-König whose questions and comments during a Colloquium talk the author gave in Stuttgart have resulted in this article.

2. The Gamma Operators I

Let the sequence $\{\gamma_i\}$, $i\geqq 0$, satisfy

$$0=\gamma_0<\gamma_1<\cdots<\gamma_n<\cdots\to\infty, \qquad \sum_{n=1}^{\infty}\frac{1}{\gamma_n}=\infty \quad \text{and} \quad \sum_{n=1}^{\infty}\frac{1}{\gamma_n^2}<\infty. \tag{2.1}$$

Denote for a complex number ζ

$$E(\zeta,\gamma)=\prod_{\nu=1}^{\infty}\left(1+\frac{\zeta}{\gamma_\nu}\right)e^{-\zeta/\gamma_\nu}$$

where the convergence of the infinite product is guaranteed by (2.1) (see [2] Ch. I §6.1) and, for $0<\theta<\infty$,

$$\omega(\theta,\gamma)=\frac{1}{2\pi i}\int_{\delta-i\infty}^{\delta+i\infty}\frac{\theta^{-\zeta}d\zeta}{\zeta E(\zeta,\gamma)},\qquad \delta>0.$$

It can be shown that for the sequence $\gamma_n=n$, $n\geqq 0$, we have $\omega(\theta,\gamma)=\exp(-\theta e^{-c})$ where c is Euler's constant (see [1] p. 495). In this case $\omega(xt,\gamma)$ is the Laplace kernel (after rescaling one or both coordinates). In general, $\omega(\theta,\gamma)$ is differentiable infinitely often in $0<\theta<\infty$. We associate with it the sequence of functions

$$\Omega_0(x,t)=\omega(xt,\gamma),\quad \Omega_1(x,t)=\frac{\partial}{\partial x}\Omega_0(x,t),\quad \Omega_{n+1}(x,t)=\frac{\partial}{\partial x}\left[\frac{\Omega_n(x,t)}{x^{\gamma_n-\gamma_{n-1}-1}}\right],\tag{2.2}$$

$$n\geqq 1,\qquad 0<x,t<\infty.$$

As it turns out

$$(-1)^n\Omega_n(x,t)\geqq 0,\qquad n=0,1,2,\ldots,\qquad 0<x,t<\infty.$$

and therefore, as a function of x, $\omega(xt,\gamma)$ is a generalized completely monotonic function (with respect to the sequence of derivatives defined by (2.2)). In our first generalization of the Gamma-operators we shall use the kernel $\omega(xt,\gamma)$ and the sequence of derivatives (2.2) instead of the Laplace kernel and the sequence of ordinary derivatives.

Let $f(t)$ be a bounded function in $[0,\infty)$ and associate with f the functions $G_n(f;\gamma;\bullet)$ defined by

$$G_n(f;\gamma;x)=\frac{1}{\gamma_n!}\int_0^{\infty}(-1)^{n+1}\Omega_{n+1}(u,x)\,u^{\gamma_n}f\left(\frac{\beta_n}{u}\right)du,\qquad n\geqq 0,\tag{2.3}$$

where

$$\gamma_n!=\gamma_1\cdot\cdots\cdot\gamma_n,\qquad n\geqq 1,\ \gamma_0!=1\quad\text{and}\quad \beta_n=\exp\left(\sum_{i=1}^{n}\frac{1}{\gamma_i}\right),\qquad n\geqq 1,\ \beta_0=0.$$

Our discussion above shows that for the sequence $\gamma_n=n$, $n\geqq 0$, the functions $G_n(f;\gamma;x)$ reduce to those in (1.1) (after rescaling) thus we shall call the operators $G_n\colon f\to G_n(f;\gamma;\bullet)$ the Gamma-operators. This is a new setup for the inversion operators of the convolution transform. For by [1] (1.38)

$$\frac{(-1)^{x+1}}{\gamma_n!}\Omega_{n+1}(u,x)\,u^{\gamma_n+1}=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{(ux)^{-\zeta}d\zeta}{\beta_n^{-\zeta}E_{n+1}(\zeta,\gamma)},$$

where $E_{n+1}(\zeta,\gamma)\equiv\prod_{\nu=n+1}^{\infty}\left(1+\frac{\zeta}{\gamma_\nu}\right)e^{-\zeta/\gamma_\nu}$ and a change of variables $u=e^v$ and $x=e^y$ yields

$$\frac{(-1)^{n+1}}{\gamma_n!}\Omega_{n+1}(u,x)\,u^{\gamma_n+1}=G_n(v+y-\lambda_n).\tag{2.4}$$

Here $G_n(t)$ is the convolution kernel of [2] §4.1 and $\lambda_n=\sum_{\nu=1}^{n}\frac{1}{\gamma_\nu}$. Therefore

$$G_n(f;\gamma;x)=\int_{-\infty}^{\infty} G_n(v+y-\lambda_n)\,f(e^{\lambda_n-v})\,dv$$
$$=\int_{-\infty}^{\infty} G_n(y-u)\,f(e^u)\,du.$$

Hence a slight modification of [2] Ch. III Th. 7.1 yields the following.

Theorem 1. *Let $f(t)$ be bounded in $(0,\infty)$ and let $x>0$ be a point of continuity of $f(t)$. Then*

$$\lim_{n\to\infty} G_n(f;\gamma;x)=f(x). \tag{2.5}$$

Furthermore if $f(t)$ is continuous in $[a,b]$, $0<a<b<\infty$, then (2.5) *holds uniformly in $x\in[a,b]$.*

3. The Gamma Operators II

The above discussion brings us to defining Gamma approximation operators with respect to yet another kernel. Let $\{w_i(t)\}$ $(i\geqq 0)$ be an infinite sequence of functions which are positive on $[0,\infty)$. Define the sequence of differential operators, D_i, by

$$D_i\,\phi(t)=\frac{d}{dt}\,\frac{\phi(t)}{w_i(t)}, \qquad i=0,1,2,\ldots. \tag{3.1}$$

Studden [7] defined the concept of generalized complete monotonicity with the operators D_i replacing ordinary differentiation. Also under the assumptions that $w_0(t)=1$ and that there exists a function $f(t)$ on $[0,\infty)$ such that

$$\begin{aligned}&\text{(i) } 0<f(t)<1,\ t\in[0,\infty);\ \lim_{t\to\infty} f(t)=0;\ \int_0^\infty f(t)\,dt<\infty;\ \text{and } \inf[1-f(t)]>0,\\ &\text{(ii) } 1-f(t)\leqq w_i(t)\leqq 1+f(t), \qquad t\in[0,\infty),\ i=1,2,\ldots,\end{aligned} \tag{3.2}$$

Studden [7] produced a kernel $\psi(t;x)$ which as a function of t, belongs to $C^\infty[0,\infty)$ and is generalized completely monotonic in t, i.e., $\psi(t;x)\geqq 0$, $t,x\in[0,\infty)$ and $(-1)^n D_{n-1}D_{n-2}\ldots D_0\,\psi(t;x)\geqq 0$, $n=1,2,\ldots$, $t,x\in[0,\infty)$, where the differentiation is with respect to t. If we denote

$$\begin{aligned}&G_0\,\psi(t;x)=\psi(t;x)\\ &G_k\,\psi(t;x)=(-1)^k D_{k-1}\ldots D_0\,\psi(t;x), \qquad k\geqq 1,\end{aligned} \tag{3.3}$$

then it was proved in [7] that

$$0\leqq G_k\,\psi(t;x)\leqq x^k\,w_k(t)\,e^{-x(t-2F)}, \qquad k\geqq 0, \tag{3.4}$$

where $F=\int_0^\infty f(t)\,dt<\infty$.

Set $\psi_0(t;x)=1$, $0\leqq t\leqq x$ and $\psi_0(t;x)=0$, $t>x$, and for $n\geqq 1$

$$\psi_n(t;x)=\begin{cases}\int\limits_t^x w_n(\xi_n)\int\limits_t^{\xi_n} w_{n-1}(\xi_{n-1})\dots\int\limits_t^{\xi_2} w_1(\xi_1)\,d\xi_1\dots d\xi_n, & 0\leqq t\leqq x,\\ 0, & t>x.\end{cases}$$

Then by [7] (3.1)

$$\psi_n(t;x)\leqq\frac{(x-t+F)^n}{n!}. \tag{3.5}$$

Repeated integration by parts yields (see [7] (4.12)) for arbitrary x, $0\leqq x<\infty$, and $n\geqq 0$,

$$\psi(t;x)=\sum_{k=0}^{n}\psi_k(t;u)\frac{G_k\,\psi(u;x)}{W_k(u)}+\int\limits_t^u\psi_n(t;v)\,G_{n+1}\,\psi(v;x)\,dv,\qquad u\geqq t\geqq 0.$$

Now $\psi(0;x)=1$, therefore

$$1=\sum_{k=0}^{n}\psi_k(0;u)\frac{G_k\,\psi(u;x)}{w_k(u)}+\int\limits_0^u\psi_n(0;v)\,G_{n+1}\,\psi(v;x)\,dv,\qquad u\geqq 0.$$

Letting $u\to\infty$ it follows by (3.4) and (3.5) that for each $k\geqq 0$

$$0\leqq\lim_{u\to\infty}\psi_k(0;u)\frac{G_k\,\psi(u;x)}{w_k(u)}\leqq\lim_{u\to\infty}\frac{(u+F)^k}{k!}\,x^k\,e^{-x(u-2F)}=0,$$

hence

$$1=\int\limits_0^\infty\psi_n(0;v)\,G_{n+1}\,\psi(v;x)\,dv,\qquad 0\leqq x\leqq\infty. \tag{3.6}$$

For the sequence of functions $w_i(t)\equiv 1$, $i=0,1,2,\dots$ which obviously satisfy the requirements (3.2), the kernel $\psi(t;x)$ reduces to the Laplace kernel e^{-xt}. In this case

$$\psi_n(0;u)\,G_{n+1}\,\psi(u;x)=\frac{x^{n+1}}{n!}\,u^n\,e^{-ux},$$

the kernel defining the Gamma-operators. Let us therefore associate with a bounded function f on $(0,\infty)$ the functions $\Gamma_n(f;w;\bullet)$ defined by

$$\Gamma_n(f;w;x)=\int\limits_0^\infty\psi_n(0;u)\,G_{n+1}\,\psi(u;x)\,f\left(\frac{n}{u}\right)du,\qquad n=0,1,2,\dots$$

which are well defined by virtue of (3.6). In fact (3.6) turns out to be

$$\Gamma_n(1;w;x)\equiv 1\qquad n=0,1,2,\dots.$$

We prove the following

Theorem 2. *Let $f(t)$ be bounded in $(0,\infty)$ and let $0<x<\infty$ be a point of continuity of $f(t)$. Then*

$$\lim_{n\to\infty}\Gamma_n(f;w;x)=f(x). \tag{3.7}$$

Furthermore if $f(t)$ is continuous in $[a,b]$, $0<a<b<\infty$, then (3.7) holds uniformly in $[a,b]$.

Proof. By (3.6)

$$\Gamma_n(f; w; x) - f(x) = \int_0^\infty \psi_n(0; u)\, G_{n+1}\, \psi(u; x) \left[f\left(\frac{n}{u}\right) - f(x) \right] dx$$

and so

$$|\Gamma_n(f; w; x) - f(x)| \leqq \omega(\delta) + 2M \int_{\left|\frac{n}{u} - x\right| \geqq \delta} \psi_n(0; u)\, G_{n+1}\, \psi(u; x)\, du$$

where $\omega(\delta)$ is the modulus of continuity of f at x and $M = \sup |f(t)|$. Now by (3.4) and (3.5)

$$\int_{\left|\frac{n}{u} - x\right| \geqq \delta} \psi_n(0; u)\, G_{n+1}\, \psi(u; x)\, du \leqq \frac{x^{n+1}}{n!} \int_{\left|\frac{n}{u} - x\right| \geqq \delta} (u + F)^n e^{-x(u - 2F)} du$$

$$\leqq \frac{x^{n+1}}{n!} e^{3xF} \int_{\left|\frac{n}{v} - x\right| \geqq \eta} v^n e^{-xv} dv$$

for some $0 < \eta < \delta$ and all $n \geqq n_0$. Hence by [5] Hilfssatz 2.3,

$$\lim_{n \to \infty} \int_{\left|\frac{n}{u} - x\right| \geqq \delta} \psi_n(0; u)\, G_{n+1}\, \psi(u; x)\, du = 0,$$

uniformly in x in any finite interval.

References

1. Badaljan, G.V.: The application of a convolution transform to the theory of the generalized Stieltjes moment problem. Math. USSR-Izv. **1**, 475–514 (1967). [English trans. of Izvestija Akad. Nauk SSSR, Ser. Mat. 31 (1967).]
2. Hirschman, I. I., Widder, D.V.: The convolution transform. Princeton: University Press 1955.
3. Leviatan, D.: An application of a convolution transform to the sequence-to-function analogues of Hausdorff transformations. J. Analyse Math. **24**, 173–189 (1971).
4. — A representation theorem and approximation operators arising from inequalities involving differential operators. Trans. Amer. Math. Soc. (to appear).
5. Lupas, A., Müller, M.: Approximationseigenschaften der Gammaoperatoren. Math. Z. **98**, 208–226 (1967).
6. Müller, M.: Die Folge der Gammaoperatoren. Dissertation, Stuttgart 1967.
7. Studden, W.J.: A generalization of Bernstein's theorem and a differential inversion formula. Trans. Amer. Math. Soc. **142**, 81–92 (1969).

Prof. D. Leviatan
Department of Mathematics
Tel Aviv University,
Tel Aviv, Israel

(Received May 11, 1971)

Math. Z. 124, 213–214 (1972)

A Note on Hemi-Primal Algebras

ALDEN F. PIXLEY

Let H be a universal algebra. In [1] Foster defined a function $f: H^n \to H$ to be *h-conservative* if for each congruence relation θ on H and elements x_i, y_i of H for which $x_i \theta y_i$, $i=1, \dots, n$, it follows that $f(x_1, \dots, x_n) \theta f(y_1, \dots, y_n)$. If H is a finite non-trivial algebra H is then defined to be *hemi-primal* if each h-conservative $f: H^n \to H$ is representable by a polynomial. H is *linear* hemi-primal provided the congruence lattice of H is a chain. In [1] the following is established:

Theorem. *A linear hemi-primal algebra H is hemi-categorial (i.e.: each member of the variety generated by H is isomorphic with a subdirect product of homomorphic images of H).*

In this note we apply the methods of [3] and [4] to provide a major simplification of the proof. Our proof requires the following

Lemma. *If H is a linear hemi-primal algebra, there is a polynomial $q(x, y, z)$ such that*

$$q(x, x, z) = q(z, x, z) = q(z, x, x) = z \tag{1}$$

are identities of H. Hence Id(H) *are permutable and distributive.*[1]

Proof. Let the congruences of H be $\theta_0 = 0 < \theta_1 < \cdots < \theta_n < I$. For each i, let $[x]_i$ denote the θ_i congruence class containing x. On H/θ_n define f_n by

$$\begin{aligned} f_n([x]_n, [y]_n, [z]_n) &= [z]_n \quad \text{if } [x]_n = [y]_n, \\ &= [x]_n \quad \text{otherwise.} \end{aligned}$$

Then f_n satisfies (1) on H/θ_n.

Now assume f_i has been defined on H/θ_i. Define f_{i-1} on H/θ_{i-1} by

$$\begin{aligned} f_{i-1}([x]_{i-1}, [y]_{i-1}, [z]_{i-1}) &= [z]_{i-1} \quad \text{if } [x]_{i-1} = [y]_{i-1}, \\ &= [x]_{i-1} \quad \text{if } [z]_{i-1} = [x]_{i-1} \text{ or } [y]_{i-1}, \\ &= \text{an arbitrary element in} \\ &\quad\ f_i([x]_i, [y]_i, [z]_i) \quad \text{otherwise.}^2 \end{aligned}$$

[1] i.e.: each algebra in the variety generated by H has permutable congruence relations and a distributive congruence lattice. (See [4] where the equivalence of (1) and these conditions is established.)

[2] For convenience we are considering the elements of θ_i congruence classes to be θ_{i-1} congruence classes.

Observe that f_i satisfies (1) for all $i=0, 1, \ldots, n$. Also, since $\theta_{i-1}<\theta_i$, we have for all x, y, z and all $i \geqq 1$,

$$f_{i-1}([x]_{i-1}, [y]_{i-1}, [z]_{i-1}) \in f_i([x]_i, [y]_i, [z]_i).^2$$

Now suppose $x, y, z, x', y', z' \in H$ and for some congruence θ_i, $x\,\theta_i\,x'$, $y\,\theta_i\,y'$, $z\,\theta_i\,z'$. Then

$$f_0(x, y, z) \in f_1([x]_1, [y]_1, [z]_1) \in \cdots \in f_i([x]_i, [y]_i, [z]_i),$$

and

$$f_0(x', y', z') \in f_1([x']_1, [y']_1, [z']_1) \in \cdots \in f_i([x']_i, [y']_i, [z']_i).$$

But $f_i([x]_i, [y]_i, [z]_i) = f_i([x']_i, [y']_i, [z']_i)$. Therefore $f_0(x, y, z)\,\theta_i\,f_0(x', y', z')$. Hence f_0 is h-conservative on $H = H/\theta_0$, so that there is a polynomial q representing f_0 in H. Hence q satisfies (1). By Lemma 2.3 of [4] it follows that $\mathrm{Id}(H)$ are permutable and distributive.

Proof of Theorem. Let H be hemi-primal and let B satisfy the identities of H. By Theorem 2.3 of [3] we need only consider the case of finitely generated B.

Hence for some integer k and congruence θ,

$$B \cong F_k(H)/\theta$$

where $F_k(H)$ is the H-free algebra with k free generators. But since H has no proper subalgebras (see [1]), $F_k(H)$ is isomorphic with a finite subdirect power of H. Also by the Lemma, the congruence lattice of B is distributive. The result then follows by Theorem 2.5 of [2].

Added in Proof. By a slight extension of our present argument one can easily extend the above lemma and theorem to the more general case where H is any hemi-primal algebra with permutable congruences and a distributive congruence lattice.

References

1. Foster, A. L.: Congruence relations and functional completeness in universal algebra; structure theory of hemi-primals I. Math. Z. **113**, 293–308 (1970).
2. – Pixley, A. F.: Semi-categorial algebras II. Math. Z. **85**, 169–184 (1964).
3. Pixley, A. F.: Functionally complete algebras generating distributive and permutable classes. Math. Z. **114**, 361–372 (1970).
4. – The ternary discriminator function in universal algebra. Math. Ann. **191**, 167–180 (1971).

Prof. Alden F. Pixley
Dept. of Mathematics
Harvey Mudd College
Claremont, California 91711
USA

(Received February 19, 1971)

Math. Z. 124, 215–236 (1972)

Kan-Bedingungen und abstrakte Homotopietheorie

Klaus Heiner Kamps

Einleitung und Überblick

Wir zeigen in dieser Arbeit, daß einige zentrale Sätze der Homotopietheorie, wie etwa der Homotopiesatz für h-Faserungen ([4], (7.22)) und ein Satz von Dold über fasernweise Homotopieäquivalenzen ([5], 6.1, [4], (6.21)) von formaler Natur sind und auf Kan-Bedingungen geeigneter semikubischer Komplexe beruhen. Der einfache geometrische Hintergrund für die benötigten Kan-Bedingungen ist die Tatsache, daß die Vereinigung dreier Seiten des Einheitsquadrats $[0, 1] \times [0, 1]$ im $\mathbb{R}^2$ Retrakt von $[0, 1] \times [0, 1]$ ist, und die entsprechende Tatsache in der Dimension 3.

Unserem axiomatischen Zugang zur Homotopietheorie liegt der Begriff des *Homotopiesystems* (*Zylinderfunktor*) in einer Kategorie zugrunde, ein Begriff, der auf Kan [11] zurückgeht (1.1). Mit seiner Hilfe definieren wir in 1. die Grundbegriffe der Homotopietheorie, wie z.B. Faserung, h-Faserung, Cofaserung, h-Cofaserung. Die Definitionen in 1. sind konsistent mit den Definitionen in [10], wo die Grundbegriffe der Homotopietheorie mit Hilfe des allgemeineren Begriffs des verallgemeinerten Homotopiesystems in einer Kategorie erklärt werden (s. [10], 6.). Wir können daher in dieser Arbeit auf die Ergebnisse von [10] zurückgreifen. Im übrigen verwenden wir die kategorientheoretische Terminologie wie in [10].

In 2. befassen wir uns mit semikubischen Komplexen und Kan-Bedingungen für semikubische Komplexe, in 3. werden Kan-Bedingungen für Homotopiesysteme definiert. 4. enthält die Zerlegung eines Morphismus' in eine Cofaserung und eine Homotopieäquivalenz, 5. und 6. unsere kategorientheoretische Fassung des Homotopiesatzes für h-Faserungen (5.4) und des erwähnten Satzes von Dold (6.1).

In 7. dualisieren wir die Begriffe der Abschnitte 1. und 3. Dual zum Begriff des Homotopiesystems hat man den Begriff des *Cohomotopiesystems* in einer Kategorie. Ein Cohomotopiesystem wird insbesondere immer dann induziert, wenn der Zylinderfunktor eines Homotopiesystems einen adjungierten Funktor hat. Als Anwendung erhält man aus dem zu (6.1) dualen Satz bei Vorliegen einer solchen Adjungiertheit einen (6.1) entsprechenden formalen Satz über Homotopieäquivalenzen unter A (7.17).

In 8. wenden wir unsere Methoden an und beweisen einen formalen Satz über das Zusammenkleben (*glueing*) von Homotopieäquivalenzen (8.2) (vgl. [1], 7.5.7, [3], Lemma 1). Dualisierung liefert einen entsprechenden Satz über das Co-Zusammenkleben (*coglueing*) von Homotopieäquivalenzen (8.7) (vgl. [2], 1.2).

Diese Arbeit ist mit Ausnahme von Abschnitt 8 in [8] enthalten. Über einen Teil der Ergebnisse wurde bereits in [9] berichtet.

Herrn Professor D. Puppe danke ich für zahlreiche wertvolle Anregungen.

1. Homotopiesysteme. Faserungen. Cofaserungen

(1.1) **Definition.** Ein *Homotopiesystem* $\mathbf{I}=(I, j_0, j_1, q)$ in einer Kategorie $\mathscr{C}$ besteht aus einem kovarianten Funktor $I: \mathscr{C} \to \mathscr{C}$ (*Zylinderfunktor*) und drei natürlichen Transformationen $j_0, j_1: \mathrm{id}_{\mathscr{C}} \to I$, $q: I \to \mathrm{id}_{\mathscr{C}}$ mit $q j_0 = q j_1 = \mathrm{id}$.

(1.2) Das für uns interessanteste Beispiel eines Homotopiesystems ist das Homotopiesystem $\mathbf{T}$ in der Kategorie der topologischen Räume, dessen Zylinderfunktor einem Raum X den *Zylinder* über X, d.h. das Produkt $X \times [0, 1]$ von X mit dem Einheitsintervall der reellen Zahlen zuordnet und dessen natürliche Transformationen j_0, j_1, q durch die Abbildungen $x \mapsto (x, 0)$, $x \mapsto (x, 1)$, $(x, t) \mapsto x$ gegeben sind ($x \in X$, $t \in [0, 1]$).

Im folgenden sei $\mathscr{C}$ eine Kategorie und $\mathbf{I}=(I, j_0, j_1, q)$ ein Homotopiesystem in $\mathscr{C}$. Relativ zu $\mathbf{I}$ definieren wir die wichtigsten Begriffe der Homotopietheorie (vgl. [4]).

(1.3) **Definition.** Sind $f, g: X \to Y$ Morphismen von $\mathscr{C}$, so heißt f *homotop* zu g (in Zeichen: $f \simeq g$), genau wenn ein Morphismus $\varphi: IX \to Y$ von $\mathscr{C}$ existiert mit $\varphi_0 = f$ und $\varphi_1 = g$.

Dabei sei $\varphi_\nu := \varphi j_{\nu X}$ ($\nu = 0, 1$).

Ein solches φ heißt *Homotopie* von f nach g ($\varphi: f \simeq g$).

Die Homotopierelation ist i.a. weder symmetrisch noch transitiv, sie ist aber reflexiv und mit der Komposition von Morphismen verträglich ([10], 2.4).

Für Objekte B, A von $\mathscr{C}$ induziert $\mathbf{I}$ ferner Relationen in der Kategorie $\mathscr{C}_B$ der Objekte über B und der Kategorie $\mathscr{C}^A$ der Objekte unter A.

(1.4) **Definition.** Sind $p: E \to B$, $p': E' \to B$ Objekte von $\mathscr{C}_B$ und $f, g: p \to p'$ Morphismen über B (d.h. $p' f = p' g = p$), so heißt f *homotop über B* zu g ($f \underset{B}{\simeq} g$), wenn eine Homotopie $\varphi: f \simeq g$ existiert mit $p' \varphi = p\, q_E$.

φ heißt dann *Homotopie über B*.

(1.5) **Definition.** Sind $i: A \to X$, $i': A \to X'$ Objekte von $\mathscr{C}^A$, $f, g: i \to i'$ Morphismen unter A (d.h. $f i = g i = i'$), dann heißt f *homotop unter A* zu g ($f \overset{A}{\simeq} g$), wenn eine Homotopie $\varphi: f \simeq g$ existiert mit $i' q_A = \varphi I i$.

φ heißt dann *Homotopie unter A*.

(1.6) Wenn wir voraussetzen, daß $\simeq$ (bzw. $\underset{B}{\simeq}$, $\overset{A}{\simeq}$) eine Äquivalenzrelation ist, können wir die Quotientkategorie ([13], I.3) $\mathscr{C} h := \mathscr{C}/\simeq$ (bzw. $\mathscr{C}_B h := \mathscr{C}_B/\underset{B}{\simeq}$, $\mathscr{C}^A h := \mathscr{C}^A/\overset{A}{\simeq}$) bilden.

Ein Morphismus f von $\mathscr{C}$ ($\mathscr{C}_B$, $\mathscr{C}^A$) heißt dann *h-Äquivalenz* (*h-Äquivalenz über B, unter A*), wenn seine Klasse $[f]$ ($[f]_B$, $[f]^A$) in $\mathscr{C} h$ ($\mathscr{C}_B h$, $\mathscr{C}^A h$) ein Isomorphismus ist.

(1.7) **Definition.** Ein Morphismus $p: E \to B$ von $\mathscr{C}$ heißt *Faserung* (bzw. *h-Faserung*), wenn zu jedem kommutativen Diagramm in $\mathscr{C}$ der Form

$$\begin{array}{ccc} X & \xrightarrow{f} & E \\ {\scriptstyle j_0 X}\downarrow & & \downarrow{\scriptstyle p} \\ IX & \xrightarrow{\varphi} & B \end{array}$$

ein Morphismus $\Phi: IX \to E$ existiert mit (1) $p\Phi = \varphi$ und (2) $f = \Phi_0$ (bzw. $f \underset{B}{\simeq} \Phi_0$).

Man beachte: (1) impliziert $p\Phi_0 = \varphi_0 = pf$; Φ_0 und f sind daher Morphismen über B von pf nach p.

(1.8) **Definition.** Ein Morphismus $i: A \to X$ von $\mathscr{C}$ heißt *Cofaserung* (bzw. *h-Cofaserung*), wenn zu jedem kommutativen Diagramm in $\mathscr{C}$ der Form

$$\begin{array}{ccc} A & \xrightarrow{j_0 A} & IA \\ {\scriptstyle i}\downarrow & & \downarrow{\scriptstyle \varphi} \\ X & \xrightarrow[f]{} & Y \end{array}$$

ein Morphismus $\Phi: IX \to Y$ existiert mit (1) $\Phi I i = \varphi$ und (2) $f = \Phi_0$ (bzw. $f \overset{A}{\simeq} \Phi_0$).

2. Semikubische Komplexe. Kan-Bedingungen

(2.1) **Definition** (vgl. [11]). Ein *semikubischer Komplex* besteht aus einer Folge von Mengen Q_n ($n \in \mathbb{N}$) und drei Familien von Abbildungen

$$0_n^i, 1_n^i: Q_n \to Q_{n-1} \quad (i, n \in \mathbb{N},\ 1 \leqq i \leqq n) \quad (\textit{Seitenoperatoren}),$$
$$\zeta_n^j: Q_n \to Q_{n+1} \quad (j, n \in \mathbb{N},\ 1 \leqq j \leqq n+1) \quad (\textit{Ausartungsoperatoren}),$$

so daß die folgenden Axiome (A.1)–(A.5) gelten.

(A.1) $\quad \varepsilon_{n-1}^i \omega_n^j = \omega_{n-1}^{j-1} \varepsilon_n^i, \quad i<j,$

(A.2) $\quad \zeta_{n+1}^i \zeta_n^j = \zeta_{n+1}^{j+1} \zeta_n^i, \quad i \leqq j,$

(A.3) $\quad \varepsilon_{n+1}^i \zeta_n^j = \zeta_{n-1}^{j-1} \varepsilon_n^i, \quad i<j,$

(A.4) $\quad \varepsilon_{n+1}^i \zeta_n^i = \mathrm{id}_{Q_n},$

(A.5) $\quad \varepsilon_{n+1}^i \zeta_n^j = \zeta_{n-1}^j \varepsilon_n^{i-1}, \quad i>j,$

wobei $\varepsilon = 0, 1$, $\omega = 0, 1$.

Wir verwenden die folgenden Bezeichnungen:

$$\varphi_\nu := \nu_1^1 \varphi \quad \text{für } \varphi \in Q_1, \quad \nu = 0, 1,$$
$$\psi_\varepsilon^i := \varepsilon_2^i \psi \quad \text{für } \psi \in Q_2, \quad i = 1, 2, \ \varepsilon = 0, 1.$$

Seien Q, Q' semikubische Komplexe. Eine *semikubische Abbildung* $f: Q \to Q'$ ist eine Folge von Abbildungen $f_n: Q_n \to Q'_n$, die mit den Seiten- und Ausartungsoperatoren vertauschbar ist.

Wir erhalten so eine Kategorie, die *Kategorie der semikubischen Komplexe.* Wir bezeichnen sie im folgenden stets mit $\mathscr{K}$.

Für den Rest des Abschnitts sei (n, ν, k) ein Tripel natürlicher Zahlen mit $0 \leqq \nu \leqq 1$ und $1 \leqq k \leqq n$.

(2.2) **Definition.** Eine *(n, ν, k)-Gleichung* γ in einem semikubischen Komplex Q ist eine Familie

$$(\gamma_\varepsilon^i \mid \varepsilon = 0, 1,\ i = 1, \dots, n,\ (\varepsilon, i) \neq (\nu, k)),$$

so daß $\gamma_\varepsilon^i \in Q_{n-1}$ und $\varepsilon_{n-1}^i \gamma_\omega^j = \omega_{n-1}^{j-1} \gamma_\varepsilon^i$ für $i < j$ und (ε, i), $(\omega, j) \neq (\nu, k)$.

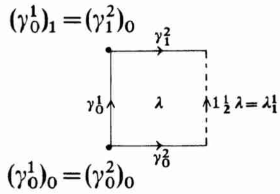
Die Menge der (n, ν, k)-Gleichungen in Q bezeichnen wir mit $Q_{(n, \nu, k)}$.

$\gamma \in Q_{(n, \nu, k)}$ heißt *lösbar*, wenn eine *Lösung* λ von γ existiert, d.h. ein Element $\lambda \in Q_n$, so daß

$$\varepsilon_n^i \lambda = \gamma_\varepsilon^i \quad \text{für alle } (\varepsilon, i) \neq (\nu, k).$$

Wir sagen, Q *erfüllt die Kan-Bedingung* $\mathrm{E}(n, \nu, k)$, wenn jede (n, ν, k)-Gleichung in Q lösbar ist.

Eine $(2, 1, 1)$-Gleichung $\gamma = (\gamma_0^1, \gamma_0^2, \gamma_1^2)$ und eine Lösung λ von γ kann durch die folgende Figur veranschaulicht werden:



(2.3) Wir definieren eine Reihe von Funktoren und natürlichen Transformationen.

$\mathscr{S}$ bezeichne die Kategorie der Mengen, Q einen semikubischen Komplex.

Für $n \in \mathbb{N}$ definieren wir einen kovarianten Funktor $G_n: \mathscr{K} \to \mathscr{S}$ (*n-Gerüst*) durch $G_n(Q) := Q_n$, $G_n(f) := f_n$, wo f eine semikubische Abbildung ist.

Für $n \in \mathbb{N}$ sei $\zeta_{nQ}: Q_0 \to Q_n$ die Abbildung $\zeta_{n-1}^1 \circ \cdots \circ \zeta_2^1 \zeta_1^1 \zeta_0^1$, falls $n \geqq 1$, und id_{Q_0}, falls $n = 0$. Wir erhalten eine natürliche Transformation $\zeta_n: G_0 \to G_n$.

Wir definieren ferner einen kovarianten Funktor $G_{(n, \nu, k)}: \mathscr{K} \to \mathscr{S}$. $G_{(n, \nu, k)}(Q)$ sei die Menge $Q_{(n, \nu, k)}$ der (n, ν, k)-Gleichungen in Q. Ist $f: Q \to Q'$ eine semikubische Abbildung, γ eine (n, ν, k)-Gleichung in Q mit den Komponenten γ_ε^i $((\varepsilon, i) \neq (\nu, k))$, so sei $G_{(n, \nu, k)}(f)(\gamma)$ die (n, ν, k)-Gleichung in Q' mit den Komponenten $f_{n-1} \gamma_\varepsilon^i$.

Schließlich erklären wir eine natürliche Transformation $\zeta_{(n, \nu, k)}: G_0 \to G_{(n, \nu, k)}$. Ist $\varphi \in Q_0$, so sei $\zeta_{(n, \nu, k)Q}(\varphi)$ diejenige (n, ν, k)-Gleichung in Q, deren Komponenten sämtlich $\zeta_{n-1Q}(\varphi)$ sind.

(2.4) **Definition.** Ist Q ein semikubischer Komplex, dann heißt eine Abbildung $\lambda\colon Q_{(n,\nu,k)} \to Q_n$ *Lösungsfunktion*, wenn $\lambda(\gamma)$ für alle $\gamma \in Q_{(n,\nu,k)}$ eine Lösung von γ ist.

Ist λ eine Lösungsfunktion, die das Diagramm

$$\begin{tikzcd} Q_{(n,\nu,k)} \arrow[rr, "\lambda"] & & Q_n \\ & Q_0 \arrow[ul, "\zeta_{(n,\nu,k)Q}"] \arrow[ur, "\zeta_{nQ}"'] & \end{tikzcd}$$

kommutativ macht, sagen wir, λ ist *mit Ausartungen verträglich.*

3. Kan-Bedingungen für Homotopiesysteme

Sei $\mathscr{C}$ eine Kategorie, $\mathbf{I} = (I, j_0, j_1, q)$ ein Homotopiesystem in $\mathscr{C}$. $\mathbf{I}$ induziert einen 1-kontra-, 2-kovarianten Funktor $Q_{\mathbf{I}}\colon \mathscr{C} \times \mathscr{C} \to \mathscr{K}$ in die Kategorie der semikubischen Komplexe.

Durch $I^0 := \mathrm{id}_{\mathscr{C}}$, $I^n := I I^{n-1}$ $(n \geqq 1)$ wird zunächst eine Folge von kovarianten Funktoren $I^n\colon \mathscr{C} \to \mathscr{C}$ definiert. Für $i = 1, \ldots, n$, $j = 1, \ldots, n+1$ erhalten wir natürliche Transformationen

$$0_n^i, 1_n^i\colon I^{n-1} \to I^n, \qquad \zeta_n^j\colon I^{n+1} \to I^n$$

durch $\varepsilon_n^i := I^{i-1} j_\varepsilon I^{n-i}$ $(\varepsilon = 0, 1)$, $\zeta_n^j := I^{j-1} q I^{n+1-j}$.

(3.1) **Definition.** Sind X, Y Objekte von $\mathscr{C}$, so erhalten wir einen semikubischen Komplex $Q_{\mathbf{I}}(X, Y)$ durch die Folge von Mengen $\mathscr{C}(I^n X, Y)$, $n = 0, 1, 2, \ldots$, und die Familien von Abbildungen

$$\mathscr{C}(\varepsilon_{nX}^i, \mathrm{id}_Y)\colon \mathscr{C}(I^n X, Y) \to \mathscr{C}(I^{n-1} X, Y), \qquad \varepsilon = 0, 1,\ i = 1, \ldots, n,$$
$$\mathscr{C}(\zeta_{nX}^j, \mathrm{id}_Y)\colon \mathscr{C}(I^n X, Y) \to \mathscr{C}(I^{n+1} X, Y), \qquad j = 1, \ldots, n+1.$$

Die Axiome (A.1)–(A.5) von (2.1) verifiziert man mit Hilfe von Godement's fünf Regeln über die Zusammensetzung von Funktoren und natürlichen Transformationen ([7], S. 269, 270).

Sind $f\colon X' \to X$, $g\colon Y \to Y'$ Morphismen von $\mathscr{C}$, so liefert die Folge von Abbildungen $\mathscr{C}(I^n f, g)$, $n = 0, 1, 2, \ldots$, eine semikubische Abbildung $Q_{\mathbf{I}}(f, g)\colon Q_{\mathbf{I}}(X, Y) \to Q_{\mathbf{I}}(X', Y')$.

Wir vermerken: $G_0 Q_{\mathbf{I}}\colon \mathscr{C} \times \mathscr{C} \to \mathscr{S}$ ist der 1-kontra-, 2-kovariante Morphismenfunktor $\mathscr{C}(-,-)$.

Wir sind jetzt in der Lage, Kan-Bedingungen für Homotopiesysteme zu definieren.

(3.2) **Definition.** Sei (n, ν, k) ein Tripel natürlicher Zahlen mit $0 \leqq \nu \leqq 1$, $1 \leqq k \leqq n$. Das Homotopiesystem $\mathbf{I}$ *erfüllt die Kan-Bedingung* $\mathrm{E}(n, \nu, k)$, wenn für alle Objekte X, Y von $\mathscr{C}$ der semikubische Komplex $Q_{\mathbf{I}}(X, Y)$ die Kan-Bedingung $\mathrm{E}(n, \nu, k)$ erfüllt (vgl. (2.2)).

I *erfüllt die Kan-Bedingung* NE(n, ν, k), wenn eine natürliche Transformation λ: $G_{(n,\nu,k)}\mathbf{Q_1} \to G_n\mathbf{Q_1}$: $\mathscr{C}\times\mathscr{C}\to\mathscr{S}$ existiert, so daß $\lambda_{(X,Y)}$ für alle Objekte X, Y von $\mathscr{C}$ eine Lösungsfunktion ist.

Ist darüber hinaus jedes $\lambda_{(X,Y)}$ mit Ausartungen verträglich (vgl. (2.4)), so sagen wir, **I** *erfüllt die Kan-Bedingung* DNE(n, ν, k).

Erfüllt ein Homotopiesystem **I** bei festgehaltenem n die Kan-Bedingung E(n, ν, k) (NE(n, ν, k), DNE(n, ν, k)) für alle $\nu = 0, 1$, $k = 1, \ldots, n$, so sagen wir, **I** *erfüllt die Kan-Bedingung* E(n) (NE(n), DNE(n)).

(3.3) **Lemma.** *Erfüllt das Homotopiesystem* **I** *die Kan-Bedingung* E(2), *so ist die Homotopierelation* $\simeq$ (vgl. (1.3)) *eine Äquivalenzrelation.*

Beweis. Symmetrie. Seien f, g: $X \to Y$ Morphismen von $\mathscr{C}$, φ: $IX \to Y$ eine Homotopie von f nach g. Dann erhalten wir durch $\gamma_0^2 := \varphi$, $\gamma_0^1 := \gamma_1^2 := f q_X$ eine (2, 1, 1)-Gleichung γ im semikubischen Komplex $\mathbf{Q_1}(X, Y)$. Da **I** die Kan-Bedingung E(2) erfüllt, existiert eine Lösung $\lambda \in \mathscr{C}(IIX, Y)$ von γ. $\varphi' := \lambda_1^1 = \lambda j_{1IX}$ ist dann eine Homotopie von g nach f.

Transitivität. Gegeben seien Morphismen f, g, h: $X \to Y$ und Homotopien φ, ψ: $IX \to Y$, φ: $f \simeq g$, ψ: $g \simeq h$. Dann liefert $\gamma_0^2 := f q_X$, $\gamma_0^1 := \varphi$, $\gamma_1^2 := \psi$ eine (2, 1, 1)-Gleichung γ in $\mathbf{Q_1}(X, Y)$. Da **I** E(2) erfüllt, können wir eine Lösung λ: $IIX \to Y$ von γ wählen. Durch $\chi := \lambda j_{1IX}$ erhalten wir eine Homotopie von f nach h.

(3.4) **Lemma.** *Erfüllt* **I** *die Kan-Bedingung* NE(2), *so existiert eine natürliche Transformation* r: $I \to I$ *mit* $r j_0 = j_1$, $r j_1 = j_0$.

Beweis. Da **I** NE(2) erfüllt, existiert eine natürliche Transformation λ: $G_{(2,1,1)}\mathbf{Q_1} \to G_2\mathbf{Q_1}$, so daß $\lambda_{(X,Y)}$ für alle Objekte X, Y von $\mathscr{C}$ eine Lösungsfunktion ist. Ist X ein Objekt von $\mathscr{C}$, dann ist $\gamma_X := (\gamma_0^1, \gamma_0^2, \gamma_1^2)$ mit $\gamma_0^1 := \gamma_1^2 := j_{0X} q_X$, $\gamma_0^2 := \mathrm{id}_{IX}$ eine (2, 1, 1)-Gleichung in $\mathbf{Q_1}(X, IX)$. Wir setzen $r_X := (\lambda_{(X,IX)}\gamma_X) j_{1IX}$ und erhalten die gesuchte natürliche Transformation r.

Wir kehren kurz zu Beispiel (1.2) zurück. Da die Vereinigung aller $(n-1)$-dimensionalen Seiten des Würfels $[0, 1]^n$ bis auf eine Retrakt von $[0, 1]^n$ ist, überlegt man leicht:

(3.5) **Satz.** *Das in* (1.2) *definierte Homotopiesystem* **T** *in der Kategorie der topologischen Räume erfüllt die Kan-Bedingung* DNE(n) *für alle* n.

(3.6) Ist B ein Objekt von $\mathscr{C}$, so induziert das Homotopiesystem **I** ein Homotopiesystem $\mathbf{I}_B = (I', j_0', j_1', q')$ in der Kategorie $\mathscr{C}_B$ der Objekte über B. Ist p: $E \to B$ ein Objekt von $\mathscr{C}_B$, so sei $I'p := q_B I p = p q_E$: $IE \to B$. Ist f ein Morphismus über B von p nach p', so ist If ein Morphismus über B von $I'p$ nach $I'p'$. Wir setzen daher $I'f := If$. Die natürlichen Transformationen j_ε': $\mathrm{id}_{\mathscr{C}_B} \to I'$ $(\varepsilon = 0, 1)$, q': $I' \to \mathrm{id}_{\mathscr{C}_B}$ schließlich definieren wir durch $j_{\varepsilon p}' := j_{\varepsilon E}$, $q_p' := q_E$, wenn p: $E \to B$ ein Objekt von $\mathscr{C}_B$ ist.

(3.7) **Satz.** *Erfüllt* **I** *die Kan-Bedingung* DNE(n, ν, k), *so auch* $\mathbf{I}_B$.

Beweis. Wir beschränken uns auf eine Beweisskizze. Nach Voraussetzung existiert eine natürliche Transformation λ: $G_{(n,\nu,k)}\mathbf{Q_1} \to G_n\mathbf{Q_1}$: $\mathscr{C}\times\mathscr{C}\to\mathscr{S}$, so daß $\lambda_{(X,Y)}$ für alle Objekte X, Y von $\mathscr{C}$ eine mit Ausartungen verträgliche Lösungsfunktion ist. Wir haben eine entsprechende natürliche Transformation $\lambda_B\colon G_{(n,\nu,k)}\, Q_{\mathbf{I}_B} \to G_n Q_{\mathbf{I}_B}\colon \mathscr{C}_B \times \mathscr{C}_B \to \mathscr{S}$ zu konstruieren. $p\colon E \to B$, $p'\colon E' \to B$ seien Objekte von $\mathscr{C}_B$. Man überzeugt sich zunächst von den Teilmengenbeziehungen

$$G_{(n,\nu,k)}\, Q_{\mathbf{I}_B}(p, p') \subset G_{(n,\nu,k)}\, Q_{\mathbf{I}}(E, E'),$$

$$G_n Q_{\mathbf{I}_B}(p, p') \subset G_n Q_{\mathbf{I}}(E, E'),$$

rechnet dann nach

$$\lambda_{(E,E')}\, G_{(n,\nu,k)}\, Q_{\mathbf{I}_B}(p, p') \subset G_n Q_{\mathbf{I}_B}(p, p')$$

und definiert $\lambda_{B(p,p')}$ als Einschränkung von $\lambda_{(E,E')}$.

Man überzeugt sich leicht, daß die von $\mathbf{I}_B$ in $\mathscr{C}_B$ nach (1.3) induzierte Relation gerade die Relation $\underset{B}{\simeq}$ aus (1.4) ist. Daher gilt wegen (3.3) und (3.7):

(3.8) **Lemma.** *Erfüllt* $\mathbf{I}$ *die Kan-Bedingung* DNE(2), *so ist die Relation* $\underset{B}{\simeq}$ *eine Äquivalenzrelation.*

Bemerkung. (3.6) ist ein Beispiel dafür, daß der Begriff des Homotopiesystems den Übergang zu anderen Kategorien gestattet. So induziert ein Homotopiesystem in einer Kategorie $\mathscr{C}$ u.a. stets ein Homotopiesystem in der Kategorie der Paare (Morphismenkategorie) und, falls $\mathscr{C}$ cokartesische Quadrate hat, in der Kategorie $\mathscr{C}^A$ der Objekte unter einem Objekt A von $\mathscr{C}$. Im Fall der Kategorie der Paare vererben sich dabei die Kan-Bedingungen NE(n, ν, k) und DNE(n, ν, k); im Fall $\mathscr{C}^A$ überträgt sich DNE(n, ν, k), sofern der Zylinderfunktor des gegebenen Homotopiesystems cokartesische Quadrate erhält (vgl. [8], 2.).

Wir beschließen den Abschnitt mit einem Lemma, das für den Beweis des Homotopiesatzes für h-Faserungen benötigt wird (vgl. (5.5)).

(3.9) **Lemma.** *$\mathscr{C}$ sei eine Kategorie,* $\mathbf{I}$ *ein Homotopiesystem in $\mathscr{C}$ mit* DNE(2). *Ist $p\colon E \to B$ eine h-Faserung, dann läßt sich jedes kommutative Diagramm der Form*

$$\begin{tikzcd} X \arrow[r, "g"] \arrow[d, "j_1 X"'] & E \arrow[d, "p"] \\ IX \arrow[r, "\psi"'] & B \end{tikzcd}$$

durch $\Psi\colon IX \to E$ ergänzen mit $p\Psi = \psi$ und $g \underset{B}{\simeq} \Psi_1$.

Beweis. Wir notieren zunächst, daß $\underset{B}{\simeq}$ wegen der Kan-Bedingung DNE(2) und (3.8) eine Äquivalenzrelation ist. Da $\mathbf{I}$ NE(2) erfüllt, existiert nach (3.4) eine natürliche Transformation $r\colon I \to I$ mit $r j_0 = j_1$, $r j_1 = j_0$. Da $(\psi\, r_X)_0 = \psi_1 = p g$ und da p eine h-Faserung ist, existiert $\Phi\colon IX \to E$ mit $p\Phi = \psi\, r_X$ und $g \underset{B}{\simeq} \Phi_0$. Da $\mathbf{I}$ E(2) erfüllt, existiert $\lambda\colon IIX \to B$ mit

(3.10) $$\lambda j_{0\,IX} = \psi\, r_X\, r_X,$$

(3.11) $$\lambda j_{1IX}=\psi,$$

(3.12) $$\lambda I j_{1X}=\psi_1 q_X=p g q_X.$$

Da p eine h-Faserung ist, existiert wegen (3.10) Λ: $IIX \to E$ mit $p\Lambda=\lambda$ und $\Phi r_X \underset{B}{\simeq} \Lambda j_{0IX}$. Setze $\Psi := \Lambda j_{1IX}$. Wegen (3.11) gilt $p\Psi=\psi$, wegen (3.12) ist $\varphi := \Lambda I j_{1X}$ eine Homotopie über B. Da nun, wie man leicht nachrechnet, $\varphi_0 \underset{B}{\simeq} g$ und $\varphi_1=\Psi_1$, gilt $g \underset{B}{\simeq} \Psi_1$.

4. Kan-Bedingungen und Cofaserungen

$\mathscr{C}$ sei eine Kategorie, $\mathbf{I}=(I, j_0, j_1, q)$ ein Homotopiesystem in $\mathscr{C}$.

Sei

(4.1)
$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "j_{0X}"'] & Y \arrow[d, "k"] \\ IX \arrow[r, "l"'] & Z_f \end{tikzcd}$$

ein cokartesisches Quadrat in $\mathscr{C}$. (Z_f heißt *Abbildungszylinder* von f.)

Da $\mathrm{id}_Y f = f q_X j_{0X}$, existiert genau ein Morphismus p: $Z_f \to Y$ von $\mathscr{C}$ mit $pk=\mathrm{id}_Y$ und $pl=fq_X$.

Damit ist das Diagramm

$$\begin{tikzcd} & X \arrow[dl, "l_1"'] \arrow[dr, "f"] & \\ Z_f \arrow[rr, "p"'] & & Y \end{tikzcd}$$

kommutativ ($l_1=l j_{1X}$).

(4.2) **Satz.** *Erfüllt* $\mathbf{I}$ *die Kan-Bedingung* E(2) *und erhält der Funktor* I *cokartesische Quadrate, so gilt*

(1) $kp \overset{Y}{\simeq} \mathrm{id}_{Z_f}$.

(2) l_1 *ist eine Cofaserung.*

Beweis. (1) Wegen E(2) existiert zunächst ein Morphismus λ: $IIX \to Z_f$ von $\mathscr{C}$ mit $\lambda j_{0IX}=\lambda I j_{0X}=kfq_X=kq_Y If$ und $\lambda j_{1IX}=l$. Da I cokartesische Quadrate erhält, existiert wegen der ersten Gleichung genau ein Morphismus ψ: $IZ_f \to Z_f$ mit $\psi Il=\lambda$ und $\psi Ik=kq_Y$. ψ ist also eine Homotopie unter Y. Da $\psi_0 k=kpk$ und $\psi_0 l=kpl$, wie man leicht nachrechnet, und da (4.1) cokartesisch ist, gilt $\psi_0=kp$; entsprechend erhält man $\psi_1=\mathrm{id}_{Z_f}$.

(2) Gegeben seien g: $Z_f \to V$, φ: $IX \to V$ mit $g l_1=\varphi_0$. Wegen E(2) existiert dann ein Morphismus λ: $IIX \to V$ mit

$$\lambda j_{0IX}=gl, \quad \lambda I j_{0X}=gkq_Y If, \quad \lambda I j_{1X}=\varphi.$$

Da I cokartesische Quadrate erhält, existiert wegen der zweiten Gleichung genau ein Morphismus Φ: $IZ_f \to V$ mit $\Phi Il=\lambda$ und $\Phi Ik=gkq_Y$. Dann gilt

$\Phi I\, l_1 = \lambda I\, j_{1X} = \varphi$ und $\Phi_0 = g$, da $\Phi_0\, l = g\, l$ und $\Phi_0\, k = g\, k$ und da (4.1) cokartesisch ist.

(4.3) Ist

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "a"'] & Y \arrow[d, "b"] \\ X' \arrow[r, "f'"'] & Y' \end{tikzcd}$$

ein kommutatives Diagramm in $\mathscr{C}$ und ist neben (4.1) ein entsprechendes und entsprechend bezeichnetes Diagramm für f' cokartesisch, so bestimmen die Gleichungen $c\,k = k'\,b$, $c\,l = l'\,I\,a$ eindeutig einen Morphismus $c\colon Z_f \to Z_{f'}$. c macht das Diagramm

$$\begin{tikzcd} & X \arrow[dl, "l_1"'] \arrow[dr, "f"] \arrow[dd, "a"] & \\ Z_f \arrow[rr, "p"] \arrow[dd, "c"'] & & Y \arrow[dd, "b"] \\ & X' \arrow[dl, "l_1'"'] \arrow[dr, "f'"] & \\ Z_{f'} \arrow[rr, "p'"'] & & Y' \end{tikzcd}$$

kommutativ. Daher gilt wegen (4.2):

(4.4) **Satz.** *Hat $\mathscr{C}$ cokartesische Quadrate, erhält I cokartesische Quadrate und erfüllt* **I** *die Kan-Bedingung* E(2), *so lassen sich die Morphismen von $\mathscr{C}$ funktoriell in eine Cofaserung, gefolgt von einer h-Äquivalenz zerlegen.*

An Hand der Definitionen beweist man leicht

(4.5) **Lemma.** *Erfüllt* **I** E(2), *so sind j_{0X}, j_{1X} für alle Objekte X von $\mathscr{C}$ Cofaserungen.*

Sei $X \xrightarrow{i_0} X+X \xleftarrow{i_1} X$ ein Coproduktdiagramm in $\mathscr{C}$. $j\colon X+X \to IX$ sei der Morphismus mit den Komponenten j_{0X} und j_{1X}.

(4.6) **Lemma.** *Erfüllt* **I** E(2) *und erhält I Coprodukte, so ist j eine Cofaserung.*

Beweis. Sind $f\colon IX \to Y$, $\varphi\colon I(X+X) \to Y$ Morphismen mit $fj = \varphi_0$, so existiert wegen E(2) $\Phi\colon IIX \to Y$ mit $\Phi\, j_{0IX} = f$ und $\Phi I j_{\nu X} = \varphi I\, i_\nu$ $(\nu = 0, 1)$. Aus der letzten Beziehung folgt $\Phi I j I\, i_\nu = \varphi I\, i_\nu$ für $\nu = 0, 1$ und daher $\Phi I j = \varphi$, da I Coprodukte erhält.

5. Der Homotopiesatz für *h*-Faserungen

Sei $\mathscr{C}$ eine Kategorie mit kartesischen Quadraten, **I** ein Homotopiesystem in $\mathscr{C}$.

$\alpha\colon A \to B$ sei ein Morphismus von $\mathscr{C}$. Wir definieren einen kovarianten Funktor $\alpha^*\colon \mathscr{C}_B \to \mathscr{C}_A$.

Zu jedem Objekt $p: E \to B$ von $\mathscr{C}_B$ wählen wir ein kartesisches Quadrat

(5.1)
$$\begin{tikzcd} E_\alpha \arrow[r, "\tilde{\alpha}"] \arrow[d, "p_\alpha"'] & E \arrow[d, "p"] \\ A \arrow[r, "\alpha"'] & B \end{tikzcd}$$

und setzen $\alpha^* p := p_\alpha$. $p': E' \to B$ sei ein weiteres Objekt von $\mathscr{C}_B$,

(5.1′)
$$\begin{tikzcd} E'_\alpha \arrow[r, "\tilde{\alpha}'"] \arrow[d, "p'_\alpha"'] & E' \arrow[d, "p'"] \\ A \arrow[r, "\alpha"'] & B \end{tikzcd}$$

das zu p' gewählte kartesische Quadrat. $f: p \to p'$ sei ein Morphismus von $\mathscr{C}_B$. Da $p' f \tilde{\alpha} = \alpha p_\alpha$ und da (5.1′) ein kartesisches Quadrat ist, existiert genau ein Morphismus von $\mathscr{C}$ $f_\alpha: E_\alpha \to E'_\alpha$ mit $\tilde{\alpha}' f_\alpha = f \tilde{\alpha}$ und $p'_\alpha f_\alpha = p_\alpha$. Wir setzen $\alpha^* f := f_\alpha: \alpha^* p \to \alpha^* p'$.

Bemerkung. Die Definition von α^* hängt selbstverständlich von der Auswahl der kartesischen Quadrate ab. Verschiedene Auswahlen ergeben jedoch äquivalente Funktoren. Daraus folgt:

(1) $(\mathrm{id}_B)^*$ *ist äquivalent zu* $\mathrm{id}_{\mathscr{C}_B}$.

(2) $(\beta\alpha)^*$ *ist äquivalent zu* $\alpha^* \beta^*$, *wo* $\beta: B \to C$ *ein weiterer Morphismus von* $\mathscr{C}$ *ist.*

(2) beruht auf der Tatsache, daß das Aneinandersetzen von kartesischen Quadraten wieder ein kartesisches Quadrat liefert ([10], 1.5(a)).

(5.2) **Lemma.** *Sind* $f, g: p \to p'$ *Morphismen von* $\mathscr{C}_B$, *so gilt*

$$(f \underset{B}{\simeq} g) \Rightarrow (\alpha^* f \underset{A}{\simeq} \alpha^* g).$$

Beweis. [10], 4.9.

Wir setzen jetzt voraus, $\underset{B}{\simeq}$ und $\underset{A}{\simeq}$ sind Äquivalenzrelationen in $\mathscr{C}_B$ bzw. $\mathscr{C}_A$. Wegen (5.2) induziert α^* einen Funktor der Quotientkategorien $\mathscr{C}_B h \to \mathscr{C}_A h$.

(5.3) **Definition.** $\mathscr{F}_B h$ sei die volle Unterkategorie von $\mathscr{C}_B h$, deren Objekte die h-Faserungen von $\mathscr{C}$ mit Ziel B sind.

Da mit $p: E \to B$ auch $\alpha^* p$ eine h-Faserung ist (*induzierte h-Faserung*) ([10], 4.10), induziert α^* auch einen Funktor $\mathscr{F}_B h \to \mathscr{F}_A h$. Wir bezeichnen diesen Funktor ebenfalls mit α^*.

(5.4) **Satz.** (Homotopiesatz für h-Faserungen.) *$\mathscr{C}$ sei eine Kategorie mit kartesischen Quadraten,* **I** *ein Homotopiesystem in $\mathscr{C}$ mit der Kan-Bedingung* DNE(2). *Sind* $\alpha_0, \alpha_1: A \to B$ *Morphismen von* $\mathscr{C}$, *so daß* $\alpha_0 \simeq \alpha_1$, *so existiert eine natürliche Äquivalenz* $\alpha_0^* \to \alpha_1^*: \mathscr{F}_B h \to \mathscr{F}_A h$.

Der Beweis von Satz (5.4) erfordert einige Vorbereitungen. Für $\nu=0,1$ sei

$$(\mathrm{D}_\nu) \qquad \begin{array}{ccc} E_\nu & \xrightarrow{i_\nu} & E \\ {\scriptstyle p_\nu}\downarrow & & \downarrow{\scriptstyle p} \\ B & \xrightarrow[j_{\nu B}]{} & IB \end{array}$$

ein kommutatives Diagramm in $\mathscr{C}$. Da $qj_\nu=\mathrm{id}$, ist dann auch

$$\begin{tikzcd} E_\nu \arrow[rr, "i_\nu"] \arrow[dr, "p_\nu"'] & & E \arrow[dl, "q_B p"] \\ & B & \end{tikzcd}$$

kommutativ, d.h. i_ν ist ein Morphismus über B von p_ν nach $q_B p$.

(5.5) **Lemma.** *Erfüllt* **I** *die Kan-Bedingung* DNE(2) *und sind* (D_0), (D_1) *kartesische Quadrate, dann sind* i_0 *und* i_1 *h-Äquivalenzen über* B, *falls* p *eine h-Faserung ist.*

Beweis. Wir führen den Beweis für i_0. Da **I** DNE(2) erfüllt, genügt das von **I** in $\mathscr{C}_B$ induzierte Homotopiesystem $\mathbf{I}_B$ der Kan-Bedingung E(2) (3.7).

Mit ihrer Hilfe konstruiert man einen Morphismus $\varphi\colon IIB\to IB$ von $\mathscr{C}$ mit

$$(5.6) \qquad q_B\varphi=q_B q_{IB},$$

$$(5.7) \qquad \varphi j_{0IB}=\mathrm{id}_{IB},$$

$$(5.8) \qquad \varphi I j_{0B}=\varphi j_{1IB}=j_{0B}q_B.$$

Wegen (5.7) ist das äußere Quadrat des Diagramms

$$\begin{tikzcd} E \arrow[r, "\mathrm{id}_E"] \arrow[d, "j_{0E}"'] & E \arrow[d, "p"] \\ IE \arrow[r, "\varphi I p"'] \arrow[ur, dashed, "\Phi"] & IB \end{tikzcd}$$

kommutativ. Da p eine h-Faserung ist, existiert $\Phi\colon IE\to E$ mit $p\Phi=\varphi Ip$ und $\mathrm{id}_E \underset{IB}{\simeq} \Phi_0$. Wegen (5.6) ist Φ eine Homotopie über B. Aus (5.8) folgt $p\Phi_1=j_{0B}q_B p$. Da (D_0) kartesisch ist, existiert daher genau ein Morphismus $r_0\colon E\to E_0$ mit $i_0 r_0=\Phi_1$ und $p_0 r_0=q_B p$. r_0 ist ein Morphismus über B von $q_B p$ nach p_0.

Behauptung. r_0 *ist h-invers über* B *zu* i_0.

Beweis. (1) $i_0 r_0=\Phi_1 \underset{B}{\simeq} \Phi_0 \underset{B}{\simeq} \mathrm{id}_E$.

(2) Zu zeigen bleibt $r_0 i_0 \underset{B}{\simeq} \mathrm{id}_{E_0}$. Mit (5.8) rechnet man $p\Phi I i_0=j_{0B}p_0 q_{E_0}$ nach. Da (D_0) kartesisch ist, existiert daher genau ein Morphismus $\pi\colon IE_0\to E_0$ mit $i_0\pi=\Phi I i_0$ und $p_0\pi=p_0 q_{E_0}$. π ist also eine Homotopie über B. Ferner

gilt für $\nu=0,1$ $i_0\,\pi_\nu=\Phi_\nu\, i_0$ und $p_0\,\pi_\nu=p_0$. Daraus folgt zunächst $\pi_1=r_0\,i_0$, zum andern wegen $\mathrm{id}_E \underset{IB}{\simeq} \Phi_0$ und Lemma (5.2) $\mathrm{id}_{E_0} \underset{B}{\simeq} \pi_0$ und daher schließlich $\mathrm{id}_{E_0} \underset{B}{\simeq} r_0\, i_0$.

Der Beweis für i_1 verläuft ähnlich. Man beachte nur, daß man dabei auf Lemma (3.9) zurückzugreifen hat.

Beweis von Satz (5.4). Sei $\varphi\colon \alpha_0 \simeq \alpha_1$. Da $(\varphi\, j_{\nu A})^*$ äquivalent ist zu $(j_{\nu A})^*\,\varphi^*$, können wir annehmen $\alpha_\nu^*=(j_{\nu A})^*\,\varphi^*$ $(\nu=0,1)$.

Sei $p\colon E\to B$ eine h-Faserung. Wir haben dann ein kommutatives Diagramm mit drei kartesischen Quadraten (0), (1), (2):

$$\begin{tikzcd}
E_{\alpha_0} \arrow[drr, "i_0"] \arrow[drrr, "\tilde\alpha_0"] \arrow[ddd, "\alpha_0^* p"'] \arrow[ddr, phantom, "(0)"] & & & \\
& & E_\varphi \arrow[r, "\tilde\varphi"] \arrow[ddd, "\varphi^* p"] & E \arrow[ddd, "p"] \\
& E_{\alpha_1} \arrow[ur, "i_1"] \arrow[urr, "\tilde\alpha_1"'] \arrow[ddd, "\alpha_1^* p"'] \arrow[r, phantom, "(1)"] & \arrow[r, phantom, "(2)"] & \\
A \arrow[drr, "j_{0A}"'] \arrow[drrr, "\alpha_0"] & & & \\
& & IA \arrow[r, "\varphi"'] & B \\
& A \arrow[ur, "j_{1A}"] \arrow[urr, "\alpha_1"'] & &
\end{tikzcd}$$

Nach Lemma (5.5) sind $[i_0]_A$ und $[i_1]_A$ Isomorphismen von $\mathscr{C}_A h$. Mit $\mu_p := [i_1]_A^{-1}[i_0]_A\colon \alpha_0^* p \to \alpha_1^* p$ erhält man dann eine natürliche Äquivalenz zwischen den Funktoren α_0^* und α_1^*. Damit ist Satz (5.4) bewiesen.

Bemerkung. Fassen wir den eben definierten Morphismus μ_p von $\mathscr{C}_A h$ als Morphismus von $\mathscr{C} h$ auf, $\mu_p\colon E_{\alpha_0}\to E_{\alpha_1}$, so gilt in $\mathscr{C} h$ die Beziehung

$$[\tilde\alpha_1]\,\mu_p=[\tilde\alpha_0]. \tag{5.9}$$

μ_p ist dabei ein Isomorphismus von $\mathscr{C} h$. Aus (5.9) gewinnt man mit denselben formalen Schlüssen wie in [4], (7.30) das folgende Korollar zum Homotopiesatz für h-Faserungen.

(5.10) **Satz.** *Ist $\mathscr{C}$ eine Kategorie mit kartesischen Quadraten,* **I** *ein Homotopiesystem in $\mathscr{C}$ mit der Kan-Bedingung* DNE(2), *so gilt: Ist*

$$\begin{tikzcd}
E_\alpha \arrow[r, "\tilde\alpha"] \arrow[d, "p_\alpha"'] & E \arrow[d, "p"] \\
A \arrow[r, "\alpha"'] & B
\end{tikzcd}$$

ein kartesisches Quadrat in $\mathscr{C}$, p eine h-Faserung und α eine h-Äquivalenz, so ist $\tilde{\alpha}$ eine h-Äquivalenz.

6. h-Faserungen und h-Äquivalenzen über B

(6.1) **Satz** (vgl. [5], 6.1). *Sei $\mathscr{C}$ eine Kategorie, $\mathbf{I}$ ein Homotopiesystem in $\mathscr{C}$ mit den Kan-Bedingungen* DNE(2) *und* E(3).

Ist

$$\begin{tikzcd} E' \arrow[rr, "f"] \arrow[dr, "p'"'] & & E \arrow[dl, "p"] \\ & B & \end{tikzcd}$$

ein kommutatives Diagramm in $\mathscr{C}$, sind p und p' h-Faserungen und ist f eine h-Äquivalenz, so ist f, aufgefaßt als Morphismus über B, eine h-Äquivalenz über B.

Beweis. Wir erinnern zunächst, daß wegen (3.8) $\simeq$ und $\underset{B}{\simeq}$ Äquivalenzrelationen sind. Wir zeigen:

(6.2) *Der Morphismus $[f]_B$ von $\mathscr{C}_B h$ hat ein Rechtsinverses.*

Satz (6.1) ergibt sich dann nach bekanntem Muster durch zweimalige Anwendung von (6.2).

Da f eine h-Äquivalenz ist, können wir zunächst $f''\colon E \to E'$ wählen, so daß $ff'' \simeq \mathrm{id}_E$. Dann gilt $p'f'' = pff'' \simeq p$. Da p' eine h-Faserung ist, existiert nach [10], 4.2 $f'\colon E \to E'$ mit $f' \simeq f''$ und $p'f' = p$. Wir setzen $g := ff'$ und erhalten einen Morphismus über B $g\colon p \to p$ mit $g \simeq \mathrm{id}_E$.

Der wesentliche Schritt besteht nun darin, einen Morphismus über B $g'\colon p \to p$ mit $gg' \underset{B}{\simeq} \mathrm{id}_E$ zu konstruieren. $[f'g']_B$ ist dann nämlich rechtsinvers zu $[f]_B$ in $\mathscr{C}_B h$. Zur Konstruktion von g' wählen wir zunächst eine Homotopie φ von g nach id_E. Dann ist

$$\begin{tikzcd} E \arrow[r, "\mathrm{id}_E"] \arrow[d, "j_{0E}"'] & E \arrow[d, "p"] \\ IE \arrow[r, "p\varphi"'] & B \end{tikzcd}$$

kommutativ. Da p eine h-Faserung ist, existiert $\psi\colon IE \to E$ mit $p\psi = p\varphi$ und $\psi_0 \underset{B}{\simeq} \mathrm{id}_E$. ψ_1 ist ein Morphismus über B von p nach p. Wir setzen $g' := \psi_1$ und haben zu zeigen: $gg' \underset{B}{\simeq} \mathrm{id}_E$. Sei ξ eine Homotopie über B von ψ_0 nach id_E. Da $\mathbf{I}$ die Kan-Bedingung NE(2) erfüllt, können wir mit Hilfe einer natürlichen Transformation $\lambda\colon G_{(2,1,1)} Q_{\mathbf{I}} \to G_2 Q_{\mathbf{I}}$ der Reihe nach die in den folgenden Figuren beschriebenen (2, 1, 1)-Gleichungen lösen.

$$\begin{tikzcd} & & \mathrm{id}_E \\ \cdot \arrow[r, "\varphi"] & \cdot \arrow[ur, dash] & \\ \cdot \arrow[u, "g q_E"] \arrow[r, "g q_E"'] & \cdot \arrow[u, dashed, "\varphi'"'] \arrow[dr, dash] & \\ & & g \end{tikzcd} \qquad \begin{tikzcd} & & g' \\ \cdot \arrow[r, "\psi"] & \cdot \arrow[ur, dash] & \\ \cdot \arrow[u, "\psi_0 q_E"] \arrow[r, "\xi"'] & \cdot \arrow[u, dashed, "\psi'"'] \arrow[dr, dash] & \\ & & \mathrm{id}_E \end{tikzcd} \qquad \begin{tikzcd} & & \mathrm{id}_E \\ \cdot \arrow[r, "\varphi'"] & \cdot \arrow[ur, dash] & \\ \cdot \arrow[u, "g q_E"] \arrow[r, "g \psi'"'] \arrow[ur, phantom, "\alpha"] & \cdot \arrow[u, dashed, "\Phi"'] \arrow[dr, dash] & \\ & & g g' \end{tikzcd}$$

Wir erhalten Morphismen $\varphi', \psi', \Phi: IE \to E$, $\alpha: IIE \to E$. Wegen der Natürlichkeit von λ gilt $p\varphi' = p\psi'$. Wir setzen

$$\gamma_0^2 := p\alpha, \quad \gamma_0^1 := p q_E q_{IE}, \quad \gamma_0^3 := \gamma_1^2 := \gamma_1^3 := p\varphi' I q_E = p\psi' I q_E$$

und erhalten eine (3, 1, 1)-Gleichung γ. Da **I** E(3) erfüllt, existiert eine Lösung $\mu: I^3 E \to B$ von γ. Für $\pi := \mu j_{1\,IIE}: IIE \to B$ gilt dann

(6.3) $$\pi_0^1 = \pi j_{0\,IE} = p\Phi,$$

(6.4) $$\pi_1^1 = \pi_0^2 = \pi_1^2 = p q_E.$$

Da p eine h-Faserung ist, existiert wegen (6.3) $\bar{\pi}: IIE \to E$ mit $p\bar{\pi} = \pi$ und $\Phi \underset{B}{\simeq} \bar{\pi} j_{0\,IE}$. Wegen (6.4) sind $\bar{\pi}_1^1, \bar{\pi}_0^2, \bar{\pi}_1^2$ Homotopien über B. Daher gilt

$$\begin{aligned} g g' &= \Phi_0 \underset{B}{\simeq} (\bar{\pi}_0^1)_0 = (\bar{\pi}_0^2)_0 \underset{B}{\simeq} (\bar{\pi}_0^2)_1 = (\bar{\pi}_1^1)_0 \underset{B}{\simeq} (\bar{\pi}_1^1)_1 \\ &= (\bar{\pi}_1^2)_1 \underset{B}{\simeq} (\bar{\pi}_1^2)_0 = (\bar{\pi}_0^1)_1 \underset{B}{\simeq} \Phi_1 = \mathrm{id}_E \end{aligned}$$

und somit $g g' \underset{B}{\simeq} \mathrm{id}_E$, was noch zu zeigen war.

7. Cohomotopiesysteme

Wir dualisieren die Definitionen von 1. und 3. Dual zum Begriff des Homotopiesystems in einer Kategorie ist der des Cohomotopiesystems.

(7.1) **Definition.** Ein *Cohomotopiesystem* in einer Kategorie $\mathscr{C}$ ist ein Homotopiesystem in der dualen Kategorie $\mathscr{C}^{op}$, d.h. ein Cohomotopiesystem in $\mathscr{C}$ ist ein Quadrupel $\mathbf{W} = (W, p_0, p_1, s)$ aus einem kovarianten Funktor $W: \mathscr{C} \to \mathscr{C}$ und natürlichen Transformationen $p_0, p_1: W \to \mathrm{id}_{\mathscr{C}}$, $s: \mathrm{id}_{\mathscr{C}} \to W$ mit $p_0 s = p_1 s = \mathrm{id}$.

Sei jetzt $\mathbf{W} = (W, p_0, p_1, s)$ ein Cohomotopiesystem in einer Kategorie $\mathscr{C}$.

(7.2) **Definition** (vgl. (1.3)). $f, g: X \to Y$ seien Morphismen von $\mathscr{C}$. f heißt *cohomotop* zu g, wenn eine *Cohomotopie* φ von f nach g existiert, d.h. ein Morphismus $\varphi: X \to WY$ von $\mathscr{C}$ mit $p_{0Y}\varphi = f$ und $p_{1Y}\varphi = g$.

(7.3) **Definition** (vgl. (1.4)). A sei ein Objekt von $\mathscr{C}$, $i: A \to X$, $i': A \to X'$ seien Objekte von $\mathscr{C}^A$, $f, g: i \to i'$ Morphismen unter A. f heißt *cohomotop unter A* zu g, wenn eine Cohomotopie φ von f nach g mit $\varphi i = s_{X'} i'$ existiert.

(7.4) **Definition** (vgl. (1.7)). Ein Morphismus $i\colon A \to X$ von $\mathscr{C}$ heißt *Co-Faserung* (bzw. *Co-h-Faserung*), wenn zu jedem kommutativen Diagramm in $\mathscr{C}$ der Form

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[d, "i"'] & WY \arrow[d, "p_{0Y}"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

ein Morphismus $\Phi\colon X \to WY$ existiert mit $\Phi i = \varphi$ und $f = p_{0Y}\Phi$ (bzw. f cohomotop unter A zu $p_{0Y}\Phi$).

(7.5) Ist B ein Objekt von $\mathscr{C}$, so definiert man dual zu (1.5) in $\mathscr{C}_B$ die Relation *cohomotop über B*. Dual zu (1.8) erklärt man die Begriffe *Co-Cofaserung* und *Co-h-Cofaserung*. Dual zu den Begriffen h-Äquivalenz (über B, unter A) hat man die Begriffe *Co-h-Äquivalenz* (*unter A, über B*).

(7.6) Dual zu (3.1)[1] induziert $\mathbf{W}$ einen 1-kontra-, 2-kovarianten Funktor $Q^{\mathbf{W}}\colon \mathscr{C} \times \mathscr{C} \to \mathscr{K}$ in der Kategorie der semikubischen Komplexe. Sind X, Y Objekte von $\mathscr{C}$, so besteht $Q^{\mathbf{W}}(X, Y)$ aus der Folge von Mengen $\mathscr{C}(X, W^n Y)$, $n=0, 1, 2, \ldots$, und den Familien von Abbildungen $\mathscr{C}(\mathrm{id}_X, \varepsilon^i_{nY})\colon \mathscr{C}(X, W^n Y) \to \mathscr{C}(X, W^{n-1} Y)$, $\mathscr{C}(\mathrm{id}_X, \zeta^j_{nY})\colon \mathscr{C}(X, W^n Y) \to \mathscr{C}(X, W^{n+1} Y)$, wobei $\varepsilon^i_n := W^{i-1} p_\varepsilon W^{n-i}$, $\zeta^j_n := W^{j-1} s W^{n+1-j}$ ($\varepsilon = 0, 1, i = 1, \ldots, n, j = 1, \ldots, n+1$).

Dual zu (3.2) definiert man für $\mathbf{W}$ die Kan-Bedingungen $\mathrm{E}(n, v, k)$, $\mathrm{NE}(n, v, k)$, $\mathrm{DNE}(n, v, k)$, $\mathrm{E}(n)$, $\mathrm{NE}(n)$, $\mathrm{DNE}(n)$. Man hat dann die zu den Sätzen von 3.–6. dualen Sätze. So ist z. B. der folgende Satz dual zu (6.1).

(7.7) **Satz.** *Sei $\mathscr{C}$ eine Kategorie, $\mathbf{W}$ ein Cohomotopiesystem in $\mathscr{C}$ mit den Kan-Bedingungen* DNE(2) *und* E(3). *Ist*

$$\begin{tikzcd} & A \arrow[dl, "i"'] \arrow[dr, "i'"] & \\ X \arrow[rr, "f"'] & & X' \end{tikzcd}$$

ein kommutatives Diagramm in $\mathscr{C}$, sind i und i' Co-h-Faserungen und ist f eine Co-h-Äquivalenz, so ist f, aufgefaßt als Morphismus unter A, eine Co-h-Äquivalenz unter A.

(7.8) Cohomotopiesysteme treten insbesondere dann auf, wenn der Zylinderfunktor eines Homotopiesystems einen adjungierten Funktor hat.

Wir erinnern: Sind $S\colon \mathscr{D} \to \mathscr{C}$, $T\colon \mathscr{C} \to \mathscr{D}$ kovariante Funktoren, so besteht eine *Adjungiertheit* $(\alpha, \beta)\colon S \dashv T$ aus zwei Familien

$$\alpha_{(Y,X)}\colon \mathscr{C}(SY, X) \to \mathscr{D}(Y, TX), \qquad \beta_{(Y,X)}\colon \mathscr{D}(Y, TX) \to \mathscr{C}(SY, X)$$

[1] Unter Dualisierung von (3.1) wird Vertauschen der Faktoren in $\mathscr{C} \times \mathscr{C}$ und Übergang von $\mathscr{C}$ zu $\mathscr{C}^{\mathrm{op}}$ verstanden.

(X ein Objekt von $\mathscr{C}$, Y ein Objekt von $\mathscr{D}$) zueinander inverser Abbildungen, die natürlich in X und Y sind.

Wir lassen die Indizes an α und β im folgenden weg.

Existiert eine Adjungiertheit (α, β): $S \dashv T$, so heißt T *adjungiert* zu S, S *coadjungiert* zu T.

(7.9) $\mathscr{C}$ sei eine Kategorie, $\mathbf{I}=(I, j_0, j_1, q)$ ein Homotopiesystem in $\mathscr{C}$. Zu I existiere ein adjungierter Funktor W: $\mathscr{C} \to \mathscr{C}$. Zu einer Adjungiertheit (α, β): $I \dashv W$ konstruieren wir ein Cohomotopiesystem $\mathbf{I}_\dashv$ in $\mathscr{C}$ (vgl. [6], § 6).

Die Definitionen $p_{vX} := \beta(\mathrm{id}_{WX})\, j_{vWX}$: $WX \to X$, $s_X := \alpha(q_X)$: $X \to WX$ ($v=0,1$, X ein Objekt von $\mathscr{C}$) liefern nach dem zu [13], V. Proposition (2.1) dualen Satz natürliche Transformationen p_0, p_1: $W \to \mathrm{id}_{\mathscr{C}}$, s: $\mathrm{id}_{\mathscr{C}} \to W$, so daß für alle Objekte X, Y von $\mathscr{C}$ die Diagramme

(7.10)
$$\begin{tikzcd} \mathscr{C}(IX, Y) \arrow[r, shift left, "\alpha"] \arrow[d, "j_{vX}^*"'] & \mathscr{C}(X, WY) \arrow[l, shift left, "\beta"] \arrow[dl, "p_{vY*}"] \\ \mathscr{C}(X, Y) & \end{tikzcd},$$

(7.11)
$$\begin{tikzcd} \mathscr{C}(X, Y) \arrow[d, "q_X^*"'] \arrow[dr, "s_{Y*}"] & \\ \mathscr{C}(IX, Y) \arrow[r, shift left, "\alpha"] & \mathscr{C}(X, WY) \arrow[l, shift left, "\beta"] \end{tikzcd}$$

kommutativ sind ($v=0,1$). Dabei sei $j_{vX}^* = \mathscr{C}(j_{vX}, \mathrm{id}_Y)$, $p_{vY*} = \mathscr{C}(\mathrm{id}_X, p_{vY})$, entsprechend q_X^*, s_{Y*}. Man rechnet leicht nach: $p_0 s = p_1 s = \mathrm{id}$. $\mathbf{I}_\dashv := (W, p_0, p_1, s)$ ist also ein Cohomotopiesystem in $\mathscr{C}$.

Dual konstruiert man zu einem Cohomotopiesystem $\mathbf{W}=(W, p_0, p_1, s)$ in $\mathscr{C}$ und einer Adjungiertheit (α, β): $I \dashv W$ ein Homotopiesystem $\mathbf{W}^\dashv$ in $\mathscr{C}$.

Man rechnet aus (bei gegebener Adjungiertheit)

(7.12)
$$(\mathbf{I}_\dashv)^\dashv = \mathbf{I}, \quad (\mathbf{W}^\dashv)_\dashv = \mathbf{W}.$$

(7.13) *Beispiel.* Adjungiert zum Zylinderfunktor des in (1.2) definierten Homotopiesystems $\mathbf{T}$ in der Kategorie der topologischen Räume ist der Wegeraumfunktor, der einem Raum X die Menge der Wege in X, versehen mit der Kompakt-Offen-Topologie zuordnet. Man hat zwischen diesen Funktoren eine kanonische Adjungiertheit ([4], 4.5). Für die nach (7.9) induzierten natürlichen Transformationen gilt: p_{0X} ordnet einem Weg in X den Anfangspunkt, p_{1X} den Endpunkt zu; s_X bildet einen Punkt von X in den konstanten Weg in diesem Punkt ab.

Wir kehren zu $\mathbf{I}_\dashv$ zurück.

(7.14) **Satz.** *Die Begriffe, die in der folgenden Tabelle in einer Zeile stehen, sind äquivalent.*

bezüglich $\mathbf{I}$	bezüglich $\mathbf{I}_{\dashv}$
homotop	cohomotop
homotop unter A	cohomotop unter A
homotop über B	cohomotop über B
h-Äquivalenz	Co-h-Äquivalenz
— unter A, über B	— unter A, über B
Cofaserung	Co-Faserung
h-Cofaserung	Co-h-Faserung
Faserung	Co-Cofaserung
h-Faserung	Co-h-Cofaserung

Beweis. Wir beschränken uns auf die wesentlichen Hinweise. Zu Zeile 1. Ist φ eine Homotopie von f nach g, dann ist $\alpha(\varphi)$ wegen der Kommutativität von (7.10) eine Cohomotopie von f nach g, umgekehrt schließt man entsprechend. Die Äquivalenz der Begriffe in der zweiten Zeile ergibt sich aus der Natürlichkeit von α und β und der Kommutativität von (7.11). Wendet man die zur Äquivalenzaussage der zweiten Zeile duale Aussage auf das Cohomotopiesystem $\mathbf{I}_{\dashv}$ an, so erhält man wegen (7.12) die Äquivalenz der Begriffe in der dritten Zeile. Den Rest überlassen wir dem Leser.

Das folgende Lemma stellt Beziehungen zwischen Kan-Bedingungen für $\mathbf{I}$ und $\mathbf{I}_{\dashv}$ her. α^n und β^n mögen aus α bzw. β durch Iteration entstehen.

(7.15) **Lemma.** *Seien X, Y Objekte von $\mathscr{C}$. Für $\varepsilon = 0, 1, i = 1, \dots, n, j = 1, \dots, n+1$ haben wir folgende kommutative Diagramme:*

$$\begin{tikzcd}
\mathscr{C}(I^n X, Y) \arrow[r, "\varepsilon_n^i"] \arrow[d, shift right, "\alpha^n"'] & \mathscr{C}(I^{n-1} X, Y) \arrow[d, shift right, "\alpha^{n-1}"'] \\
\mathscr{C}(X, W^n Y) \arrow[r, "\varepsilon_n^{n+1-i}"'] \arrow[u, shift right, "\beta^n"'] & \mathscr{C}(X, W^{n-1} Y), \arrow[u, shift right, "\beta^{n-1}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathscr{C}(I^n X, Y) \arrow[r, "\zeta_n^j"] \arrow[d, shift right, "\alpha^n"'] & \mathscr{C}(I^{n+1} X, Y) \arrow[d, shift right, "\alpha^{n+1}"'] \\
\mathscr{C}(X, W^n Y) \arrow[r, "\zeta_n^{n+2-j}"'] \arrow[u, shift right, "\beta^n"'] & \mathscr{C}(X, W^{n+1} Y). \arrow[u, shift right, "\beta^{n+1}"']
\end{tikzcd}$$

In den Zeilen mögen dabei die Seiten- bzw. Ausartungsoperatoren von $Q_{\mathbf{I}}(X, Y)$ bzw. $Q^{\mathbf{I}_{\dashv}}(X, Y)$ stehen.

Da der Beweis von (7.15) aus einer leichten Rechnung besteht, übergehen wir ihn hier.

Aus (7.15) folgert man:

(7.16) **Satz.** $\mathbf{I}$ *erfüllt die Kan-Bedingung* $\mathrm{E}(n, v, k)$ *genau dann, wenn* $\mathbf{I}_{\dashv}$ *die Kan-Bedingung* $\mathrm{E}(n, v, n+1-k)$ *erfüllt* $(0 \leqq v \leqq 1,\ 1 \leqq k \leqq n)$.

Entsprechende Aussagen gelten für die übrigen Kan-Bedingungen. $\mathbf{I}$ erfüllt also genau dann die Kan-Bedingung $\mathrm{E}(n)$, $(\mathrm{NE}(n)$, $\mathrm{DNE}(n))$, wenn $\mathbf{I}_{\dashv}$ $\mathrm{E}(n)$ $(\mathrm{NE}(n)$, $\mathrm{DNE}(n))$ erfüllt.

Wir beschließen den Abschnitt mit zwei Anwendungen.

(7.17) **Satz** (vgl. [4], (2.18)). *Sei $\mathscr{C}$ eine Kategorie, $\mathbf{I}$ ein Homotopiesystem in $\mathscr{C}$, das die Kan-Bedingungen* $\mathrm{DNE}(2)$ *und* $\mathrm{E}(3)$ *erfüllt und zu dessen Zylinderfunktor ein adjungierter Funktor existiert. Ist*

$$\begin{tikzcd} & A \arrow[dl, "i"'] \arrow[dr, "i'"] & \\ X \arrow[rr, "f"'] & & X' \end{tikzcd}$$

ein kommutatives Diagramm in $\mathscr{C}$, sind i und i' h-Cofaserungen und ist f eine h-Äquivalenz, so ist f, aufgefaßt als Morphismus unter A, eine h-Äquivalenz unter A.

Beweis. Man wählt einen Funktor $W: \mathscr{C} \to \mathscr{C}$ und eine Adjungiertheit $(\alpha, \beta): I \dashv W$ und wendet Satz (7.7) auf das Cohomotopiesystem $\mathbf{I}_{\dashv}$ an. Wegen (7.14) und (7.16) folgt (7.17).

Ähnlich folgt aus dem zu (5.10) dualen Satz:

(7.18) **Satz.** *Ist $\mathscr{C}$ eine Kategorie mit cokartesischen Quadraten, $\mathbf{I}$ ein Homotopiesystem in $\mathscr{C}$, das die Kan-Bedingung* $\mathrm{DNE}(2)$ *erfüllt und dessen Zylinderfunktor einen adjungierten Funktor hat, so gilt: Ist*

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "i"'] & A' \arrow[d, "i'"] \\ X \arrow[r, "\xi"'] & X' \end{tikzcd}$$

ein cokartesisches Quadrat in $\mathscr{C}$, i eine h-Cofaserung und α eine h-Äquivalenz, so ist ξ eine h-Äquivalenz.

8. Das Zusammenkleben von h-Äquivalenzen

Das wichtigste Hilfsmittel beim Zusammenkleben von h-Äquivalenzen wird durch das folgende Lemma bereitgestellt.

(8.1) **Lemma** (vgl. [3], Lemma 2). *Sei $\mathscr{C}$ eine Kategorie, $\mathbf{I}$ ein Homotopiesystem in $\mathscr{C}$ mit den Kan-Bedingungen* $\mathrm{DNE}(2)$ *und* $\mathrm{E}(3)$. *Sei*

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] \arrow[d, "h_1"'] & A_0 \arrow[d, "h_0"] & A_2 \arrow[l, "f_2"'] \arrow[d, "h_2"] \\ B_1 \arrow[r, "g_1"'] & B_0 & B_2 \arrow[l, "g_2"] \end{tikzcd}$$

ein kommutatives Diagramm in $\mathscr{C}$, wobei h_0, h_1, h_2 h-Äquivalenzen und f_1, f_2, g_1, g_2 h-Faserungen sind.

Dann existieren h-Äquivalenzen $H_r\colon B_r \to A_r$ $(r=0,1,2)$ mit $f_s H_s = H_0 g_s$ $(s=1,2)$ und Homotopien $\varphi_r\colon IB_r \to B_r$, $\varphi_r\colon h_r H_r \simeq \mathrm{id}_{B_r}$ $(r=0,1,2)$, so daß $g_s \varphi_s = \varphi_0 I g_s$ $(s=1,2)$.

Beweis. Im folgenden sei $s=1,2$. $H_0\colon B_0 \to A_0$ sei h-invers zu h_0, d.h. $h_0 H_0 \simeq \mathrm{id}_{B_0}$ und $H_0 h_0 \simeq \mathrm{id}_{A_0}$. $H'_s\colon B_s \to A_s$ sei h-invers zu h_s. Dann gilt $f_s H'_s \simeq H_0 g_s$. Da f_s eine h-Faserung ist, können wir nach [10], 4.2 H'_s innerhalb seiner Homotopieklasse zu H''_s mit $f_s H''_s = H_0 g_s$ abändern. Wähle $\psi\colon h_0 H_0 \simeq \mathrm{id}_{B_0}$. Dann ist

$$\begin{tikzcd}
B_s \arrow[r, "h_s H''_s"] \arrow[d, "j_{0B_s}"'] & B_s \arrow[d, "g_s"] \\
IB_s \arrow[r, "\psi I g_s"'] & B_0
\end{tikzcd}$$

kommutativ. Da g_s eine h-Faserung ist, läßt sich das Diagramm durch eine Diagonale Φ_s ergänzen, die das rechte untere Dreieck kommutativ, das linke obere Dreieck kommutativ bis auf Homotopie über B_0 macht. Dann ist $k_s := \Phi_s j_{1B_s}$ eine h-Äquivalenz (!), die das Diagramm

$$\begin{tikzcd}
B_s \arrow[rr, "k_s"] \arrow[dr, "g_s"'] & & B_s \arrow[dl, "g_s"] \\
& B_0 &
\end{tikzcd}$$

kommutativ macht. Da g_s eine h-Faserung ist, existiert nach Satz (6.1) ein Morphismus über B_0 $k'_s\colon g_s \to g_s$ mit $k_s k'_s \underset{B_0}{\simeq} \mathrm{id}_{B_s}$. Wir setzen jetzt $H_s := H''_s k'_s$ und erhalten h-Äquivalenzen mit $f_s H_s = H_0 g_s$.

Zu konstruieren bleiben die Homotopien φ_r $(r=0,1,2)$. Wir wählen dazu Homotopien $\rho_s\colon \Phi_s j_{0B_s} \underset{B_0}{\simeq} h_s H''_s$ und $\sigma_s\colon k_s k'_s \underset{B_0}{\simeq} \mathrm{id}_{B_s}$. Da $\mathbf{I}$ NE(2) erfüllt, können wir mit Hilfe einer natürlichen Transformation $\lambda\colon G_{(2,1,1)} Q_{\mathbf{I}} \to G_2 Q_{\mathbf{I}}$ die folgenden (2, 1, 1)-Gleichungen γ_s und γ_0 in $Q_{\mathbf{I}}(B_s, B_s)$ bzw. $Q_{\mathbf{I}}(B_0, B_0)$ lösen:

$$\begin{tikzcd}
{} \arrow[r, "\sigma_s"] & {} & & {} \arrow[r, "q_{B_0}"] & {} \\
{} \arrow[u, "\Phi I k'_s"] \arrow[r, "\rho_s I k'_s"'] & {} \arrow[u, dashed, "\varphi_s"'] & & {} \arrow[u, "\psi"] \arrow[r, "h_0 H_0 q_{B_0}"'] & {} \arrow[u, dashed, "\varphi_0"']
\end{tikzcd}$$

Wir erhalten Homotopien $\varphi_r\colon h_r H_r \simeq \mathrm{id}_{B_r}$ $(r=0,1,2)$. Wendet man g_s auf γ_s an, so ergibt sich dieselbe (2, 1, 1)-Gleichung in $Q_{\mathbf{I}}(B_s, B_0)$ wie aus γ_0 durch Vorschalten von $I g_s$. Wegen der Natürlichkeit von λ gilt daher $g_s \varphi_s = \varphi_0 I g_s$.

(8.2) **Satz.** (Zusammenkleben von h-Äquivalenzen.) *Sei $\mathscr{C}$ eine Kategorie mit cokartesischen Quadraten, $\mathbf{I}$ ein Homotopiesystem in $\mathscr{C}$, das die Kan-Bedingungen*

DNE(2) *und* E(3) *erfüllt und dessen Zylinderfunktor einen adjungierten Funktor hat. In dem kommutativen Diagramm*

(8.3)
$$\begin{tikzcd}
 & A_1 \arrow[dd, "h_1"] \arrow[drr, "f_1'"] & & \\
A_0 \arrow[ur, "f_1"] \arrow[dd, "h_0"'] \arrow[ddrr, "f_2"] & & & A \arrow[dd, "h"] \\
 & B_1 \arrow[drr, "g_1'"] & & \\
B_0 \arrow[ur, "g_1"] \arrow[drr, "g_2"'] & & A_2 \arrow[uur, "f_2'"] \arrow[d, "h_2"] & B \\
 & & B_2 \arrow[ur, "g_2'"'] &
\end{tikzcd}$$

seien das obere und untere Quadrat cokartesisch. Dann gilt: Sind f_1 *und* g_1 *h-Cofaserungen und* h_0, h_1, h_2 *h-Äquivalenzen, so ist h eine h-Äquivalenz.*

Bemerkung. Satz (8.2) verallgemeinert im topologischen Fall [1], 7.5.7 von abgeschlossenen Cofaserungen auf h-Cofaserungen.

Beweis. 1. Wir nehmen zunächst an, daß auch f_2 und g_2 h-Cofaserungen sind. Wir wählen einen Funktor $W: \mathscr{C} \to \mathscr{C}$ und eine Adjungiertheit $I \dashv W$, wenden das zu (8.1) duale Lemma auf das Cohomotopiesystem $\mathbf{I}_{\dashv}$ an (vgl. (7.9)) und erhalten h-Äquivalenzen $H_r: B_r \to A_r$ $(r=0,1,2)$ mit

(8.4) $$H_s g_s = f_s H_0 \qquad (s=1,2)$$

sowie Homotopien $\varphi_r: IA_r \to A_r$ $(r=0,1,2)$, $\varphi_r: H_r h_r \simeq \mathrm{id}_{A_r}$ mit

(8.5) $$\varphi_s I f_s = f_s \varphi_0 \qquad (s=1,2).$$

Da das untere Quadrat von (8.3) cokartesisch ist, existiert wegen (8.4) genau ein Morphismus $H: B \to A$ mit $H g_s' = f_s' H_s$ $(s=1,2)$. Da I einen adjungierten Funktor hat, erhält I cokartesische Quadrate ([13], II.12.1*). Daher existiert wegen (8.5) genau ein Morphismus $\varphi: IA \to A$ mit $\varphi I f_s' = f_s' \varphi_s$ $(s=1,2)$. Man rechnet leicht nach $\varphi: Hh \simeq \mathrm{id}_A$. In (8.3) ersetzen wir jetzt h_0, h_1, h_2, h durch

H_0, H_1, H_2, H, wenden den Schluß noch einmal an und erkennen, daß H und damit h eine h-Äquivalenz ist.

2. Wir haben uns jetzt noch von der Voraussetzung zu befreien, daß f_2' und g_2 h-Cofaserungen sind. Nach Satz (4.4) können wir ein kommutatives Diagramm in $\mathscr{C}$

$$\begin{tikzcd} A_0 \arrow[r, "k"] \arrow[d, "h_0"'] & A' \arrow[r, "u"] \arrow[d, "h'"'] & A_2 \arrow[d, "h_2"] \\ B_0 \arrow[r, "l"'] & B' \arrow[r, "v"'] & B_2 \end{tikzcd}$$

wählen, so daß die obere Zeile f_2, die untere Zeile g_2 ist und u und v h-Äquivalenzen, k und l Cofaserungen und daher h-Cofaserungen sind. Da h_2 eine h-Äquivalenz ist, ist dann auch h' eine h-Äquivalenz. In dem Diagramm

(8.6)
$$\begin{tikzcd} A_1 \arrow[r, "k'"] & A'' \arrow[r, "u'"] & A \\ A_0 \arrow[r, "k"'] \arrow[u, "f_1"] & A' \arrow[r, "u"'] \arrow[u, "a"] & A_2 \arrow[u, "f_2'"'] \end{tikzcd}$$

seien zunächst k' und a so gewählt, daß das linke Teilquadrat cokartesisch ist. u' ist dann eindeutig dadurch bestimmt, daß die obere Zeile f_1 ist und das rechte Teilquadrat kommutativ ist. Entsprechend zerlegt man das untere Quadrat von (8.3) durch Morphismen l', v', b. Da f_1 eine h-Cofaserung ist, ist a nach dem zu [10], 4.10 dualen Satz eine h-Cofaserung. Ferner ist das rechte Teilquadrat von (8.6) nach dem zu [10], 1.5(b) dualen Satz cokartesisch. Also ist u' nach (7.18) eine h-Äquivalenz, entsprechend v'. Durch die Forderungen $h''k'=l'h_1$, $h''a=bh'$ ist eindeutig ein Morphismus $h'': A'' \to B''$ bestimmt. Es gilt $hu'=v'h''$. Nach Teil 1 des Beweises ist h'' eine h-Äquivalenz, also auch h. Das war zu zeigen.

Aus Satz (8.2) folgern wir abschließend:

(8.7) **Satz** (vgl. [2], (1.2)). *Sei $\mathscr{C}$ eine Kategorie mit kartesischen Quadraten, I ein Homotopiesystem in $\mathscr{C}$, das die Kan-Bedingungen* DNE(2) *und* E(3) *erfüllt und dessen Zylinderfunktor einen adjungierten Funktor hat. Gegeben sei ein kommutatives Diagramm* (8.3), *in dem das obere und das untere Quadrat kartesisch sind. Dann gilt: Sind f_1', g_1' h-Faserungen und h, h_1, h_2 h-Äquivalenzen, so ist h_0 eine h-Äquivalenz.*

Beweis. Man wählt einen Funktor $W: \mathscr{C} \to \mathscr{C}$ und eine Adjungiertheit $I \dashv W$ und wendet den zu (8.2) dualen Satz auf das Cohomotopiesystem $\mathbf{I}_\dashv$ an.

Literatur

1. Brown, R.: Elements of modern topology. London: McGraw-Hill 1968.
2. – Heath, P. R.: Coglueing homotopy equivalences. Math. Z. 113, 313–325 (1970).

3. tom Dieck, T.: Partitions of unity in homotopy theory. Compositio Math. **23**, 159–167 (1971).
4. – Kamps, K. H., Puppe, D.: Homotopietheorie. Lecture Notes in Mathematics No. 157. Berlin-Heidelberg-New York: Springer 1970.
5. Dold, A.: Partitions of unity in the theory of fibrations. Ann. of Math. **78**, 223–255 (1963).
6. Eckmann, B., Hilton, P. J.: Group-like structures in general categories II. Math. Ann. **151**, 150–186 (1963).
7. Godement, R.: Topologie algébrique et théorie des faisceaux. Paris: Hermann 1958.
8. Kamps, K. H.: Faserungen und Cofaserungen in Kategorien mit Homotopiesystem. Dissertation, Saarbrücken (1968).
9. – Fibrations and cofibrations in categories with homotopy system. Topol. Appl., Sympos. Herceg-Novi (Yugoslavia) 1968, 211–218 (1969).
10. – Über einige formale Eigenschaften von Faserungen und h-Faserungen. manuscripta math. **3**, 237–255 (1970).
11. Kan, D. M.: Abstract homotopy I. Proc. Nat. Acad. Sci. USA **41**, 1092–1096 (1955).
12. – Abstract homotopy II. Proc. Nat. Acad. Sci. USA **42**, 255–258 (1956).
13. Mitchell, B.: Theory of categories. New York: Academic Press 1965.

Dr. Klaus Heiner Kamps
D-7750 Konstanz
Universität, Fachbereich Mathematik
Deutschland

(Eingegangen am 22. Februar 1971)

Math. Z. 124, 237–249 (1972)

Covering-Avoidance for Saturated Formations of Solvable Lie Algebras*

ERNEST L. STITZINGER

Hallahan and Overbeck in [8] have shown that if L is a Lie algebra of nilpotent length at most two, then a subalgebra C of L is a Cartan subalgebra of L if and only if C covers all central chief factors of L and avoids all eccentric chief factors of L. If L has nilpotent length greater than two, then this result no longer holds. This then leads to the following question: In a solvable Lie algebra L do there exist subalgebras which cover all central chief factors of L and avoid all eccentric chief factors of L? The problem can be stated in a more general way using the idea of saturated formation defined in [1]. These remarks bring to mind the $\mathfrak{F}$-normalizers of Carter and Hawkes in finite solvable groups (see [4]). We shall develop the Lie algebra analogues to the $\mathfrak{F}$-normalizers. In so doing, we shall answer the question raised at the beginning of the paragraph in the affirmative. All Lie algebras considered here are solvable and finite dimensional.

I thank D. W. Barnes for his substantial contribution to this paper. The fundamental definitions of $\mathfrak{F}$-normal and $\mathfrak{F}$-central, along with various of the results, have been suggested by him.

1. Preliminaries

Recall that a class $\mathfrak{F}$ of finite dimensional solvable Lie algebras is called a formation if $\mathfrak{F}$ satisfies the following:

(1) If $L \in \mathfrak{F}$ and N is an ideal of L, then $L/N \in \mathfrak{F}$.

(2) If N and M are ideals of L and L/N, $L/M \in \mathfrak{F}$, then $L/N \cap M \in \mathfrak{F}$.

Let $\mathfrak{F}$ be a formation. A subalgebra H of L is called an $\mathfrak{F}$-projector (or $\mathfrak{F}$-subalgebra) if H satisfies the following:

(1) $H \in \mathfrak{F}$.

(2) If K is a subalgebra of L containing H and K_0 is an ideal of K such that $K/K_0 \in \mathfrak{F}$, then $H + K_0 = K$.

A formation $\mathfrak{F}$ is called saturated if all solvable Lie algebras have $\mathfrak{F}$-projectors. For properties of all these concepts see [1].

Throughout this paper, $\mathfrak{F}$ will denote a saturated formation of solvable Lie algebras. $\mathfrak{N}$ will be the formation of nilpotent Lie algebras and $\mathfrak{U}$ the formation of supersolvable Lie algebras. L will denote a solvable Lie algebra,

* This work was partially supported by a grant from the Engineering Foundation of North Carolina.

$N(L)$ the nil-radical of L and $\Phi(L)$ the Frattini subalgebra of L ($\Phi(L)$ is an ideal of L by Lemma 3.4 of [1]). If M is a maximal subalgebra of L, then $\operatorname{core}_L(M)$ will be the maximal ideal of L contained in M. A subalgebra H of L is said to cover the chief factor U/V of L if $(H+V)\cap U=U$ and said to avoid U/V if $(H+V)\cap U=V$. All formations considered here are non-zero.

Definition. A chief factor U/V of the solvable Lie algebra L is called $\mathfrak{F}$*-central* if the split extension of U/V by $L/C_L(U/V)\in\mathfrak{F}$ and $\mathfrak{F}$-eccentric otherwise.

Note that if any extension of U/V by $L/C_L(U/V)$ is in $\mathfrak{F}$, then the split extension is also in $\mathfrak{F}$ by Lemma 1.16 of [1]. Note that U/V is $\mathfrak{N}$-central if and only if it is central in the usual sense and is $\mathfrak{U}$-central if and only if $\dim U/V=1$.

The question raised in the first paragraph may now be formulated in the following way. Let $\mathfrak{F}$ be a saturated formation. If L is a solvable Lie algebra, then does L contain subalgebras which cover the $\mathfrak{F}$-central chief factors of L and avoid the rest? This will be answered in the affirmative and then some properties of members of this class are considered; in particular, their relations with $\mathfrak{F}$-projectors.

Definition. Let M be a maximal subalgebra of L. M is called $\mathfrak{F}$*-normal* in L if $L/\operatorname{core}_L(M)\in\mathfrak{F}$ and $\mathfrak{F}$-abnormal otherwise.

Note that $L/\operatorname{core}_L(M)$ is a primitive Lie algebra in the sense that $L/\operatorname{core}_L(M)$ contains a minimal ideal which is its own centralizer. A maximal subalgebra M of L is $\mathfrak{N}$-normal if and only if M is an ideal and is $\mathfrak{U}$-normal if and only if M has codimension one in L.

Lemma 1. *A maximal subalgebra M of L is $\mathfrak{F}$-normal if and only if it complements an $\mathfrak{F}$-central chief factor of L.*

Proof. Suppose M is $\mathfrak{F}$-normal. Put $K=\operatorname{core}_L(M)$. Let H/K be the unique minimal ideal of L/K. Then M complements H/K and L/K is the split extension of H/K by $L/C_L(H/K)$. Now $L/K\in\mathfrak{F}$ since M is $\mathfrak{F}$-normal.

Suppose M complements an $\mathfrak{F}$-central chief factor H/K of L. Then $H+M=L$, $M\cap H=K$. Put $C=C_L(H/K)$ and $D=M\cap C$. Now D is an ideal of L and any ideal of L contained in M centralizes H/K, hence $D=\operatorname{core}_L(M)$. Now L/D is the split extension of H/K by L/C. Hence $L/D\in\mathfrak{F}$ and M is $\mathfrak{F}$-normal.

Lemma 2. *$L\in\mathfrak{F}$ if and only if every minimal ideal of $L/\Phi(L)$ is $\mathfrak{F}$-central.*

Proof. Since $L\in\mathfrak{F}$ if and only if $L/\Phi(L)\in\mathfrak{F}$ (Lemma 3.5 of [1]), we may suppose $\Phi(L)=0$. By Theorem 3 of [12], $N(L)=N_1+\cdots+N_r$ where each N_i is a minimal ideal of L. Let $C_i=C_L(N_i)$. Then $N(L)=\bigcap_i C_i$ by Theorem 3 of [12]. Since $\Phi(L)=0$ and $N(L)$ is abelian, $N(L)$ is complemented by a subalgebra, say B. Let $M_i=N_1+\cdots+\hat{N}_i+\cdots+N_r+B$ for $i=1,\ldots,r$. Clearly each M_i is maximal in L. Now $M_i\cap C_i=\operatorname{core}_L(M_i)$ and $L/M_i\cap C_i$ is the split extension of N_i by L/C_i. Thus $L/M_i\cap C_i\in\mathfrak{F}$ if and only if N_i is $\mathfrak{F}$-central. Since $\mathfrak{F}$ is a formation and $\bigcap_i(M_i\cap C_i)=0$, $L\in\mathfrak{F}$ if and only if all N_i are $\mathfrak{F}$-central.

Corollary. *The following are equivalent:*

(1) *Every chief factor of L is $\mathfrak{F}$-central.*

(2) *Every maximal subalgebra of L is $\mathfrak{F}$-normal.*

(3) $L \in \mathfrak{F}$.

(4) $L/\Phi(L) \in \mathfrak{F}$.

Proof. That (1) implies (2) follows from Lemma 1. Suppose that (2) holds. Then every maximal subalgebra of $\bar{L} = L/\Phi(L)$ is $\mathfrak{F}$-normal and each minimal ideal of $\bar{L}$ is complemented by an $\mathfrak{F}$-normal maximal subalgebra. Hence each minimal ideal is $\mathfrak{F}$-central and (3) holds by Lemma 2. Now (3) and (4) are equivalent using Lemma 3.5 of [1]. Assume now that (3) holds and let U/V be a chief factor of L. Since $L/V \in \mathfrak{F}$, the split extension, L^* of U/V by L/U is in $\mathfrak{F}$ by Lemma 1.16 of [1]. Let M be a complement of U/V in L^*. Now $\mathrm{core}_{L^*}(M) = M \cap C_{L^*}(U/V)$ and $L^*/\mathrm{core}_{L^*}(M)$ is the split extension of U/V by $L^*/C_{L^*}(U/V)$ and is also the split extension of U/V by $L/C_L(U/V)$. Since $L^*/\mathrm{core}_{L^*}(M) \in \mathfrak{F}$, (1) holds.

2. $\mathfrak{F}$-Normalizers

Definition. Let M be a maximal subalgebra of L. M will be called *$\mathfrak{F}$-critical* if M is $\mathfrak{F}$-abnormal and $M + N(L) = L$.

Note that if $\bar{L}$ is a homomorphic image of L and $\overline{M} \subset \bar{L}$, then $\overline{M}$ is $\mathfrak{F}$-critical if M is $\mathfrak{F}$-critical.

Definition. A chief factor U/V of L is called *$\mathfrak{F}$-critical* if it is $\mathfrak{F}$-eccentric and all factors below it are either $\mathfrak{F}$-central or uncomplemented.

Lemma 3. *Let M be a maximal subalgebra of L. Then M is $\mathfrak{F}$-critical if and only if M complements an $\mathfrak{F}$-critical chief factor of L.*

Proof. The assumptions and assertions are unaffected by passage to $L/\Phi(L)$, hence we may assume $\Phi(L) = 0$.

Let M be $\mathfrak{F}$-critical and set $N = N(L)$. Then $N(L) = A_1 + \cdots + A_r$ for some minimal ideals A_i of L. Since $M \not\supseteq N$, M complements some A_i, and this A_i is clearly $\mathfrak{F}$-critical.

Suppose conversely that M complements the $\mathfrak{F}$-critical chief factor H/K of L. Then M is $\mathfrak{F}$-abnormal. If $K = 0$, then $H \subseteq N = N(L)$ and so M is $\mathfrak{F}$-critical. Suppose $K \neq 0$ and let $A \subseteq K$ be a minimal ideal of L. Let $N^*/A = N(L/A)$ and put $C = C_L(A)$. Then $N^* \cap C = N$. Suppose M is not $\mathfrak{F}$-critical. Then $M \supseteq N$. By induction, M/A is $\mathfrak{F}$-critical in L/A, so that $M \not\supseteq N^*$. Thus M complements a chief factor between N^* and $N^* \cap C$. Since $A \subseteq K$ and is complemented, A is $\mathfrak{F}$-central and therefore $L/C \in \mathfrak{F}$. Hence all chief factors of L above C are $\mathfrak{F}$-central. But $N^*/N^* \cap C$ and $N^* + C/C$ are operator isomorphic and therefore all chief factors between $N^* \cap C$ and N^* are $\mathfrak{F}$-central. Hence M complements an $\mathfrak{F}$-central chief factor of L contrary to M being $\mathfrak{F}$-abnormal.

Theorem 1. *L has an $\mathfrak{F}$-critical maximal subalgebra if and only if $L \notin \mathfrak{F}$.*

Proof. If $L \in \mathfrak{F}$, then each maximal subalgebra is $\mathfrak{F}$-normal and the result is clear.

Conversely, if $L \notin \mathfrak{F}$, then $\bar{L} = L/\Phi(L)$ has an $\mathfrak{F}$-eccentric minimal ideal, N, by Lemma 2. Then $N/\Phi(L)$ is an $\mathfrak{F}$-critical chief factor and is evidently complemented by a maximal subalgebra M of L. By Lemma 3, M is $\mathfrak{F}$-critical.

Definition. Let L be a solvable Lie algebra. An $\mathfrak{F}$-*normalizer* of L will be a subalgebra H of L such that

(1) there exists a chain $H = H_0 \subset \cdots \subset H_{n+1} = L$ such that H_i is an $\mathfrak{F}$-critical maximal subalgebra of H_{i+1} and

(2) H has no $\mathfrak{F}$-critical maximal subalgebras.

The term $\mathfrak{F}$-normalizer is used because of the analogy with the subgroups of Carter and Hawkes. However, the subalgebras defined here do not appear to normalize anything in particular.

Corollary. *Each solvable Lie algebra has $\mathfrak{F}$-normalizers and these are in $\mathfrak{F}$. In particular, $L \in \mathfrak{F}$ if and only if it is an $\mathfrak{F}$-normalizer of itself.*

Corollary. *If $\bar{L}$ is a homomorphic image of L and H is an $\mathfrak{F}$-normalizer of L, then $\bar{H}$ is an $\mathfrak{F}$-normalizer of $\bar{L}$.*

Lemma 4. *Let M be a maximal subalgebra of L. Then M contains an $\mathfrak{F}$-normalizer of L if and only if M is $\mathfrak{F}$-abnormal in L.*

Proof. Suppose M is $\mathfrak{F}$-abnormal. If M is $\mathfrak{F}$-critical, then trivially M contains an $\mathfrak{F}$-normalizer of L, so suppose M is not $\mathfrak{F}$-critical. Then $M \supseteq N(L)$. Now $\bar{M} = M/\Phi(L)$ is $\mathfrak{F}$-abnormal in $\bar{L} = L/\Phi(L)$ and $\bar{M} \supseteq N(L)/\Phi(L) = N(\bar{L})$. $\bar{L}$ has an $\mathfrak{F}$-eccentric minimal ideal A by Lemma 2, $\bar{M} \supseteq A$ and A has a complement $\bar{V} = V/\Phi(L)$ which is $\mathfrak{F}$-critical in $\bar{L}$. Since $\bar{V} \simeq \bar{L}/A$, $\bar{M} \cap \bar{V}$ is an $\mathfrak{F}$-abnormal maximal subalgebra of $\bar{V}$. Now V is $\mathfrak{F}$-critical in L and $M \cap V$ is an $\mathfrak{F}$-abnormal maximal subalgebra of V. By induction $M \cap V$ contains an $\mathfrak{F}$-normalizer D of V. Now D is also an $\mathfrak{F}$-normalizer of L and $D \subseteq M$.

Conversely, suppose D is an $\mathfrak{F}$-normalizer of L and $D \subseteq M$. Let $K = \mathrm{core}_L(M)$. Then $D+K/K$ is an $\mathfrak{F}$-normalizer of L/K and $D+K \subseteq M \subset L$. Therefore $L/K \notin \mathfrak{F}$.

Theorem 2. *Let K be an $\mathfrak{F}$-normalizer of L. Then K covers the $\mathfrak{F}$-central chief factors of L and avoids the $\mathfrak{F}$-eccentric chief factors of L.*

Proof. We may assume that $L \notin \mathfrak{F}$. Let A/B be a chief factor of L such that $B \neq 0$. Then $K+B/B$ is an $\mathfrak{F}$-normalizer of L/B and the result holds by induction. Hence we may suppose that A is a minimal ideal of L. Let $K = K_0 \subset \cdots \subset K_{n+1} = L$ be a chain which makes K an $\mathfrak{F}$-normalizer of L. Suppose $K_n + A = L$, hence $K_n \cap A = 0$. Then K avoids A. Since K_n is $\mathfrak{F}$-abnormal, A is $\mathfrak{F}$-eccentric and the result holds in this case.

Suppose that $K_n + A \subset L$, hence $A \subset K_n$. Since $K_n + N(L) = L$ and $C_L(A) \supseteq N(L)$, A is a minimal ideal of K_n and is $\mathfrak{F}$-central in K_n if and only if it is $\mathfrak{F}$-central in L. Since K is an $\mathfrak{F}$-normalizer of K_n, the result follows by induction.

Corollary. *Let K be an $\mathfrak{F}$-normalizer of L. Then the dimension of K is the sum of the dimensions of the $\mathfrak{F}$-central chief factors in any given chief series of L. In particular, if K and J are both $\mathfrak{F}$-normalizers of L, then* $\dim K = \dim J$.

Note that if U is a $\mathfrak{U}$-normalizer of L, then $\dim U$ is the number of one-dimensional factors in any given chief series of L.

In general, if K and J are $\mathfrak{F}$-normalizers of L, they need not be conjugate. Also there exist Lie algebras L which have subalgebras which satisfy the covering-avoidance property but are not $\mathfrak{F}$-normalizers of L. Examples of these situations appear in the final section of the paper.

This covering-avoidance property may be used to show the next result. For groups, see Lemma 3.1 of [11].

Corollary. *Let D be an $\mathfrak{F}$-normalizer of L and H and K be ideals of L. Then*
(1) $(D+H)\cap(D+K)=D+(H\cap K)$ *and*
(2) $(D\cap H)+(D\cap K)=D\cap(H+K)$.

Proof. To show (1) we may assume that $H\neq 0$, $K\neq 0$ and $H\cap K=0$. We use induction on $\dim L$. Let H_1 be a minimal ideal of L contained in H. Then since $D+H_1/H_1$ is an $\mathfrak{F}$-normalizer of L/H_1, the induction hypothesis implies that $(D+H/H_1)\cap(D+H_1+K/H_1)=D+H_1/H_1$ so that $(D+H)\cap(D+K)=(D+H_1)\cap(D+K)$. Hence we may suppose that $H=H_1$; that is, H is a minimal ideal of L. Also we may suppose that K is a minimal ideal of L.

By the covering property of D, we may suppose that H and K are $\mathfrak{F}$-eccentric chief factors of L. Then $H+K/K$ is also an $\mathfrak{F}$-eccentric chief factor of L, and hence, by the avoidance property of D, $D\cap(H+K)=0$. Suppose $x\in(D+H)\cap(D+K)$. Then $x=d_1+h=d_2+k$ with $d_1, d_2\in D$, $h\in H$ and $k\in K$. Then $d_1-d_2=k-h\in D\cap(H+K)=0$. Hence $h=k$ and $x\in D$. Therefore $(D+H)\cap(D+K)=D$.

Now (2) follows from (1) and the Lie algebra version of Hilfssatz 2.5 of [10].

3. Two Characterizations of $\mathfrak{F}$-Projectors

In this section we find two characterizations of $\mathfrak{F}$-projectors which will be of use in the next section.

Lemma 5. *The $\mathfrak{F}$-projectors K of L can be characterized by the following conditions:*

(1) $K\in\mathfrak{F}$.

(2) *Every link in every maximal chain joining K to L is $\mathfrak{F}$-abnormal.*

Proof. Induct on $\dim L$. Let D be an $\mathfrak{F}$-projector of L and $D\subseteq V\subset U\subseteq L$ where V is a maximal subalgebra of U. D is an $\mathfrak{F}$-projector of U so that if $U\subset L$, then V is $\mathfrak{F}$-abnormal in U by induction. Hence assume $U=L$ and V is a maximal subalgebra of L. Let $K=\operatorname{core}_L(V)$ and H/K be the minimal ideal of L/K. Suppose V is $\mathfrak{F}$-normal in L. Then $L/K\in\mathfrak{F}$. Therefore $L=D+K$ by the definition of D as an $\mathfrak{F}$-projector of L. But $K+D\subseteq V\subset L$, a contradiction. Hence V is $\mathfrak{F}$-abnormal in L.

Conversely, suppose $D\in\mathfrak{F}$ and that every link in every maximal chain from D to L is $\mathfrak{F}$-abnormal. Let $D\subseteq S\subseteq L$ and S_0 be an ideal of S such that $S/S_0\in\mathfrak{F}$. If $S\subset L$, then D is an $\mathfrak{F}$-projector of S by induction, hence $S=S_0+D$. Suppose $S_0+D\subset S=L$ and let M be a maximal subalgebra of L containing S_0+D.

Since $M \supseteq D$, M is $\mathfrak{F}$-abnormal in L. But $L/S_0 \in \mathfrak{F}$, hence all maximal subalgebras of L/S_0 are $\mathfrak{F}$-normal. Hence M/S_0 is $\mathfrak{F}$-normal in L/S_0, so that M is $\mathfrak{F}$-normal in L and we have a contradiction.

Definition. A maximal subalgebra M of L will be called $\mathfrak{F}$-*crucial* if

(1) *M is $\mathfrak{F}$-abnormal in L and*

(2) $M/\mathrm{core}_L(M) \in \mathfrak{F}$.

Lemma 6. *L has an $\mathfrak{F}$-crucial maximal subalgebra if and only if $L \notin \mathfrak{F}$. If $L \notin \mathfrak{F}$ and C is an $\mathfrak{F}$-projector of L, then L has an $\mathfrak{F}$-crucial maximal subalgebra which contains C.*

Proof. The result holds if $L \in \mathfrak{F}$. Suppose $L \notin \mathfrak{F}$. Let C be an $\mathfrak{F}$-projector of L and let M be a maximal subalgebra of L containing C. Put $K = \mathrm{core}_L(M)$ and let N/K be the minimal ideal of L/K. Then $L/K \notin \mathfrak{F}$ since $C + K \subseteq M \subset L$. If $L/N \notin \mathfrak{F}$, then by induction, $C + N/N$ is contained in an $\mathfrak{F}$-crucial maximal subalgebra U/N of L/N and U is $\mathfrak{F}$-crucial in L. If $L/N \in \mathfrak{F}$, then $M/K \in \mathfrak{F}$ and M is $\mathfrak{F}$-crucial.

Lemma 7. *Let L be a solvable Lie algebra and A be a subalgebra of L. Then A is an $\mathfrak{F}$-projector of L if and only if there exists a chain $A = A_0 \subset \cdots \subset A_{n+1} = L$ in which A_i is $\mathfrak{F}$-crucial in A_{i+1} and A contains no $\mathfrak{F}$-crucial subalgebras.*

Proof. Suppose the condition holds. Then $A \in \mathfrak{F}$ by Lemma 6. If $A = L$, then the result is clear. Suppose $A \subset L$. Now A is $\mathfrak{F}$-crucial in A_1, hence by Lemma 5, A is an $\mathfrak{F}$-projector of A_1. If $A_1 = L$, then we are done. Suppose $A_1 \subset L$. Let $K = \mathrm{core}_{A_2}(A_1)$. Then $A_1/K \in \mathfrak{F}$ and A_1/K is $\mathfrak{F}$-abnormal in A_2/K so that A_1/K is an $\mathfrak{F}$-projector of A_2/K. Hence, by Lemma 1.8 of [1], A is an $\mathfrak{F}$-projector of A_2. We may continue this procedure up the chain to see that A is an $\mathfrak{F}$-projector of L.

Conversely, suppose A is an $\mathfrak{F}$-projector of L. If $L \in \mathfrak{F}$, then the result is trivial. If $L \notin \mathfrak{F}$, then, by Lemma 6, there exists an $\mathfrak{F}$-abnormal maximal subalgebra M of L which contains A and $M/\mathrm{core}_L(M) \in \mathfrak{F}$. A is an $\mathfrak{F}$-projector of M so we may repeat this procedure in M until we reach a subalgebra T of L, $A \subseteq T$ and $T \in \mathfrak{F}$. Since A is an $\mathfrak{F}$-projector of T, $A = T$ and the result holds.

4. $\mathfrak{F}$-Projectors and $\mathfrak{F}$-Normalizers

Denote the class of solvable Lie algebras L which have a nilpotent ideal whose factor is in $\mathfrak{F}$ by $\mathfrak{N}\mathfrak{F}$.

Theorem 3. *If $L \in \mathfrak{N}\mathfrak{F}$, then the $\mathfrak{F}$-projectors coincide with the $\mathfrak{F}$-normalizers.*

Proof. If $L \in \mathfrak{F}$, then the result is clear. Suppose $L \notin \mathfrak{F}$. Then L has an $\mathfrak{F}$-abnormal maximal subalgebra M and clearly $L = N(L) + M$. Hence every $\mathfrak{F}$-abnormal maximal subalgebra of L is $\mathfrak{F}$-critical. Also $N(L) \cap M$ is an ideal in M and not self-normalizing in $N(L)$, hence $N(L) \cap M$ is an ideal in L. Then $N(L) \cap M \subseteq \mathrm{core}_L(M)$. Hence $M/\mathrm{core}_L(M) \in \mathfrak{F}$. Therefore each $\mathfrak{F}$-abnormal maximal subalgebra of L is $\mathfrak{F}$-crucial. Consequently, the $\mathfrak{F}$-crucial and $\mathfrak{F}$-

critical maximal subalgebras of L coincide. Hence the $\mathfrak{F}$-projectors and $\mathfrak{F}$-normalizers of L coincide using Lemma 7.

We note that if L has nilpotent length at most two and $\mathfrak{F}=\mathfrak{N}$, then Theorem 2, Theorem 3 and the fact that $\mathfrak{F}$-projectors are precisely Cartan subalgebras shows that Cartan subalgebras have the cover-avoidance property shown in [8]. Also, the condition $L\in\mathfrak{N}\mathfrak{F}$ can not be removed as is seen in the example in the final section.

Corollary. *Let C be an $\mathfrak{F}$-projector of L and H, K be ideals of L. Then*

(1) $(C+H)\cap(C+K)=C+(H\cap K)$ and

(2) $(C\cap H)+(C\cap K)=C\cap(H+K)$.

Proof. To show (1) we may assume $H\supset 0$, $K\supset 0$ and $H\cap K=0$. Induct on $\dim L$. Let H_1 be a minimal ideal of L contained in H. Then $C+H_1/H_1$ is an $\mathfrak{F}$-projector of L/H_1, hence the induction hypothesis yields

$$(C+H)/H_1\cap(C+H_1+K/H_1)=C+H_1/H_1$$

so that

$$(C+H)\cap(C+H_1+K)=C+H_1.$$

Similarly, if K_1 is a minimal ideal of L contained in K, then

$$(C+K)\cap(C+H+K_1)=C+K_1.$$

Hence,

$$(C+H)\cap(C+K)=(C+H_1)\cap(C+K_1).$$

Now H_1+K_1 is an abelian ideal of L and $C\in\mathfrak{F}$, hence $H_1+K_1+C\in\mathfrak{N}\mathfrak{F}$. Also, C is an $\mathfrak{F}$-projector of H_1+K_1+C. Hence C is an $\mathfrak{F}$-normalizer of H_1+K_1+C. Therefore, the Corollary to Theorem 2 yields

$$(C+H_1)\cap(C+K_1)=C+(H_1\cap K_1)=C.$$

Hence,

$$(C+H)\cap(C+K)=C.$$

Therefore (1) holds.

Now (2) follows from (1) and Hilfssatz 2.5 of [10].

Theorem 4. *Let $L\in\mathfrak{N}\mathfrak{F}$ and H a subalgebra of L which covers all $\mathfrak{F}$-central chief factors of L in some chief series of L. Then H contains an $\mathfrak{F}$-normalizer of L. In particular, if H also avoids the $\mathfrak{F}$-eccentric factors in the chief series, then H is an $\mathfrak{F}$-normalizer of L.*

Proof. If $H=L$, then the result is clear. Assume that $L\notin\mathfrak{F}$ and $H\subset L$. Let M be a maximal subalgebra of L containing H. Since H covers all $\mathfrak{F}$-central factors in the given chief series, M complements an $\mathfrak{F}$-eccentric chief factor of L. Hence M is $\mathfrak{F}$-abnormal in L and so M is $\mathfrak{F}$-critical since clearly $M+N(L)=L$. Now in the given series, M avoids one $\mathfrak{F}$-eccentric factor contained in $N(L)$ and covers the remaining factors. Let A/B be a chief factor of L covered by M. Since $N(L)$ centralizes A/B,

$$M/C_M(A/B)=M/C_L(A/B)\cap M\simeq M+C_L(A/B)/C_L(A/B)=L/C_L(A/B)$$

and the split extensions of A/B by $L/C_L(A/B)$ and A/B by $M/C_M(A/B)$ are seen to be isomorphic. Furthermore, $A=(A\cap M)+B$, hence A/B and $A\cap M/B\cap M$ are operator isomorphic. Hence $C_M(A/B)=C_M(A\cap M/B\cap M)$ and the split extensions of A/B by $M/C_M(A/B)$ and $A\cap M/B\cap M$ by $M/C_M(A\cap M/B\cap M)$ are also isomorphic. Hence $A\cap M/B\cap M$ is a chief factor of M and is $\mathfrak{F}$-central in M if and only if A/B is $\mathfrak{F}$-central in L. Therefore H covers all $\mathfrak{F}$-central chief factors of M in the chief series induced in M by the given chief series of L. By induction H contains an $\mathfrak{F}$-normalizer of M. But M is $\mathfrak{F}$-critical, hence H contains an $\mathfrak{F}$-normalizer of L. If H also avoids the $\mathfrak{F}$-eccentric chief factors of L, then H has the same dimension as the $\mathfrak{F}$-normalizers of L and so must be one of them.

It is noted that if L has nilpotent length at most two, $\mathfrak{F}=\mathfrak{N}$ and H is a subalgebra of L which has the cover-avoidance property, then the above two results show that H must be a Cartan subalgebra of L.

Lemma 8. *Let $L\in\mathfrak{N}\mathfrak{F}$, H a subalgebra of L such that $H\in\mathfrak{F}$ and $L=N(L)+H$. Then $N_L(H)$ is contained in some $\mathfrak{F}$-projector of L.*

Proof. The result is clear if $L\in\mathfrak{F}$. Suppose $L\notin\mathfrak{F}$ and induct on $\dim L$. Let N be a minimal ideal of L. Then $N+H/N\in\mathfrak{F}$ and $N+H/N+N(L/N)=L/N$, hence $N_{L/N}(N+H/N)$ is contained in an $\mathfrak{F}$-projector B/N of L/N and this is of the form $N+E/N$ where E is an $\mathfrak{F}$-projector of B and hence of L by Lemma 1.8 of [1]. Therefore $N_L(H)\subseteq N_L(N+H)\subseteq N+E$. Suppose $N+E\subset L$. Then H satisfies the hypothesis when considered as a subalgebra of $N+E$. Then $N_L(H)$ is contained in an $\mathfrak{F}$-projector D of $N+E$ by induction. Since $N+E/N$ is an $\mathfrak{F}$-projector of L/N, D is an $\mathfrak{F}$-projector of L. Hence we may assume that $L=N+E$ and $L/N\in\mathfrak{F}$. Further, we may assume $L/N\in\mathfrak{F}$ for every minimal ideal N of L, and so need only consider the case where N is the only minimal ideal of L. But then L is primitive, $N=N(L)$ and $H+N(L)=L$ implies that H complements N in L, $H=N_L(H)$ and H is an $\mathfrak{F}$-projector of L.

Theorem 5. *Let K be a subalgebra of L such that $L=N(L)+K$. Then every $\mathfrak{F}$-projector of K is of the form $K\cap E$ where E is an $\mathfrak{F}$-projector of L.*

Proof. K and L induce isomorphic derivation algebras on each chief factor of L since $N(L)+K=L$ and $N(L)$ centralizes each chief factor of L. As in the proof of Theorem 4, if A/B is a chief factor of L covered by K, then $A\cap K/B\cap K$ is a chief factor of K, $A/B\simeq A\cap K/B\cap K$, $L/C_L(A/B)\simeq K/C_K(A\cap K/B\cap K)$ and the corresponding split extensions are isomorphic.

Suppose $L\in\mathfrak{N}\mathfrak{F}$. Then $K/K\cap N(L)\simeq L/N(L)\in\mathfrak{F}$. Hence any $\mathfrak{F}$-projector E_1 of K satisfies $(K\cap N(L))+E_1=K$. Therefore $L=N(L)+E_1$. By Lemma 8, $E_1\subseteq K\cap E$ for some $\mathfrak{F}$-projector E of L. Now K intersects an $\mathfrak{F}$-central chief factor of L which it covers in an $\mathfrak{F}$-central chief factor of K. Let $0=N_0\subset N_1\subset\cdots\subset N_k=L$ be a chief series of L. Then K either covers or avoids any chief factor. Since E covers the $\mathfrak{F}$-central chief factors and avoids the rest, $\dim(E\cap K)\leqq$ sum of the dimensions of the $\mathfrak{F}$-central chief factors in the series which K covers. But if N_i/N_{i-1} is one of these, then $N_i\cap K/N_{i-1}\cap K$ is an $\mathfrak{F}$-central chief factor of K and $N_i \cap K/N_{i-1} \cap K \simeq N_i/N_{i-1}$. Then dim E_1 is the sum of the dimensions of the $\mathfrak{F}$-central chief factors of K in the chief series induced in K by the given chief series of L. Hence $\dim(E \cap K) \leqq \dim E_1$ and $E \cap K = E_1$.

Suppose $L \notin \mathfrak{N}\mathfrak{F}$. Let C be an $\mathfrak{F}$-projector of K. Now $C + N(L)/N(L)$ is an $\mathfrak{F}$-projector of $K + N(L)/N(L) = L/N(L)$ but is properly contained in $L/N(L)$. Hence there exists an $\mathfrak{F}$-crucial maximal subalgebra $M/N(L)$ which contains $C + N(L)/N(L)$ and M is an $\mathfrak{F}$-crucial maximal subalgebra of L which contains $C + N(L)$. C is an $\mathfrak{F}$-projector of $M \cap K$. Since $N(M) \supseteq N(L)$, $M = N(M) + (M \cap K)$. By induction on M, there exists an $\mathfrak{F}$-projector E of M such that $E \cap (M \cap K) = C$. But M is $\mathfrak{F}$-crucial, hence E is an $\mathfrak{F}$-projector of L and $E \cap (M \cap K) = E \cap K = C$.

We are now able to find a relation between $\mathfrak{F}$-normalizers and $\mathfrak{F}$-projectors.

Lemma 9. *Let $\mathfrak{F}$ and $\mathfrak{F}^*$ be saturated formations. Suppose D is an $\mathfrak{F}$-normalizer of L. Then every $\mathfrak{F}^*$-projector of D is the intersection with D of an $\mathfrak{F}^*$-projector of L.*

Proof. Join D to L by an $\mathfrak{F}$-critical chain and work up applying Theorem 5.

Theorem 6. *Each $\mathfrak{F}$-normalizer of L is contained in an $\mathfrak{F}$-projector of L.*

Proof. Follows immediately from Lemma 9.

Lemma 10. *Let $\mathfrak{F}$ and $\mathfrak{F}^*$ be saturated formations such that $\mathfrak{F} \subseteq \mathfrak{F}^*$, and let D be an $\mathfrak{F}^*$-normalizer of L and E be an $\mathfrak{F}$-normalizer of D. Then E is an $\mathfrak{F}$-normalizer of L.*

Proof. An $\mathfrak{F}^*$-critical maximal subalgebra is $\mathfrak{F}$-critical.

Theorem 7. *If U is a $\mathfrak{U}$-normalizer of L, then there exists a Cartan subalgebra C of L such that $U \cap C$ is an $\mathfrak{N}$-normalizer of L.*

Proof. Let D be a Cartan subalgebra of U. Then $D = U \cap C$ for some Cartan subalgebra C of L by Lemma 9. But $U \in \mathfrak{U} \subseteq \mathfrak{N}\mathfrak{N}$ and therefore the $\mathfrak{N}$-projectors and $\mathfrak{N}$-normalizers of U coinside. Therefore D is an $\mathfrak{N}$-normalizer of U and hence D is an $\mathfrak{N}$-normalizer of L by Lemma 10.

Lemma 11. *Suppose $\mathfrak{F}$ and $\mathfrak{F}^*$ are saturated formations, that $\mathfrak{F} \subseteq \mathfrak{F}^*$ and that the $\mathfrak{F}$- and $\mathfrak{F}^*$-projectors of L coincide. Then a chief factor A/B is $\mathfrak{F}$-central if and only if it is $\mathfrak{F}^*$-central.*

Proof. If A/B is $\mathfrak{F}$-central, it is trivially $\mathfrak{F}^*$-central. Let L be a minimal counterexample and let A be a minimal ideal of L. The $\mathfrak{F}$- and $\mathfrak{F}^*$-projectors of L/A coincide, so every $\mathfrak{F}^*$-central factor of L/A is also $\mathfrak{F}$-central. Since L is a counterexample, A is $\mathfrak{F}^*$-central but not $\mathfrak{F}$-central. Since this holds for every minimal ideal, L has only one minimal ideal. Let $C = C_L(A)$ and let U/A be an $\mathfrak{F}^*$-projector of L/A, and so also an $\mathfrak{F}$-projector of L/A. Since A is $\mathfrak{F}^*$-central, $L/C \in \mathfrak{F}^*$ and therefore $C + U = L$. It follows that A is a minimal ideal of U and the split extension of A by L/C is the split extension of A by $U/C_U(A)$. Thus A is $\mathfrak{F}$-(respectively $\mathfrak{F}^*$-)central in L if and only if it is $\mathfrak{F}$-($\mathfrak{F}^*$-)central in U. A subalgebra V of U is an $\mathfrak{F}$-($\mathfrak{F}^*$-)projector of U if and only if it is an $\mathfrak{F}$-

($\mathfrak{F}^*$-)projector of L. Thus U satisfies the conditions of the lemma and the result follows by induction unless $U=L$. Thus we have $L/A\in\mathfrak{F}\subseteq\mathfrak{F}^*$. Since A is $\mathfrak{F}^*$-central, $L\in\mathfrak{F}^*$ and is an $\mathfrak{F}^*$-projector of L. Therefore L is an $\mathfrak{F}$-projector of L and $L\in\mathfrak{F}$.

Corollary. *If $\mathfrak{F}$, $\mathfrak{F}^*$ are saturated formations and the $\mathfrak{F}$-, $\mathfrak{F}^*$-projectors of L coincide, then a chief factor of L is $\mathfrak{F}$-central if and only if it is $\mathfrak{F}^*$-central.*

Proof. Here the $\mathfrak{F}\cap\mathfrak{F}^*$-projectors of L coincide with the $\mathfrak{F}$- and $\mathfrak{F}^*$-projectors of L. Hence the result follows from Lemma 11 applied to $\mathfrak{F}\cap\mathfrak{F}^*\subseteq\mathfrak{F}$ and $\mathfrak{F}\cap\mathfrak{F}^*\subseteq\mathfrak{F}^*$.

Theorem 8. *Suppose the $\mathfrak{F}$- and $\mathfrak{F}^*$-projectors of L coincide. Then the $\mathfrak{F}$- and $\mathfrak{F}^*$-normalizers of L coincide.*

Proof. By the above Corollary, it is sufficient to show that if L has a chief series in which the $\mathfrak{F}$- and $\mathfrak{F}^*$-central chief factors coincide, then the $\mathfrak{F}$- and $\mathfrak{F}^*$-normalizers of L coincide. Suppose then that L is a minimal counterexample, that D is an $\mathfrak{F}$-normalizer of L and that M is an $\mathfrak{F}$-critical maximal subalgebra of L such that $D\subseteq M$. Then M complements an $\mathfrak{F}$-eccentric chief factor in the given series and it must be $\mathfrak{F}^*$-eccentric also. Hence M is $\mathfrak{F}^*$-critical. As in the proof of Theorem 4, the chief series induces a chief series in M which satisfies the hypothesis. The result can now be completed by induction.

Let $\mathfrak{N}\mathfrak{N}\mathfrak{F}$ be the collection of all solvable Lie algebras L which have a nilpotent ideal N such that $L/N\in\mathfrak{N}\mathfrak{F}$.

Theorem 9. *Let $L\in\mathfrak{N}\mathfrak{N}\mathfrak{F}$. Then each $\mathfrak{F}$-normalizer of L is contained in exactly one $\mathfrak{F}$-projector of L.*

Proof. Let D be an $\mathfrak{F}$-normalizer of L and $D\subseteq E_1, E_2$ where E_1 and E_2 are $\mathfrak{F}$-projectors of L. Then $L/N(L)\in\mathfrak{N}\mathfrak{F}$ and $E_1+N(L)/N(L)=D+N(L)/N(L)=E_2+N(L)/N(L)$ by Theorem 3. Therefore $E_1+N(L)=D+N(L)=E_2+N(L)$. Hence $E_1=E_1\cap(D+N(L))=(E_1\cap N(L))+D$ and $E_2=(E_2\cap N(L))+D$. Let A be a minimal ideal of L. By induction $A+E_1=A+E_2$. Now E_1 and E_2 are both $\mathfrak{F}$-projectors of $A+E_1$ and $A+E_1/A\in\mathfrak{F}$. By Lemma 1.11 of [1], there exists $a\in A$ such that $E_1=E_2\,(I+\mathrm{ad}\,a)$. Also A is in the center of $N(L)$, hence

$$E_1\cap N(L)=E_2(I+\mathrm{ad}\,a)\cap N(L)=(E_2\cap N(L))\,(I+\mathrm{ad}\,a)=E_2\cap N(L).$$

Then

$$E_1=(E_1\cap N(L))+D=(E_2\cap N(L))+D=E_2.$$

The final result of this section is a characterization of $\mathfrak{F}$-projectors which is analogous to Theorem A of [9].

Theorem 10. *Let E be a subalgebra of L. E is an $\mathfrak{F}$-projector of L if and only if $\theta(E)$ is maximal in $\theta(L)$, subject to the condition of being in $\mathfrak{F}$, for every homomorphism θ of L.*

Proof. The condition is clearly necessary. Suppose then that the condition holds. Let E^* be a subalgebra of L such that $\theta(E^*)$ is maximal subject to the condition $\theta(E^*)\in\mathfrak{F}$ for every homomorphism θ of L. We use induction on dim L. If N is a minimal ideal of L, then the conditions carry over to L/N; hence, by induction, E^*+N/N is an $\mathfrak{F}$-projector of L/N. Let $M=E^*+N$. Then $M/N\simeq E^*/E^*\cap N\in\mathfrak{F}$, hence $M\in\mathfrak{N}\mathfrak{F}$. Also $N\subseteq N(M)$, hence $E^*+N(M)=M$. Now by Lemma 8, $E^*\subseteq F$ for some $\mathfrak{F}$-projector F of M. By assumption, $E^*=F$, hence E^* is an $\mathfrak{F}$-projector of M. Since E^*+N/N is an $\mathfrak{F}$-projector of L/N, E^* is an $\mathfrak{F}$-projector of L.

5. Example

In this section we deal with $\mathfrak{F}=\mathfrak{N}$ and consider the $\mathfrak{N}$-normalizers in a class of Lie algebras which may be thought of as analogues to the A-groups of Hall and Taunt. We need some more notation. The derived algebra of L will be denoted by L' and the derived algebra of L' by L''. The center of L will be denoted by $Z_1(L)$ and define the upper central series of L inductively by letting $Z_i(L)$ be the subalgebra of L such that $Z_i(L)/Z_{i-1}(L)\simeq Z_1(L/Z_{i-1}(L))$. The terminal member of the upper central series of L will be called the hypercenter and denoted by $H(L)$.

We shall say that L is an A-Lie algebra if all its nilpotent subalgebras are abelian. Clearly a subalgebra of an A-Lie algebra is also an A-Lie algebra. Also this property is preserved under homomorphisms. For let L be an A-Lie algebra and L/N a homomorphic image of L. Let K/N be a nilpotent subalgebra of L/N and C be a Cartan subalgebra of K. Then $C+N/N=K/N$ and since C must be abelian, K/N is abelian.

If L is a solvable Lie algebra and K is an $\mathfrak{N}$-normalizer of L, then, since K avoids all eccentric chief factors of L, one sees that K is contained in the hypercenter of its normalizer. Since K covers all central chief factors of L, K contains $H(L)$. For the same reason, $K+L'=L$. If L is an A-Lie algebra, then $Z_1(L)=H(L)$. For by the Jacobi identity $[Z_2(L),L']=0$. Since L is an A-Lie algebra, K is abelian and $[Z_2(L),K]=0$ since $Z_2(L)\subseteq H(L)\subseteq K$. Hence

$$[Z_2(L),L]=[Z_2(L),L'+K]=0 \quad\text{and}\quad Z_1(L)=Z_2(L)=H(L).$$

Theorem 11. *Let L be an A-Lie algebra. Then $Z_1(L)\cap L'=0$.*

Proof. Let L be a minimal counterexample. Then $Z_1(L)\cap L'=H\neq 0$. Let $N\neq 0$ be an ideal of L. Then $(L/N)'=L'+N/N$ and $(Z_1(L)+N)/N\subseteq Z_1(L/N)$. Since dim $L/N<$ dim L, $L'\cap Z_1(L)=H\subseteq N$ and H is the unique minimal ideal of L.

Now $L'\subset L$. Suppose $L''\neq 0$. Then $L''\supseteq H$. Now $H=L'\cap Z_1(L)$ is contained in the center of L' so that $H\subseteq L''\cap Z_1(L')$ which contradicts the minimality of L. Hence $L''=0$.

Let C be a Cartan subalgebra of L. Then $L=C+L'$ and the Fitting one-component C_1 of L with respect to C satisfies $C_1\subseteq L'$. Since L' is abelian, C_1 is an ideal with abelian factor isomorphic to C. Hence $C_1=L'$. Now if $x\in L'\cap Z_1(L)$, then $x\in N_L(C)=C$ and $x\in C\cap L'=0$.

From this result one sees that the derived algebra of an A-Lie algebra L avoids all central chief factors. If K is an $\mathfrak{N}$-normalizer of L, then K avoids all eccentric chief factors. Hence $K \cap L' = 0$. Therefore in an A-Lie algebra, the $\mathfrak{N}$-normalizers are complements of the derived algebra. As will be seen, not all complements of the derived algebra are $\mathfrak{N}$-normalizers.

Let $L = ((u, v, e_0, \dots, e_{p-1}))$ with multiplication $[u, v] = u$, $[e_i, e_j] = 0$, $[e_i, u] = e_{i+1}$, $[e_i, v] = i e_i$ (subscripts mod p) over a field of characteristic p. L is verified to be an A-Lie algebra and $((v))$, $((u+v))$ are $\mathfrak{N}$-normalizers of L. These subalgebras are not conjugate; in fact, there is no automorphism of L which takes $((v))$ into $((u+v))$ since their normalizers are of different dimensions. L has nilpotent length 3 and so $((v))$ is contained in a unique Cartan subalgebra by Theorem 9. It is $((v, e_0))$.

Now $((v+e_0))$ has the cover-avoidance property. However $((v+e_0))$ is not an $\mathfrak{N}$-normalizer of L. For, suppose $((v+e_0))$ is an $\mathfrak{N}$-normalizer of L. Then $((v+e_0))$ is contained in a maximal critical subalgebra P. Evidently P complements $Q = ((e_0, \dots, e_{p-1}))$ in L, so that P has dimension 2. Now there must exist a vector $z = a u + b v + \sum_{i=0}^{p-1} \alpha_i e_i$, where a, b, α_i are in the ground field, such that $P = ((z, v+e_0))$ and $[z, v+e_0] \in P$. Hence there exist scalers s and t such that $[v+e_0, z] = s(v+e_0) + t z$. We see that $-a = t a$, $0 = s + t b$, $0 = s + t \alpha_0$, $-\alpha_1 + a = t \alpha_1$ and $i \alpha_i = t \alpha_i$ for $i = 2, \dots, p-1$. One sees that $a = 0$ and either $t = 0$ or $b = \alpha_0$. If $b = \alpha_0$, then $z - b(v+e_0) \in Q \cap P$, hence $z - b(v+e_0) = 0$ and $\dim P = 1$, a contradiction. If $t = 0$, then $s = \alpha_1 = \dots = \alpha_{p-1} = 0$ and therefore $z = b v + \alpha_0 e_0$. Then $(b v + \alpha_0 e_0) - b(v+e_0) = (\alpha_0 - b) e_0 \in P \cap Q$, hence $\alpha_0 = b$. As we have seen this is a contradiction. Hence $((v+e_0))$ is not an $\mathfrak{N}$-normalizer of L.

Finally notice that $((v+e_0))$ is a complement of the derived algebra of L, but as we have seen, it is not an $\mathfrak{N}$-normalizer of L.

References

1. Barnes, D.W., Gastineau Hills, H.M.: On the theory of soluble Lie algebras. Math. Z. **106**, 343–354 (1969).
2. – Newell, M.L.: Some theorems on saturated homomorphs of soluble Lie algebras. Math. Z. **115**, 179–187 (1970).
3. Carter, R.W.: Nilpotent self-normalizing subgroups and system normalizers. Proc. London Math. Soc. **12**, 535–563 (1962).
4. – Hawkes, T.O.: The $\mathfrak{F}$-normalizers of a finite soluble group. J. Algebra **5**, 175–202 (1967).
5. Gaschütz, W.: Zur Theorie der endlichen auflösbaren Gruppen. Math. Z. **80**, 300–305 (1963).
6. Hall, P.: On the system normalizers of a soluble group. Proc. London Math. Soc. **43**, 507–528 (1937).
7. – The construction of soluble groups. J. Math. **182**, 206–214 (1940).
8. Hallahan, C.B., Overbeck, J.: Cartan subalgebras of meta-nilpotent Lie algebras. Math. Z. **116**, 215–217 (1970).
9. Hawkes, T.O.: On formation subgroups of a finite soluble groups. J. London Math. Soc. **44**, 243–250 (1969).
10. Huppert, B.: Zur Theorie der Formationen. Arch. der Math. **19**, 561–574 (1968).

11. Rose, J.: Finite soluble groups with pronormal system normalizers. Proc. London Math. Soc. (3) **17**, 447–469 (1967).
12. Stitzinger, E.: On the Frattini subalgebra of a Lie algebra. J. London Math. Soc. (2) **2**, 429–438 (1970).
13. – Theorems on Cartan subalgebras like some on Carter subgroups. Trans. Amer. Math. Soc. (to appear).
14. Taunt, D.R.: On A-groups. Proc. Cambridge Philos. Soc. **45**, 24–42 (1949).

Prof. E.L. Stitzinger
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27607, U.S.A.

(Received June 8, 1971)

Math. Z. 124, 250–254 (1972)

On Gruenberg Resolutions for Augmented Algebras

FRANZ BACHMANN*

Introduction

Let G be a group and $1 \to K \to F \to G \to 1$ any free presentation of G. Using the induced map of group algebras

$$Z[F] \xrightarrow{\phi} Z[G],$$

Gruenberg [3] found a very useful $Z[G]$-free resolution of Z. As this resolution depends only on the kernel of ϕ and the augmentation ideal of $Z[F]$, it is natural to ask if a Gruenberg resolution exists for augmented k-algebras (k a commutative ring): Let $T \xrightarrow{\phi} R$ be a surjective morphism of augmented k-algebras with kernel S. Denote the augmentation ideal of T by JT. Then it is easy to see that the following is an exact sequence of left R-modules:

$$G(\phi): \quad \cdots \to S^n/S^{n+1} \to S^{n-1} JT/S^n JT \to S^{n-1}/S^n$$
$$\cdots \to S/S^2 \to JT/S \cdot JT \xrightarrow{\bar{\phi}} R \xrightarrow{\varepsilon} k \to 0.$$

Here $\bar{\phi}$ is induced by ϕ, ε is the augmentation of R and all the other arrows are induced by inclusion. When is $G(\phi)$ an R-projective resolution of k? Knopfmacher in [4] gives a sufficient condition: If both S and JT are projective T-modules, then $G(\phi)$ is an R-projective resolution. However, an example (1.7) shows that T-projectivity of S and JT is not necessary. In fact, we prove the following

Theorem. *S and JT are T-projective iff*

(i) *$G(\phi)$ is an R-projective resolution of k,*

(ii) *ϕ is special*

where a special morphism fullfills some Bourbaki-condition (1.4) on the S-adic filtration of S and JT.

In Section 2 we consider the case where S is a principal ideal, generated by a non-divisor of zero of T. It turns out that (a) ϕ is special, (b) S is T-projective. (c) If JT is T-projective, $G(\phi)$ is a periodic resolution of k. We conclude our work with some examples; in particular, if $T=k[x]$ and $S=(f)$ we find a periodic resolution of k as a $k[x]/(f)$-module.

* This work was supported by the Swiss National Science Foundation.

1. The Gruenberg Sequence

1.1. Let k be a commutative ring and $T \xrightarrow{\phi} R$ a surjective morphism of augmented k-algebras. We denote $\operatorname{Ker}\phi$ by S and the augmentation ideal of T by JT. Setting $S^0 = T$ we define ([4]) left R-modules X_n $(n \geqq 0)$ by

$$X_n = S^{m-1} JT/S^m JT \quad \text{if } n = 2m-1,$$
$$X_n = S^m/S^{m+1} \quad \text{if } n = 2m.$$

The *Gruenberg sequence* of ϕ is defined as

$$G(\phi): \ \cdots \to X_n \to X_{n-1} \to \cdots \to X_1 \to X_0 \to k \to 0$$

where $X_0 \to k$ is the augmentation of R, $X_1 \to X_0$ is induced by ϕ and $X_n \to X_{n-1}$ $(n > 1)$ is induced by inclusion. R-exactness of the following things shows us that $G(\phi)$ is an exact sequence of left R-modules:

$$0 \to S^n JT/S^{n+1} \to S^n/S^{n+1} \to S^n/S^n JT \to 0,$$
$$0 \to S^{n+1}/S^{n+1} JT \to S^n JT/S^{n+1} JT \to S^n JT/S^{n+1} \to 0.$$

1.2. *Remark.* The above construction of $G(\phi)$ can be generalized: Let $(F_p A)_{p \geqq 0}$ be a filtration of the ring A such that

(i) $F_0 A = A$.

(ii) $A/F_1 A$ is a ring.

Take a left ideal $N \subset A$ with the "induced" filtration $(F_p N)$ such that $F_{p+1} A \subset F_p N$ $(p \geqq 0)$. Then

$$\ldots F_{p+1} A/F_{p+2} A \to F_p N/F_{p+1} N \to F_p A/F_{p+1} A$$
$$\cdots \to N/F_1 N \xrightarrow{L} A/F_1 A \to \operatorname{Cok} L \to 0$$

is an exact sequence of $A/F_1 A$-modules.

1.3. We return to the situation of 1.1. Let M be a left T-module and $\operatorname{gr}(M) = \bigoplus_{n \geqq 0} \operatorname{gr}_n(M)$ the graded T-module associated to the S-adic filtration on M. There is a natural R-linear surjective map ([1])

$$u_M: \operatorname{gr}(T) \otimes_R M/SM \to \operatorname{gr}(M)$$

given by multiplication. We denote the n-th component of u_M by $u_M^{(n)}$.

Lemma ([1]). *Consider the following statements:*

(a) *M is left T-flat.*

(b) *M/SM is left R-flat and* $\operatorname{Tor}_1^T(R, M) = 0$.

(c) *M/SM is left R-flat and u_M is bijective.*

Then we have (a) $\Rightarrow$ (b) $\Rightarrow$ (c).

Proof. The proof of this lemma is well-known: (a) $\Rightarrow$(b) is trivial. (b) $\Rightarrow$(c) follows from the fact that

$$S^n \otimes_T M \cong S^n M \qquad (n \geqq 1)$$

and the commutativity of the diagram

$$\begin{tikzcd} & S^{n+1}\otimes_T M \arrow[r] \arrow[d] & S^n \otimes_T M \arrow[r] \arrow[d] & S^n/S^{n+1}\otimes_T M/SM \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & S^{n+1} M \arrow[r] & S^n M \arrow[r] & \mathrm{gr}_n(M) \arrow[r] & 0. \end{tikzcd}$$

1.4. **Definition.** Let $T \xrightarrow{\phi} R$ be a surjective morphism of augmented k-algebras. ϕ is called *special* if

$$u_{JT}\colon \mathrm{gr}(T)\otimes_R JT/S\cdot JT \to \mathrm{gr}(JT)$$

and

$$u_S\colon \mathrm{gr}(T)\otimes_R S/S^2 \longrightarrow \mathrm{gr}(S)$$

are bijective.

1.5. From 1.3 we know that ϕ is special if JT and S are projective T-modules. The converse is not true in general; however, if we put stronger conditions on ϕ, we can characterize those morphisms ϕ such that JT and S are T-projective:

Theorem. *For our morphism ϕ the following statements are equivalent:*

(a) *JT and S are projective left T-modules.*

(b) *ϕ is special and $G(\phi)$ is an R-projective resolution of k.*

Proof. Let JT and S be T-projective. We show by induction on n that S^n/S^{n+1} is left R-projective, $(n\geqq 1)$. For $n=1$, the assertion follows from Lemma 1.3. Consider then the isomorphism

$$S^n/S^{n+1} \cong S^{n-1}/S^n \otimes_R S/S^2.$$

By induction hypothesis and [2] the left hand side is left R-projective. Again by Lemma 1.3 we have isomorphisms

$$S^n\, JT/S^{n+1}\, JT \cong S^n/S^{n-1} \otimes_R JT/S\cdot JT.$$

This proves R-projectivity of $S^n\, JT/S^{n+1}\, JT$ $(n\geqq 1)$.

Assume now (b). We want to show that JT is T-projective. By a result of Roy [6] it is enough to know that $\mathrm{gr}(JT)$ is $\mathrm{gr}(T)$-projective. From (b) it follows that $\mathrm{gr}(JT)$ is R-projective. Therefore, $\mathrm{gr}(T)\otimes_R \mathrm{gr}(JT)$ is $\mathrm{gr}(T)$-projective. Using the isomorphism

$$\mathrm{gr}(T)\otimes_R \mathrm{gr}(JT) \cong \Big(\mathrm{gr}(T)\otimes_R \big(\bigoplus_{i>0} S^i\, JT/S^{i+1}\, JT\big)\Big) \oplus \mathrm{gr}(T)\otimes_R JT/S\cdot JT$$

we conclude that $\mathrm{gr}(T)\otimes_R JT/S\cdot JT \cong \mathrm{gr}(JT)$ is $\mathrm{gr}(T)$-projective too. The same idea applies to S. This completes the proof of the theorem.

1.6. *Remark.* The statements "ϕ is special" and "$G(\phi)$ is an R-projective resolution" are far from being independent. Indeed, if ϕ is special and $S/S^2 \oplus JT/SJT$ is R-projective, $G(\phi)$ is R-projective. On the other hand, projectivity of $G(\phi)$ gives us a section of u_{JT} and u_S. However, 1.7 shows that there exist non-special morphisms ϕ such that $G(\phi)$ is projective.

1.7. *Example.* Take $k = Z_2$, $T = Z_{12}$, $R = Z_6$, ϕ the obvious projection. We have then $S = 6 \cdot Z_{12}$, $JT = 2 \cdot Z_{12}$. So $S \cdot JT = S^2 = 0$ in Z_{12}. The Gruenberg sequence looks as follows:

$$0 \to S \to JT \to Z_6 \to Z_2 \to 0.$$

This is a Z_6-projective resolution of Z_2 because Z_6 is a semi-simple ring ([2]). On the other hand, the sequence

$$0 \to 2 \cdot Z_{12} \to Z_{12} \to 6 \cdot Z_{12} \to 0$$

is obviously not Z_{12}-split; hence, S and JT cannot be T-projective.

2. Principal Ideals

2.1. Let T be an augmented k-algebra and let s be a central non-divisor of zero of T such that $s \in JT$. Then the quotient T/sT is augmented in a natural way; we are interested in the morphism

$$T \xrightarrow{\phi} T/sT.$$

2.2. **Lemma.** *ϕ is special.*

Proof. We claim that the multiplication map

$$S^n \otimes_T S \xrightarrow{f_n} S^{n+1}$$

is a bijection for $n \geqq 1$: An inverse g_n for f_n is given by $g_n(s^{n+1} \cdot t) = s^n \otimes s \cdot t$. By the proof of Lemma 1.3 this implies that u_S is an isomorphism. A similar conclusion holds for u_{JT}.

2.3. **Lemma.** *S is T-projective.*

Proof. This is trivial: s establishes an isomorphism $sT = T$.

2.4. **Theorem.** *Let ϕ be as in 2.1 with $T/sT = R$. Then $G(\phi)$ is a periodic sequence:*

$$\ldots R \otimes_T JT \xrightarrow{d_1} R \xrightarrow{d_2} R \otimes_T JT \xrightarrow{d_1} \ldots \xrightarrow{d_2} R \otimes_T JT \xrightarrow{d_1} R \to k \to 0$$

where d_1 is given by multiplication and $d_2(r) = r \otimes s$ $(r \in R)$.

Moreover, $G(\phi)$ is an R-projective resolution of k iff JT is T-projective.

Proof. The second assertion follows from Theorem 1.5 and Lemmas 2.2 and 2.3. For the first assertion we observe (using again the properties of s)

$$S^n/S^{n+1} \overset{\alpha_n}{\cong} T/S \cong R$$

and

$$S^n JT/S^{n+1} JT \overset{\beta_n}{\cong} JT/S \cdot JT \cong R \otimes_T JT.$$

Of course, α_n and β_n are defined as "cutting away the s^n part", whereas their inverses are multiplications by s^n.

2.5. **Corollary.** *If JT is T-projective,* $\mathrm{l.dim}_R\, k=0$ *or* ∞.

2.6. *Example.* Let k be a commutative ring and $T=k[x]$. Let $f \in k[x]$ be a monic polynomial such that there exists a $w \in k$ with $f(w)=0$. We augment $k[x]$ by sending x to w ([5]). Let us write k_f for $k[x]/(f)$.

Proposition. *A k_f-projective resolution of k is given by*

$$\cdots \to k_f \to k_f \otimes_T JT \to k_f \to \cdots \to k_f \otimes_T JT \to k_f \to k \to 0$$

with maps defined as in 2.4.

Proof. This is an application of Theorem 2.4: JT is generated (as a $k[x]$-module) by the non-divisor of zero $(x-w)$. Therefore, JT is T-projective.

Another such resolution was found by Ramanan-Sridharan in [5].

2.7. Almost-example. Take for T the free ring on two symbols x and y, $S=(xy-yx)$. Therefore, $R=k[x,y]$. Let now k be a field; then T is a free ideal ring (e.g. [4]), so $G(\phi)$ is a projective resolution. On the other hand we know ([2]) that $\dim_{k[x,y]} k=2$. This shows that the assumption "s central" is essential in the proof of 2.4.

References

1. Bourbaki, N.: Algèbre commutative, Ch. 3. Paris: Hermann 1964.
2. Cartan, H., Eilenberg, S.: Homological algebra. Princeton University Press 1956.
3. Gruenberg, K. W.: Resolutions by relations. J. London Math. Soc. **35**, 481–494 (1960).
4. Knopfmacher, J.: Some homological formulae. J. Algebra **9**, 212–219 (1968).
5. Ramanan, S., Sridharan, R.: Resolutions for some filtered algebras. Math. Ann. **148**, 341–348 (1962).
6. Roy, A.: A note on filtered rings. Arch. der Math. **16**, 421–427 (1965).

Dr. F. Bachmann
Forschungsinstitut für Mathematik
ETH
8006 Zürich, Schweiz

(Received March 1, 1971)

Math. Z. 124, 255–258 (1972)

Note on the Integral Group Ring Problem

Robert Sandling

This note provides new approaches to three known results on the integral group ring problem, that the integral group ring determines normal abelian sections, the commutator subgroup of two normal subgroups and the upper class 2 section. The first two paragraphs give simpler proofs for several results while the third paragraph introduces a new method of attacking the problem.

Let G and G^* be finite groups for which $\mathbf{Z}G$ is isomorphic to $\mathbf{Z}G^*$ as rings and therefore, as we may assume, as augmented algebras. Glauberman and Passman [3] show that such an isomorphism φ sets up a bijective correspondence * between the normal subgroups of G and those of G^*; this is defined by $\eta_{N^*}=\varphi(\eta_N)$. Here η_H, for any finite group H, denotes the norm element of $\mathbf{Z}H$, that is, the sum of the elements of H; let $I(H)$ denote the augmentation ideal of $\mathbf{Z}H$, the annihilator of η_H. Since $I(N)\mathbf{Z}G$ is the annihilator of η_N in $\mathbf{Z}G$, we see that $I(N^*)\mathbf{Z}G^*=\varphi(I(N)\mathbf{Z}G)$.

Let H be a subgroup of G and T a set of coset representatives for H in G, which contains 1. Since $\mathbf{Z}G$ is free abelian on the elements of G, there is a homomorphism ρ from $\mathbf{Z}G$ to H/H' defined by sending x in G, expressed as ht, h in H and t in T, to the coset of h in H/H'.

The last item of notation to be explained here is that for the subgroup attached to an ideal J of $\mathbf{Z}G$. The set $G\cap 1+J$ of all elements x in G for which $x-1$ is in J is a subgroup of G. The author's thesis [4] consists largely of the determination of these subgroups for specific ideals. This study led to the present observations on the integral group ring problem. The author thanks the NSF for its support and the University of Cambridge for its extensive hospitality.

§1. Abelian Sections

Most of this paragraph is to be found in Whitcomb's thesis [6]. The mention of the $\mathbf{Z}G$-module structure is the (slight) innovation here, which makes Corollary ii available.

Let N be a normal subgroup of G. Then N/N' is a left $\mathbf{Z}G$-module by conjugation. Thus,

Lemma 1. *The map ρ from $\mathbf{Z}G$ to N/N' induces an isomorphism of left $\mathbf{Z}G$-modules from*

$$I(N)\,\mathbf{Z}G/I(N)\,I(G) \quad \text{to} \quad N/N'.$$

Furthermore, if $I(N)\,I(G)\subseteq J\subseteq I(N)\,\mathbf{Z}G$, the image of J under ρ is H/N' where $H=G\cap 1+J$.

Proof. The inverse of ρ is obtained by sending the coset of n to the coset of $n-1$; thus, ρ is an isomorphism.

The identity $x(n'-1)=(x\,n'\,x^{-1}\,n-1)\,t-(n-1)\,t$, $x=n\,t$, shows that ρ induces a left $\mathbf{Z}G$-map.

Lastly, if α is in J and α is sent to the coset of n, then $(n-1)-\alpha$ is in $I(N)\,I(G)$ so that n is in H.

Proposition 1. *If N and M are normal in G and M is in N and N/M is abelian, then φ induces a $\mathbf{Z}G$-isomorphism between N/M and N^*/M^*, where N^*/M^* is a $\mathbf{Z}G$-module via $\mathbf{Z}G^*=\varphi(\mathbf{Z}G)$.*

Proof. The map ρ induces an isomorphism of left $\mathbf{Z}G$-modules from

$$I(N)\,\mathbf{Z}G/I(N)\,I(G)+I(M)\,\mathbf{Z}G \quad \text{to} \quad N/M.$$

The same situation prevails for G^*. But φ carries these ideals to their counterparts in $\mathbf{Z}G^*$ involving N^* and M^*.

The following corollaries are then immediate, the first because N' is the smallest subgroup for which N/N' is abelian, the second because $[N, G]$ is the smallest subgroup for which $N/[N, G]$ is a trivial $\mathbf{Z}G$-module.

i) $N'^*=N^{*\prime}$; thus, $G^{(n)*}=G^{*(n)}$ and the lower derived factors of G and G^* are isomorphic as $\mathbf{Z}G$-modules (see Obayashi [2]).

ii) $[N, G]^*=[N^*, G^*]$; thus, $G_n{}^*=G^*{}_n$ and the lower central factors of G and G^* are isomorphic (see Cohn and Livingston [1]).

iii) $Z(G; N)^*=Z(G^*; N^*)$; thus, $Z_n(G)^*=Z_n(G^*)$ and the upper central factors of G and G^* are isomorphic (see Passman [3]). Here $Z(G; N)/N$ is the center of G/N, the largest subgroup of G such that $Z(G; N)/N$ is a trivial $\mathbf{Z}G$-module.

§ 2. Commutators

The proposition of this paragraph implies most of the above corollaries. The proposition was established by Passman [3] for nilpotent groups and by Whitcomb in general. Thruout, N and M are normal subgroups of G. If A and B are subsets of $\mathbf{Z}G$, define the Lie bracket (A, B) as the subgroup of $\mathbf{Z}G$ generated by all $(\alpha, \beta)=\alpha\beta-\beta\alpha$, α in A, β in B. If A and B are ideals, (A, B) need not be an ideal.

Lemma 2. $[N, M]=G\cap 1+\big(I(N)\,\mathbf{Z}G, I(M)\,\mathbf{Z}G\big)\,\mathbf{Z}G$.

Proof. Since $I([N, M])\,\mathbf{Z}G$ is contained in the ideal on the right, we may assume that $[N, M]=1$. In this case, we can even show that $1=G\cap 1+I(N)\,I(M)\,\mathbf{Z}G$. To prove this, we may assume that $G=NM$ by the following principle:

If H is a subgroup of G and J a right ideal of $\mathbf{Z}H$ contained in $I(H)$, then $G\cap 1+J\mathbf{Z}G=H\cap 1+J$.

With $G=NM$, $A=N\cap M$ is central and $I(N)\,I(M)$ is already an ideal. It suffices to choose coset representatives for A so that the associated map ρ

has $I(N)\,I(M)$ in its kernel. Let S be a set of representatives for A in N and T a set for A in M, both containing 1; then the set ST is an appropriate set of representatives for A in G. Taking sa in N and ta' in M, we see that ρ sends $(sa-1)(ta'-1)$ to $aa'a^{-1}a'^{-1}=1$.

Proposition 2. $[N, M]^* = [N^*, M^*]$.

Proof. This follows from the properties of φ and the fact that, if J is a two-sided ideal of $\mathbf{Z}G$ and $H=G\cap 1+J$, then H is normal in G and $H^*=G^*\cap 1+\varphi(J)$.

§3. The Class 2 Case

The Lemma in this paragraph is stated without proof. The proof appears in [5] as a new method of proving the dimension subgroup conjecture for class 2 groups. The proof is laborious and would contribute little in the present context. Here I^n denotes the n-th power of $I(G)$.

Lemma 3. *There is an ideal J of $\mathbf{Z}G$, not canonically defined, such that I^2/I^3 is isomorphic to the direct sum $G'/G_3 \oplus J/I^3$.*

This result is analogous to the classic result that I/I^2 is isomorphic to G/G'. The author has been unable to decide whether there are further such results such as J/I^4 isomorphic to $G_3/G_4 \oplus K/I^4$, etc.; if so, they would simultaneously resolve both the integral group ring problem for nilpotent groups and the dimension subgroup problem (whether $G_n = G\cap 1+I^n$).

Proposition 3. *G/G_3 is isomorphic to G^*/G^*_3.*

Proof. The isomorphism of the Lemma is actually one of the form

$$I^2/I^3 \approx J/I^3 \oplus \big(I(G')\,\mathbf{Z}G + I^3\big)/I^3$$

where the latter term is isomorphic to G'/G_3. Since the latter term is invariant under φ, we obtain a similar decomposition in $\mathbf{Z}G^*$.

Modulo J, I is a nilpotent ideal so that we may form the group of units $1+I/J$ which is isomorphic under φ to the group $1+I(G^*)/\varphi(J)$. Thus, it suffices to show that G/G_3 is isomorphic to $1+I/J$, which is obvious since $I^3 \subseteq J$ shows $1+I/J$ to have class 2 while Lemma 3 and the fact that $I/I^2 \approx G/G'$ show $1+I/J$ to have the correct order.

This proposition was first proved in G. Higman's thesis; subsequently, Passman [3] proved it and it is a consequence of the result of Whitcomb below. The proof here is unusual in that it finds a section of the group in terms of a noncanonical ideal. Recently other sections of the group have been found canonically in the group ring, particularly for solvable groups:

a) [6] G/G'' is isomorphic to the canonical subset $G+I(G')\,\mathbf{Z}G$ modulo $I(G')\,I(G)$.

b) [4] Let $N=G_\omega$, the intersection of the lower central series, $i_{\omega+n}=G\cap 1+I^\omega I^n$ where I^ω is the intersection of the powers of I, N_p and S_p be Sylow p-subgroups of N and G respectively for all p. Then

$$i_{\omega+n} = N' \prod [N_p, S_q, S_r] \prod [N_p, S_p, \ldots, S_p]$$

where the first product is taken over all triples of distinct primes p, q, r and the second over all primes p and in which each term contains n copies of S_p. The sections $i_\omega/i_{\omega+n}$ are determined as $I^\omega/I^{\omega+n}$ while the sections $i_{\omega+n}/i_{\omega+n+1}$ are determined as $\mathbf{Z}G$-modules as $I^{\omega+n}/I^{\omega+n+1}$. The method for establishing these results is due to Roseblade and Gruenberg.

References

1. Cohn, J. A., Livingstone, D.: On the structure of group algebras, I. Canadian J. Math. **17**, 583–593 (1965).
2. Obayashi, T.: Solvable groups with isomorphic group algebras. J. Math. Soc. Japan **18**, 394–397 (1966).
3. Passman, D. S.: Isomorphic groups and group rings. Pacific J. Math. **15**, 561–583 (1965).
4. Sandling, R.: The modular group rings of p-groups. Thesis, Univ. of Chicago, December 1969.
5. — Dimension subgroups over arbitrary coefficient rings. J. Algebra, to appear.
6. Whitcomb, A.: The group ring problem. Thesis, Univ. of Chicago, March 1968.

Dr. Robert Sandling
Department of Mathematics
The University
Manchester, U.K. M 13 9 PL

(Received July 14, 1970)

Math. Z. 124, 259–260 (1972)

Correction to "Locally Trivial Outer Automorphisms of Finite Groups"

Math. Z. **114**, 173–179 (1970)

EVERETT C. DADE

Reinhard Lane has called my attention to an error appearing twice in the proofs on page 177 of [1]. It concerns a finite group U having a normal Hall subgroup H, and an automorphism α of U centralizing both H and U/H. I had assumed that α would then fix a complement to H in U and hence be trivial. The correct result, as Lane pointed out, is the

Proposition. *The automorphisms α of U having the above property are precisely the inner automorphisms by elements of the center $\mathrm{Z}(H)$ of H.*

Proof. Obviously all these inner automorphisms have the property.

Suppose $\alpha \in \mathrm{Aut}(U)$ centralizes both H and U/H. Let V be any complement to H in U. Then V is also a complement to the normal Hall subgroup $\mathrm{Z}(H)$ in $\mathrm{C}_U(H)\,V = \mathrm{Z}(H)\,V$. Since α centralizes H, it centralizes $U/\mathrm{C}_U(H)$. It follows that

$$\mathrm{Z}(H)\,V^{\alpha} = [\mathrm{Z}(H)\,V]^{\alpha} = [\mathrm{C}_U(H)\,V]^{\alpha} = \mathrm{C}_U(H)\,V = \mathrm{Z}(H)\,V.$$

So V^{α} is another complement to $\mathrm{Z}(H)$ in $\mathrm{Z}(H)\,V$. Because $\mathrm{Z}(H)$ is an abelian normal Hall subgroup of $\mathrm{Z}(H)\,V$, there exists $\sigma \in \mathrm{Z}(H)$ such that $V^{\alpha} = V^{\sigma}$. If β is the inner automorphism of U determined by σ, then obviously $\alpha\beta^{-1}$ centralizes both U/H and H, and fixes $V \cong U/H$. So $\alpha\beta^{-1} = 1$ and $\alpha = \beta$, which proves the proposition.

Because of the above remark of Lane, we must change the proofs in [1] at two points. From now on, all numbers refer to statements in [1] and we adopt the notation of that paper.

Proof of Lemma (1.14). The original proof shows that K normalizes $L = \mathrm{N}_G(\mathrm{C}_G(K))$, and centralizes both the Hall subgroup $\mathrm{C}_G(K)$ of L and the factor group $L/\mathrm{C}_G(K)$. Setting $Z = \mathrm{Z}(\mathrm{C}_G(K))$ and using the subscript 1 to indicate images in $\mathrm{Aut}(L)$, we obtain from the above proposition the inclusion:

$$K_1 \leqq Z_1.$$

If $Z = 1$, then $K_1 = Z_1 = 1$. If $Z \neq 1$, then $K = \mathrm{K}(Z)$ by Lemma (1.11). We know from (1.5)–(1.7) that $K = [\mathrm{C}_S(Z\mathrm{C}_G(Z))_p, \mathrm{C}_T(Z)]$, where $\mathrm{C}_S(Z\mathrm{C}_G(Z))_p$ is the p-Sylow subgroup of the nilpotent group $\mathrm{C}_S(Z\mathrm{C}_G(Z))$. Obviously $\mathrm{C}_G(Z)$

centralizes $C_S(ZC_G(Z))_p$. So $C_T(Z)$ acts on $C_S(ZC_G(Z))_p$ as the q-group $C_T(Z)/C_G(Z) \cong T^*$. Since $q \neq p$, it follows that $[K, C_T(Z)] = K$. Evidently $C_T(Z)$ normalizes K, and hence normalizes $L = N_G(C_G(K))$. So we can take images in $\mathrm{Aut}(L)$, obtaining:

$$K_1 = [K_1, C_T(Z)_1] \leqq [Z_1, C_T(Z)_1] = [Z, C_T(Z)]_1 = 1.$$

Therefore $K_1 = 1$ in all cases, and K centralizes L, which proves Lemma (1.14).

Proof of the Theorem. On line 10 of page 177 of [1] the group G is a Frobenius group with kernel N, and $K(N)$ centralizes both G/N and N. Since N is then a Hall subgroup of G, the above proposition implies that $K(N) \leqq Z(N)$ (everything being already embedded in $\mathrm{Aut}(G)$). This is impossible by Lemma (1.6), since $S/G = S^* \neq 1$. The contradiction proves the theorem.

Reference

1. Dade, E.C.: Locally trivial outer automorphisms of finite groups. Math. Z. **114**, 173–179 (1970).

Prof. E.C. Dade
Department of Mathematics
University of Illinois
Urbana, Illinois 61801
USA

(Received September 23, 1971)

Math. Z. 124, 261–265 (1972)

k-Homogeneous Groups*

WILLIAM M. KANTOR

1. Introduction

A permutation group is called k-homogeneous if it is transitive on the k-sets of permuted points.

Theorem 1. *Let G be a group k-homogeneous but not k-transitive on a finite set Ω of n points, where $n \geqq 2k$. Then, up to permutation isomorphism, one of the following holds:*

(i) $k=2$ *and* $G \leqq A\Gamma L(1,q)$ *with* $n=q\equiv 3 \pmod 4$;

(ii) $k=3$ *and* $PSL(2,q) \leqq G \leqq P\Gamma L(2,q)$, *where* $n-1=q\equiv 3 \pmod 4$;

(iii) $k=3$ *and* $G=AGL(1,8)$, $A\Gamma L(1,8)$ *or* $A\Gamma L(1,32)$; *or*

(iv) $k=4$ *and* $G=PSL(2,8)$, $P\Gamma L(2,8)$ *or* $P\Gamma L(2,32)$.

Here $A\Gamma L(1,q)$ is the group of mappings $x \to a x^{\sigma}+b$ on $GF(q)$, where $a \neq 0$ and b are in $GF(q)$ and $\sigma \in \operatorname{Aut} GF(q)$. $AGL(1,q)$ consists of those mappings with $\sigma=1$. All the groups listed in the theorem are assumed to act in their usual permutation representations.

We note that, conversely, each of (i)–(iv) produces examples of k-homogeneous but not k-transitive groups. Thus, in (i) we need only consider maps of the form $x \to a^2 x+b$. Moreover, $PSL(2,q)$, $q\equiv 3 \pmod 4$, and the groups in (iii) and (iv) meet our requirements.

This theorem completes results of Livingstone and Wagner [8], who showed that k must be at most 4. Clearly $k>1$. For the case $k=2$, see [7], Proposition 3.1. The case $k=4$ was considered in [6], but there is an error in the proof. Note that the hypothesis $n \geqq 2k$ is essential, since, for example, a 2-transitive group of degree n is $(n-2)$-homogeneous.

The case $k=4$ will follow easily from the case $k=3$. If $k=3$ and neither (ii) nor (iii) holds, it is easy to show that $3 \nmid |G|$ and the stabilizer of 2 points has precisely 3 orbits on the remaining points. However, the deep group-theoretic results presently known about 3′-groups do not seem to apply to our situation. Instead of these we use a combinatorial argument, based on the proof of Gleason's lemma ([4], Lemma 1.7), in order to employ induction.

Our notation is that of Wielandt [9]. If X is a subset of a permutation group then $\Delta(X)$ is its set of fixed points.

* Research supported in part by NSF Grant GP 9584.

2. Induction

In order to use induction to prove Theorem 1 when $k=3$, we will use somewhat different hypotheses. The result will be an easy consequence of

Theorem 2. *Let G be a group 2-transitive on a finite set Ω, where $n=|\Omega|>2$. Assume:*

(a) $3\nmid|G|$; *and*

(b) *If $\alpha\neq\beta$ then $G_{\alpha\beta}$ has precisely 3 orbits on $\Omega-\{\alpha,\beta\}$.*

Then $n=5$ and $|G|=20$.

The following result will be used very frequently.

Lemma 1 (Livingstone and Wagner [8], Theorem 3; Bender [1], Lemma 3.3). *Let K be a group transitive on a finite set Φ. Let r be a prime and R an r-subgroup maximal with respect to fixing $\geqq 2$ points. Then $N(R)^{\Delta(R)}$ is transitive.*

Proof of Theorem 2. Let G be a counterexample with n minimal. By (a), $n>5$.

For $\alpha\neq\beta$ let $\Gamma_i(\alpha,\beta)$, $i=1,2,3$, be the orbits of $G_{\alpha\beta}$ on $\Omega-\{\alpha,\beta\}$; this labeling is chosen in any way. (These orbits cannot necessarily be labelled so that $\Gamma_i(\alpha^g,\beta^g)=\Gamma_i(\alpha,\beta)^g$ for all $g\in G$, as some g might interchange α and β and also interchange two of the orbits $\Gamma_i(\alpha,\beta)$.) For each i we have $|\Gamma_i(\alpha,\beta)|>1$, as otherwise by (b) $N(G_{\alpha\beta})^{\Delta(G_{\alpha\beta})}$ is 2-transitive of degree 3 or 4, contradicting (a).

Let p be a prime such that there is a nontrivial p-group fixing >2 points. Let $P\leqq G_{\alpha\beta}$ be such a p-group maximal with respect to fixing >2 points.

Lemma 2. $|\Delta(P)|=5$, $|N(P)^{\Delta(P)}|=20$ *and* $|\Delta(P)\cap\Gamma_i(\alpha,\beta)|=1$, $i=1,2,3$.

Proof. By Lemma 1, $N(P)^{\Delta(P)}$ is 2-transitive. Let $i=1,2$, or 3. If a p-subgroup of $G_{\alpha\beta}$ fixes >1 point of $\Gamma_i(\alpha,\beta)$ it certainly fixes >2 points of Ω. Thus, if $\Delta(P)\cap\Gamma_i(\alpha,\beta)\neq\emptyset$ then by Lemma 1 $N(P)_{\alpha\beta}$ is transitive on $\Delta(P)\cap\Gamma_i(\alpha,\beta)$.

If $\Delta(P)$ meets all $\Gamma_i(\alpha,\beta)$ then (a) and (b) hold for $N(P)^{\Delta(P)}$, and the lemma follows from the minimality of n. Since $N(P)^{\Delta(P)}$ is not 3-transitive by (a), $\Delta(P)$ cannot meet just one $\Gamma_i(\alpha,\beta)$.

Suppose that $\Delta(P)$ meets just two sets $\Gamma_i(\alpha,\beta)$. There is a natural 1-1 correspondence between the orbits O of $N(P)^{\Delta(P)}$ of ordered triples (α,β,γ) of distinct points of $\Delta(P)$ and the orbits of $N(P)^{\Delta(P)}_{\alpha\beta}$ on $\Delta(P)-\{\alpha,\beta\}$. If O is such an orbit then so are $O'=\{(\alpha,\beta,\gamma)|(\beta,\gamma,\alpha)\in O\}$ and $O''=\{(\alpha,\beta,\gamma)|(\gamma,\alpha,\beta)\in O\}$. Since two of O, O', O'' are the same in our case, $N(P)$ contains a 3-element $\neq 1$, which is not the case.

Lemma 3. *No nontrivial element fixes more than 5 points.*

Proof. If this is not the case, there is a prime p such that some nontrivial p-group fixes >5 points. Choose Q maximal among such p-groups. Set $\Delta=\Delta(Q)$ and $H=N(Q)^{\Delta}$. Let $P>Q$ be as in Lemma 2.

By Lemma 2, $|\Gamma_i(\alpha,\beta)|\equiv 1 \pmod p$, $i=1,2,3$. Thus, for any $\alpha^*,\beta^*\in\Delta$, $\alpha^*\neq\beta^*$, we have

$$|\Delta\cap\Gamma_i(\alpha^*,\beta^*)|\equiv 1 \pmod p,\quad i=1,2,3. \qquad (*)$$

If a p-subgroup of $G_{\alpha^*\beta^*}$ fixes >1 point of some $\Gamma_i(\alpha^*, \beta^*)$ then it fixes >5 points by Lemma 2. Thus, if $|\Delta \cap \Gamma_i(\alpha^*, \beta^*)|>1$ then Q is maximal among the p-subgroups of $G_{\alpha^*\beta^*}$ fixing >1 point of $\Gamma_i(\alpha^*, \beta^*)$. By Lemma 1, $H_{\alpha^*\beta^*}$ is transitive on $\Delta \cap \Gamma_i(\alpha^*, \beta^*)$, $i=1, 2, 3$. In particular, by the minimality of n, H is not 2-transitive on Δ. Frequent use will be made of the fact that, for any distinct $\alpha^*, \beta^* \in \Delta$, $H_{\alpha^*\beta^*}$ has precisely 5 orbits on Δ.

Let $\alpha \in \Delta$. By (*) H_α has a nontrivial orbit Φ on $\Delta - \alpha$. Let $\beta \in \Phi$.

We claim that H_α fixes no point $\delta \in \Delta - \alpha$. For otherwise, $H_{\alpha\beta} < H_\alpha = H_{\alpha\delta}$, where each of these groups has 5 orbits on Δ and β is an orbit of $H_{\alpha\beta}$ but not of $H_{\alpha\delta}$. This is clearly impossible.

In particular, H fixes no point $\delta \in \Delta$.

Suppose that H is intransitive on Δ. Let Δ' and Δ'' be (nontrivial) orbits of H on Δ. Let $\alpha' \in \Delta'$ and $\alpha'' \in \Delta''$. Both $\Delta' - \alpha'$ and $\Delta'' - \alpha''$ are unions of certain of the 5 orbits of $H_{\alpha'\alpha''}$ on Δ. By (*), $|\Delta' - \alpha'| \equiv 1$ or 2 and $|\Delta'' - \alpha''| \equiv 1$ or $2 \pmod p$, but $|\Delta' - \alpha'| \equiv |\Delta'' - \alpha''| \equiv 2 \pmod p$ does not occur. We may assume that $|\Delta'| \equiv 2 \pmod p$ and $H_{\alpha'\alpha''}$ is transitive on $\Delta' - \alpha'$. Then H is 2-transitive on Δ'. Let $\beta' \in \Delta' - \alpha'$. Note that $\Delta' \neq \{\alpha', \beta'\}$ since $H_{\alpha'}$ cannot fix the point $\beta' \in \Delta - \alpha'$. Thus $|\Delta'| > 2$. Now $H_{\alpha'\beta'}$ has at least one orbit on Δ'' and hence at most two orbits on $\Delta' - \{\alpha', \beta'\}$. Consequently, $H^{\Delta'}$ is either 3-transitive or a transitive extension of a rank 3 group. As at the end of the proof of Lemma 2 we obtain $3 \mid |H|$, contradicting (a).

Thus, H is transitive on Δ. Recall that each orbit of $H_\alpha^{\Delta-\alpha}$ is nontrivial. Since H is not 2-transitive on Δ, we can find at least two orbits Φ, Φ' of H_α on $\Delta - \alpha$.

Here $|\Phi|$ and $|\Phi'|$ are >2. For suppose $|\Phi|=2$. Let $\alpha \neq \delta \in \Delta - \Phi$. By (*), $H_{\alpha\delta}$ fixes Φ pointwise, so $H_{\alpha\delta} \leqq H_{\alpha\beta}$. Since $H_{\alpha\delta}$ and $H_{\alpha\beta}$ have 5 orbits, $H_{\alpha\beta}$ must fix δ. Consequently, $H_{\alpha\beta}=1$, contradicting the fact that $|\Delta|>5$.

Both $\Phi - \beta$ and Φ' are unions of certain of the 5 orbits of $H_{\alpha\beta}$ on Δ. By (*), $|\Phi - \beta| \equiv 1$ or 2 and $|\Phi'| \equiv 1$ or $2 \pmod p$. Interchanging Φ and Φ' we find that either (i) $|\Phi| \equiv |\Phi'| \equiv 2 \pmod p$, $\Delta = \alpha \cup \Phi \cup \Phi'$, and H_α is 2-transitive on Φ and Φ'; or (ii) $p=2$, $|\Phi| \equiv |\Phi'| \equiv 1 \pmod 2$, $\Delta = \alpha \cup \Phi \cup \Phi'$, and H_α has rank 3 on Φ and Φ'.

Note that H is imprimitive on Δ. This follows from [9], Theorem 17.7 if (i) holds. If (ii) holds and $|\Phi'| \geqq |\Phi|$ then, since in this case $H_{\alpha\beta}$ is transitive on Φ', Φ' is an orbit of H_β and hence H_α is not maximal in H.

We may thus assume that the global stabilizer K of $\alpha \cup \Phi$ in H is transitive. Then $K^{\alpha \cup \Phi}$ is either 3-transitive or a transitive extension of a rank 3 group. Once again, as in the proof of Lemma 2 we obtain $3 \mid |K|$.

This contradiction proves the lemma.

We can now complete the proof of Theorem 2. Recall that each $|\Gamma_i(\alpha, \beta)| > 1$.

For $i=1, 2, 3$, $G_{\alpha\beta}$ acts faithfully on $\Gamma_i(\alpha, \beta)$ as a regular or Frobenius group. To see this, let p be a prime and $x \in G_{\alpha\beta}$ a p-element fixing $\geqq 2$ points of some $\Gamma_i(\alpha, \beta)$. By Lemma 2, x fixes >5 points, so by Lemma 3 $x=1$.

Thus, $G_{\alpha\beta}$ has a unique normal subgroup A regular on each $\Gamma_i(\alpha, \beta)$. Here $|A| = (n-2)/3$.

If n is even so is $|A|$, and all involutions fix 0 or 2 points. By an elementary lemma of Hering [5, p. 164, (2)], $G_{\alpha\beta}$ has at most two orbits on $\Omega-\{\alpha,\beta\}$, contradicting (b).

Thus, n and $|A|$ are odd. Let $x=(\alpha\beta)\cdots$ be an involution. Then x normalizes A.

Suppose that $|G_{\alpha\beta}|$ is odd. Then x inverts A. If $\gamma^x=\gamma$ then x centralizes $G_{\alpha\beta\gamma}$. Thus, there are $|A|$ involutions $(\alpha\beta)\cdots$. Counting the pairs consisting of an involution and one of its 2-cycles shows that the set I of involutions of G_α has $|A|$ elements. Then $I\cup\{1\}$ is not a group, since $1+(n-2)/3 \nmid n-1$. By Bender's theorem [2], G has a nontrivial normal subgroup of odd order, and by the Feit-Thompson Theorem [3] G has a regular normal elementary abelian subgroup. This contradicts the semiregularity of A on $\Omega-\{\alpha,\beta\}$. (It is not difficult to replace Bender's theorem in the above argument by Bender's generalization of Burnside's theorem on permutation groups of prime degree [2], Lemma 2.5, according to which either G_α is 2-transitive on I or $G_\alpha= N(A)_\alpha\, C(I)$.)

Consequently, we may assume that $|\Delta(x)|=5$. Choose i such that $|\Delta(x)\cap\Gamma_i(\alpha,\beta)|>1$.

Since $C_A(x)$ is transitive on $\Delta(x)\cap\Gamma_i(\alpha,\beta)$ ([9], Theorem 11.2), by (a) we have $|C_A(x)|=5$, that is, $\Delta(x)\subseteq\Gamma_i(\alpha,\beta)$. Write $A=C_A(x)\times[A,x]$, where x inverts $[A,x]$.

Let $\gamma\in\Delta(x)$. Since x normalizes $G_{\alpha\beta\gamma}$ it centralizes some involution $y\in G_{\alpha\beta\gamma}$. Here y inverts A, so that $xy\in G_\gamma$ centralizes $[A,x]$. Since $\gamma^{[A,x]}\subseteq\Delta(xy)$ and xy is an involution, $|[A,x]|=1$ or 5, according to whether $|\Delta(xy)|=1$ or $\neq 1$. Consequently, $|A|=5$ or 25 and $n=17$ or 77.

It is easy to eliminate these possibilities by considering the index of the normalizer of a Sylow 17- or 19-subgroup. Alternatively, note that the pointwise stabilizer of $\Delta(x)$ is now a 2-group of order $\leqq 8$. Since this is normalized by a Sylow 5-subgroup F of $C(x)$ it follows that $F\lhd C(x)$. However, we have seen that corresponding to each 2-cycle $(\alpha\beta)$ of x there is a group of order 5 in $C(x)_{\alpha\beta}$. Thus, F fixes $\Omega-\Delta(x)$ pointwise, which is ridiculous.

This completes the proof of Theorem 2.

3. Proof of Theorem 1

As already remarked in § 1, we need only consider the cases $k=3$ and 4. G is $(k-1)$-transitive (Livingstone and Wagner [8], Theorem 2(a)).

Let $k=3$. Let Φ be a set of 3 points. The global stabilizer of Φ induces a permutation group on the ordered triples of distinct points of Φ each of whose orbits has the same length $6/f$. Here $f=2,3$ or 6. Each orbit of G of ordered triples of distinct points has length $n(n-1)(n-2)/f$. Consequently, if $\alpha\neq\beta$ then each orbit of $G_{\alpha\beta}$ on $\Omega-\{\alpha,\beta\}$ has length $(n-2)/f$.

Suppose that $|G_\alpha|$ is odd. By a result of Bender [1], either (ii) holds or G is solvable. In the latter case, G has a regular normal elementary abelian subgroup N of order $n=2^d$. By [9], Theorem 10.4, $G_\alpha^{\Omega-\alpha}$ has a regular normal nilpotent subgroup M. Here M acts fixed-point-freely on N and hence is cyclic. Then $|N_{GL(d,2)}(M)|=(2^d-1)d$ implies that $2^d-2|6d$, so that (iii) holds.

We may thus assume that there is an involution of the form $(\alpha\beta)(\gamma)\cdots$. Then $f=3$, so that $3\nmid|G|$. Now Theorem 2 applies, whereas $n\geqq 2k=6$. This completes the proof when $k=3$.

Now let $k=4$. We first show that there is a set Φ of 4 points whose global stabilizer is transitive on Φ. By Livingstone and Wagner [8], Theorem 3, we can find a set Δ with $|\Delta|\geqq 4$ whose global stabilizer induces a 3-transitive group H on Δ such that each nontrivial element of H fixes $\leqq 3$ points of Δ. Certainly $4\,|\,|H|$. If H has an element of order 4 we can find the desired Φ. If $|\Delta|\leqq 9$ our assertion is also clear. Finally, if $|\Delta|>9$ and if H contains a Klein group then this Klein group has an orbit of length 4.

It follows that G is transitive on the pairs (α,Φ) with $\alpha\in\Phi$ and $|\Phi|=4$. Then $G_\alpha^{\Omega-\alpha}$ is 3-homogeneous but not 3-transitive. If G_α is as in (ii) or (iii), then (iv) holds. This completes the proof of Theorem 1.

References

1. Bender, H.: Endliche zweifach transitive Permutationsgruppen, deren Involutionen keine Fixpunkte haben. Math. Z. **104**, 175–204 (1968).
2. – Transitive Gruppen gerader Ordnung, in denen jede Involution genau einen Punkt festläßt. J. Algebra **17**, 527–554 (1971).
3. Feit, W., Thompson, J.G.: Solvability of groups of odd order. Pacific J. Math. **13**, 771–1029 (1963).
4. Gleason, A.M.: Finite Fano planes. Amer. J. Math. **78**, 797–807 (1956).
5. Hering, C.: Zweifach transitive Permutationsgruppen, in denen zwei die maximale Anzahl von Fixpunkten von Involutionen ist. Math. Z. **104**, 150–174 (1968).
6. Kantor, W.M.: 4-Homogeneous groups. Math. Z. **103**, 67–68 (1968). Correction Math. Z. **109**, 86 (1969).
7. – Automorphism groups of designs. Math. Z. **109**, 246–252 (1969).
8. Livingstone, D., Wagner, A.: Transitivity of finite permutation groups on unordered sets. Math. Z. **90**, 393–403 (1965).
9. Wielandt, H.: Finite permutation groups. New York: Academic Press 1964.

Prof. W. M. Kantor
Department of Mathematics
University of Oregon
Eugene, Oregon 97403
USA

(Received October 12, 1970)

Math. Z. 124, 266–272 (1972)

Local D-Rings

EBEN MATLIS

An integral domain R is said to have *property D*, or to be a D-ring, if every torsion-free R-module of finite rank decomposes into a direct sum of R-modules of rank one. A *local ring* is a commutative ring with only one maximal ideal. The main theorem of this paper will prove that the integral closure of a local D-ring is a maximal valuation ring.

We have proved elsewhere [7, Th. 4] that a Noetherian domain R has property D if and only if it is a ring of type II; i.e., R is a complete, Noetherian, local domain of Krull dimension one such that every ideal can be generated by two elements. Hence there exist Noetherian local D-rings which are not complete discrete valuation rings. At the end of this paper we will provide an example of a non-Noetherian local D-ring which is not a maximal valuation ring.

The proof of the following Lemma is entirely due to Bass, and we will omit it here (see [1, Prop. 7.5]).

Lemma 1. *Let R be a local D-ring. Then every finitely generated ideal of R can be generated by two elements.*

Lemma 2. *Let R be a local integral domain with maximal ideal M. Suppose that every finitely generated ideal of R can be generated by two elements. Then the following statements are true.*

(1) *Let x, y, z be elements of Q, the quotient field of R; and let A be the R-module generated by these elements. Then two of the elements x, y and z generate A.*

(2) *If C is an R-submodule of Q, then $\dim_{R/M} C/MC \leqq 2$.*

(3) *M^{-1} can be generated by two elements, where $M^{-1}=\{q\in Q \mid qM\subset R\}$.*

Proof. (1) Since A can be generated by two elements, we have $\dim_{R/M} A/MA \leqq 2$. Let $\bar{x}, \bar{y}, \bar{z}$ be the images of x, y, z in A/MA. Then these elements are a linearly dependent generating set for A/MA over R/M. Hence we can assume that $\bar{x}$ and $\bar{y}$ generate A/MA. Let B be the R-module generated by x and y. Then $A=B+MA$, and thus $M(A/B)=\dfrac{B+MA}{B}=A/B$. Therefore, by the Nakayama Lemma we have $A=B$. Thus A is generated by x and y.

(2) Suppose that $\dim_{R/M} C/MC>2$. Then there exist elements x, y, z in C such that their images $\bar{x}, \bar{y}, \bar{z}$ in C/MC are linearly independent over R/M.

By (1) we can assume that z is a linear combination of x and y. But then $\bar{x}, \bar{y}, \bar{z}$ are linearly dependent over R/M. This contradiction shows that $\dim_{R/M} C/MC \leqq 2$.

(3) If $MM^{-1} = R$, then M is a projective ideal of R, and hence a principal ideal of R, since R is a local ring. But then M^{-1} is also principal. Thus we can assume that $MM^{-1} = M$. Hence by (2) we have $\dim_{R/M} M^{-1}/M \leqq 2$. Therefore, there exists an element $x \in M^{-1}$ such that the images of 1 and x generate M^{-1}/M. Let B be the R-module generated by 1 and x. Then $M^{-1} = B + M = B$, and so M^{-1} can be generated by two elements.

Proposition 1. *Let R be an integral domain such that every finitely generated ideal of R can be generated by two elements. Then the integral closure of R is a Prüfer ring.*

Proof. Clearly every finitely generated torsion-free R-module of rank one is isomorphic to an ideal of R, and can thus be generated by two elements. This property is obviously inherited by every ring between R and its quotient field. Thus without loss of generality we can assume that R is an integrally closed local ring, and we must prove that R is a valuation ring.

Let M be the maximal ideal of R, and let Q be the quotient field of R. Let V be a valuation ring contained in Q that dominates R; that is $R \subset V$ and $m(V) \cap R = M$, where $m(V)$ is the maximal ideal of V. We will prove that $R = V$.

Suppose that $R \neq V$. If every unit of V is contained in R, then $V = R$. Hence there exists a unit x of V which is not in R. Let A be the R-module generated by 1, x and x^2. By assumption A can be generated by two elements. Hence by Lemma 2, two of the elements 1, x, x^2 generate A. Since R is integrally closed, 1 and x can not generate A.

In fact, 1 and x^2 generate A. For if x and x^2 generate A, then there exist elements a and b in R such that $1 = ax + bx^2$. If both a and b are in M, then $1 \in VM \subset m(V)$ which is a contradiction. Hence either a or b is not in M. Since 1 and x do not generate A, we must have $b \in M$. Thus a is a unit in R, and we see that 1 and x^2 generate A.

Thus we have shown that there exist elements c and d in R such that $x = c + dx^2$. We must have $d \in M$, since 1 and x do not generate A. However, $c \notin M$, since x is a unit in V. But then $1/x$ is integral over R, and hence $1/x \in M$. Therefore, $1 = x \cdot \dfrac{1}{x} \in VM \subset m(V)$. This contradiction shows that $R = V$.

Corollary 1. *Let R be an integral domain with property D. Then the integral closure of R is a Prüfer ring.*

Proof. This is an immediate consequence of Lemmas 1 and 3, and [5, Lemma 6.2].

Proposition 2. *Let R be an integral domain such that every finitely generated ideal of R can be generated by two elements. Let x be an element of the quotient* *field of R that is integral over R. Then there exist elements $a, b \in R$ such that $x^2 = ax + b$.*

Proof. We will suppose that x^2 is not an element of $Rx + R$ and arrive at a contradiction. Let $I = \{r \in R \mid r x^2 \in Rx + R\}$. Then I is contained in a maximal ideal M of R. If $x^2 \in R_M x + R_M$, then there exists an element $s \in R - M$ such that $s x^2 \in Rx + R$. Since $s \notin I$, we have that $x^2 \notin R_M x + R_M$. Now every finitely generated ideal of R_M can be generated by two elements and x is integral over R_M. Thus without loss of generality we can assume that R is a local ring with maximal ideal M.

Let A be the R-module generated by 1, x, and x^2. By Lemma 2, two of these elements generate A. Since $x^2 \notin Rx + R$, 1 and x do not generate A. As in the proof of Proposition 1, we have in all cases that 1 and x^2 generate A. Therefore we have $x \in M x^2 + R$. An easy induction proves that for any integer $n > 0$, $x, x^2, \ldots, x^{n-1}$ are all elements of $M x^n + R$. Since x is integral over R, there exists an equation of the form

$$x^n = b_{n-1} x^{n-1} + \cdots + b_i x + b_0,$$

where $b_i \in R$. Thus we have $x^n \in M x^n + R$. But this implies that $x^n \in R$, and hence $x \in R$. Thus 1 generates A, and this contradiction shows that $x^2 \in Rx + R$.

Definition. Let R be an integral domain, and A an R-module. We denote the dual of A by $A^* = \operatorname{Hom}_R(A, R)$. There is a canonical R-homomorphism $\varphi: A \to A^{**}$ given by $\varphi(x)(f) = f(x)$ for all $x \in A$ and $f \in A^*$. A is called a *reflexive* R-module if φ is an isomorphism. R is called a *completely reflexive ring* if every reduced, torsion-free R-module of finite rank is reflexive.

Theorem 1. *Let R be a local D-ring with maximal ideal M. If M is not a principal ideal of R, and if $M^{-1} \neq R$, then R is a completely reflexive ring.*

Proof. Let Q be the quotient field of R, and let $K = Q/R$. We will first prove that K is an essential extension of M^{-1}/R. For this purpose it will be sufficient to take an R-module A such that $R \subsetneqq A \subset Q$, and show that $M^{-1} \subset A$. We will suppose that $M^{-1} \not\subset A$ and arrive at a contradiction. By Lemma 2, M^{-1}/R is a simple R-module, and thus we have that $A \not\subset M^{-1}$. Thus there exists an element $y \in A$ such that $y \notin M^{-1}$. By Lemma 2 there exists an element $x \in M^{-1}$ such that 1 and x generate M^{-1} as an R-module. Since $R \subset A$, we have $x \notin A$.

Let B be the R-module generated by 1, x and y. By Lemma 2, two of these elements generate B. Since $x \notin A$ and $y \notin M^{-1}$, it follows that there exist elements $a, b \in M$ such that $1 = ax + by$. Now M is not a principal ideal of R, and hence $M^{-1} M = M$. Therefore $ax \in M$, and $by = 1 - ax$ is a unit in R. Thus there exists an element $r \in M$ such that $y = 1/r$. Since $Rr \subset M$, we have $M^{-1} \subset Ry \subset A$. This contradiction shows that K is an essential extension of M^{-1}/R.

Since M^{-1}/R is a simple R-module, we have $M^{-1}/R \cong R/M$. Thus $K \subset E(R/M)$, the injective envelope of R/M. We will prove that $K = E(R/M)$. Now every non-zero principal ideal of R is the annihilator of an element of K, hence of an element of $E(R/M)$. But $E(R/M)$ is an indecomposable injective module, and thus every principal ideal of R is meet irreducible by [2, Th. 2.4].

Let $I \neq 0$ be an ideal of R, and consider an extension C of R by I:

$$0 \to R \to C \to I \to 0.$$

C is a torsion-free R-module of rank two, and R is a D-ring. Thus we have that $C = C_1 \oplus C_2$, where C_1, C_2 are torsion-free R-modules of rank one. Let J_1 and J_2 be the images of C_1 and C_2 in I. Then $J_1 + J_2 = I$ and $J_1 \cap J_2 \cong R$. Hence $J_1 \cap J_2$ is a principal ideal of R. Since we have shown that principal ideals of R are irreducible, we can assume that $J_1 \subset J_2$. But then $I = J_2$, which proves that $C = C_2 + R$. Since C_2 has rank one and maps onto I, we must have $C_2 \cap R = 0$; i.e., $C = C_2 \oplus R$.

Therefore, the previous exact sequence splits, which shows that $\mathrm{Ext}^1_R(I, R) = 0$. From this it follows that inj. $\dim_R R = 1$. Thus K is an injective R-module, and since $K \subset E(R/M)$, we have $K = E(R/M)$. Now a local D-ring is complete in the R-topology by [7, Th. 1]. Thus by [6, Prop. 5.1] R is a completely reflexive ring.

Remark. As a comment on Theorem 1, we note that if R is a local D-ring whose maximal ideal is principal, then it follows easily from [6, Th. 5.5] that R is a completely reflexive ring if and only if R is a maximal valuation ring.

Theorem 2. *Let R be a local D-ring, and let F be the integral closure of R. If $R \neq F$, then there exists a ring S such that $R \subsetneqq S \subset F$ and S is a completely reflexive local D-ring.*

Proof. Let M be the maximal ideal of R, let $x \in F - R$, and let $S = R + Mx$. By Lemma 1 every finitely generated ideal of R can be generated by two elements, and thus by Proposition 2, S is a ring. Let $N = M + Mx$; then N is a proper ideal of S. The elements of S that are not in N are units in F, and hence units in S, since F is the integral closure of S. Thus S is a local ring with maximal ideal N. By [5, Lemma 6.2] S is a local D-ring.

Let $T = R + Rx$; by Proposition 2, T is a ring. Since $x \notin S$, we have $T \neq S$. Let $N^{\#}$ be the inverse of N relative to S; $N^{\#} = \{q \in Q \mid qN \subset S\}$, where Q is the quotient field of R. Since $TN = N$, we have $T \subset N^{\#}$, and thus $N^{\#} \neq S$. If N is a principal ideal of S, then there exists an element $a \in S$ such that $N = Sa$. But then $Sa = TN = Ta$, and we have $S = T$. This contradiction shows that N is not a principal ideal of S. We have verified that S satisfies the hypotheses of Theorem 1, and thus S is a completely reflexive, local D-ring.

Theorem 3. *Let R be a local D-ring. Then the integral closure of R is a maximal valuation ring.*

Proof. Let F be the integral closure of R. If $F = R$, then F is a local ring. If $F \neq R$, then by Theorem 2, there exists a completely reflexive local ring S such that $R \subset S \subset F$. By [6, Th. 5.4] every ring extension of S in its quotient field is a local ring. Thus in all cases F is a local ring.

By [5, Lemma 6.2] F has property D, and by Proposition 1, F is a Prüfer ring. A local Prüfer ring is a valuation ring; hence by [7, Th. 2] F is a maximal valuation ring.

Corollary 2. *Let R be an integral domain with property D. Let F be the integral closure of R and P a prime ideal of R. Then there is only one prime ideal of F lying over P.*

Proof. R_P is an extension ring of R inside the field of quotients of R, and hence is a local D-ring by [5, Lemma 6.2]. The integral closure of R_P is F_P. Thus by Theorem 3, F_P is a maximal valuation ring. From this it follows directly that F has only one prime ideal lying over P.

The following theorem is a partial converse of Theorem 3.

Theorem 4. *Let R be a local integral domain with maximal ideal M, and let F be the integral closure of R. If F is a maximal valuation ring that is generated as an R-module by two elements, and if $MF=F$, then R is a local D-ring.*

Proof. Since a maximal valuation ring has property D by [7, Th. 2], we can assume that $R\neq F$. We will prove first that $F=M^{-1}$. Since $MF=M$, we have $F\subset M^{-1}$.

Now M is not a principal ideal of R. For if $M=Ra$, for some $a\in R$, then $Ra=FM=Fa$, and hence $R=F$, which contradicts our assumption. Thus M is not a principal ideal of R. Since invertible ideals in a local ring are principal ideals, we have $M^{-1}M=M$, and thus M^{-1} is a ring. Furthermore, since $F\subset M^{-1}$, and F is a valuation ring, M^{-1} is also a valuation ring.

Let $m(F)$ and $m(M^{-1})$ be the maximal ideals of F and M^{-1}, respectively. Then we have $m(M^{-1})\subset m(F)$. If $m(M^{-1})=m(F)$, then $M^{-1}=F$, which is our contention. Hence we will assume that $m(M^{-1})\neq m(F)$, and arrive at a contradiction.

We have $M\subset m(M^{-1})\subsetneqq m(F)\subsetneqq F$. Since F can be generated by two elements, we have $\dim_{R/M} F/M\leqq 2$. Hence the above chain of inequalities shows that $M=m(M^{-1})$. Thus M is a prime ideal of F. Let x be an element of $m(F)$ such that $x\notin M$. Then $x^2\notin M$, and we have $M\subsetneqq Fx^2\subsetneqq Fx\subsetneqq F$. This contradicts $\dim_{R/M} F/M\leqq 2$ and proves that $F=M^{-1}$.

Since $M^{-1-1}=M$, we have $F=F^{-1-1}$. Thus F is a reflexive R-module. Since F can be generated by two elements, we have by [6, Prop. 3.1] that $\mathrm{Ext}_R^1(F,R)=0$. Therefore, from the exact sequence:

$$0\to R\to Q\to K\to 0$$

where Q is the quotient field of R, and $K=Q/R$, we obtain the exact sequence

$$0\to \mathrm{Hom}_R(F,R)\to \mathrm{Hom}_R(F,Q)\to \mathrm{Hom}_R(F,K)\to 0.$$

Now $\mathrm{Hom}_R(F,R)\cong F^{-1}=M$, and $\mathrm{Hom}_R(F,Q)\cong Q$. Thus $\mathrm{Hom}_R(F,K)$ is isomorphic to Q/M as an F-module. Therefore, by [3, Th. 9] $\mathrm{Hom}_R(F,K)$ is an injective F-module. Since F is finitely generated over R, it follows from an unpublished theorem of David Eisenbud that K is an injective R-module.

Since $\dim_{R/M} M^{-1}/M \leqq 2$, it follows that M^{-1}/R is a simple R-module. Thus $M^{-1}/R \cong R/M$. Because K is an injective R-module, we then have $E(R/M) \subset K$, where $E(R/M)$ is the injective envelope of R/M. Therefore R is a reflexive ring by [5, Th. 2.1].

Now F is a cotorsion F-module by [3, Th. 9] and thus F is a cotorsion R-module by [4, Cor. (p. 575)]. Since F is isomorphic to an ideal of R, it follows that R is a cotorsion R-module. Thus R is complete in the R-topology. Hence by [6, Prop. 5.1] R is a completely reflexive ring.

Let A be an indecomposable torsion-free R module of finite rank. We want to show that rank $A=1$. We can assume that A is a reduced R-module. Since R is a completely reflexive ring, we have that $A \cong A^{**}$.

Let I be the trace ideal of A; by definition:

$$I = \left\{ \sum_{i=1}^{n} f_i(x_i) \mid f_i \in A^* \text{ and } x_i \in A \right\}.$$

If $I=R$, there exist elements $f \in A^*$ and $x \in A$ such that $f(x)=1$. Define $g\colon R \to A$ by $g(r)=rx$ for all $r \in R$. Then gf is the identity on R, and hence Rx is a direct summand of A. Since A is indecomposable, we have $A=Rx$; and A has rank one. Thus we can assume that $I \subset M$.

We now have $F=M^{-1} \subset I^{-1}$. Let $v \in F$ and $x \in A$, and define $vx \in A^{**}$ by $(vx)(f)=vf(x)$ for all $f \in A^{**}$. Since $I^{-1}I \subset R$, vx is well defined. This makes A^{**} into an F-module. Since $A \cong A^{**}$, A is a torsion free F-module of finite rank. Now F has property D by [7, Th. 2] and thus A is a direct sum of F-modules of rank one. Therefore A is a direct sum of R-modules of rank one, and hence rank $A=1$.

Thus we have proved that R is a local D-ring.

Example. We will now produce an example of a non-Noetherian local D-ring that is not a maximal valuation ring.

Let k be a field and X an indeterminate over k. Let G be the Abelian group $Z \times Z$ ordered lexicographically (where Z is the group of integers). Let F be the ring of all formal power series $\sum_{\alpha} b_\alpha X^\alpha$, where $b_\alpha \in k$, $\alpha \in G$, $\alpha \geqq (0,0)$, and the sets of exponents $\{\alpha\}$ are well ordered. By [8, Ch. 2, Cor. Th. 8] F is a maximal valuation ring and the maximal ideal N of F is the set of power series with exponents $\alpha > (0,0)$.

Let R be the subset of F consisting of those power series where the term $X^{(0,1)}$ has zero coefficient. Then R is a local ring with maximal ideal M consisting of those power series with exponents $\alpha > (0,1)$. We have $FM=M$, and F is generated as an R-module by the two elements $1=X^{(0,0)}$ and $X^{(0,1)}$. Thus by Theorem 4, R is a local D-ring. R is not a maximal valuation ring since F is the integral closure of R, and $F \neq R$. F is a rank two valuation ring, and thus is not Noetherian. Hence R is not a Noetherian ring.

References

1. Bass, H.: On the ubiquity of Gorenstein rings. Math. Z. **82**, 8–28 (1963).
2. Matlis, E.: Injective modules over Noetherian rings. Pacific J. Math. **8**, 511–528 (1958).
3. — Injective modules over Prüfer rings. Nagoya Math. J. **15**, 57–69 (1959).
4. — Some properties of Noetherian domains of dimension 1. Canadian J. Math. **13**, 569–586 (1961).
5. — Decomposable modules. Trans. Amer. Math. Soc. **125**, 147–179 (1966).
6. — Reflexive domains. J. Algebra **8**, 1–33 (1968).
7. — The decomposability of torsion-free modules of finite rank. Trans. Amer. Math. Soc. **134**, 315–324 (1968).
8. Schilling, O.: The theory of valuations. Math. Surveys No. IV, Amer. Math. Soc., 1950.

Prof. E. Matlis
Department of Mathematics
Northwestern University
Evanston, Illinois
USA

(Received April 26, 1971)

Math. Z. 124, 273 – 288 (1972)

Rechtskomplementäre Halbgruppen. Axiome, Polynome, Kongruenzen

Bruno Bosbach

0. Einleitung

Eine *Halbgruppe S* heißt *rechtskomplementär*, wenn sie $a \in a S$ erfüllt und zu jedem Paar a, b ein eindeutig bestimmtes Element $a*b$ besitzt mit der Eigenschaft

$$a x \in b S \Leftrightarrow x \in (a*b) S.$$

Rechtskomplementäre Halbgruppen sind *gleichungsdefiniert* durch das Axiomensystem

$$\text{(A.1)} \qquad a(a*b) = b(b*a),$$

$$\text{(A.2)} \qquad a b * c = b * (a * c),$$

$$\text{(A.3)} \qquad a(b*b) = a,$$

wie in [3] gezeigt wurde. Dual zur rechtskomplementären Halbgruppe ist die *linkskomplementäre* Halbgruppe erklärt. Eine Halbgruppe heißt *komplementär*, wenn sie sowohl rechtskomplementär bezüglich $*$ als auch linkskomplementär bezüglich : ist und zusätzlich $a(a*b) = (b:a)\, a$ erfüllt. Die Theorie der (rechts-) komplementären Halbgruppen wurde in [2 – 5] begründet und entwickelt. Insbesondere zeigt es sich, daß jede rechtskomplementäre Halbgruppe einen *Halbverband* bildet bei $a \leqq b :\Leftrightarrow b \in a S$ mit $a \cup b = a(a*b)$ für alle a, b und $a*a$ als Eins.

Als Beispiel einer rechtskomplementären Halbgruppe sei die Menge der Ordinalzahlen mit Realisationen von höchstens abzählbarer Mächtigkeit als Halbgruppe bezüglich der Ordinalzahladdition genannt. Zu den komplementären Halbgruppen gehören u.a. die *Boolesche Algebra, der Verbandsgruppenkern* sowie *die Zerlegungshalbgruppe*, insbesondere also jede *Potenzmenge*, betrachtet bezüglich der Vereinigung, *der Bereich der nicht negativen ganzen Zahlen*, betrachtet bezüglich der Addition, sowie *der Bereich der natürlichen Zahlen*, betrachtet bezüglich der Multiplikation.

Ziel dieser Note ist die Beantwortung einiger offener Fragen hinsichtlich der *Grundlegung* und in bezug auf die *Kongruenzen* (rechts-)komplementärer Halbgruppen, und zwar werden im einzelnen folgende Sätze bewiesen:

1. *Jede Klasse rechtskomplementärer Halbgruppen, die sich mit endlich vielen Gleichungen beschreiben läßt, kann auch mit zwei Gleichungen beschrieben werden.*

2. *Jede Klasse komplementärer Halbgruppen, die sich mit endlich vielen Gleichungen beschreiben läßt, kann auch mit einer (einzigen) Gleichung beschrieben werden.*

3. *Die Kongruenzen der rechtskomplementären Halbgruppe sind* 3-*vertauschbar, nicht aber* (2-)*vertauschbar.*

4. *Die Kongruenzen der rechtskomplementären Halbgruppe bilden einen distributiven Verband, doch erfüllen sie nicht die Bedingung des Chinesischen Restsatzes* (vgl. [15]), *oder in anderer Sprechweise: sie sind* 3-*distributiv, nicht aber* 2-*distributiv.*

Analog läßt sich zeigen, daß die Kongruenzen der rechtskomplementären Halbgruppe 3-modular, jedoch nicht notwendig 2-modular sind, was allerdings nach Resultaten aus [6] und [12] mit den oben angegebenen Sätzen implizit erledigt ist.

Wir runden die Untersuchungen dieser Note ab durch *Polynome* für komplementäre Halbgruppen, deren Existenz sich aus den Ergebnissen der Arbeit [4] ergibt.

Angeregt wurden die nachfolgenden Ausführungen von Herrn Rudolf Wille, der den Autor in Gesprächen wiederholt über Fragestellungen und Resultate der Universellen Algebra und Verbandstheorie informierte, die z.T. ihren Niederschlag erst in preprints gefunden hatten. Hierfür sei Herrn Wille an dieser Stelle gedankt.

1. Zur Grundlegung der rechtskomplementären Halbgruppe

Satz 1. *Eine Algebra $S(\cdot, *)$ ist genau dann eine rechtskomplementäre Halbgruppe, wenn sie den beiden Axiomen genügt:*

$$a(b*b)=a, \tag{B.1}$$

$$u(u*(ab*c))=(b*(a*c))((ab*c)*u). \tag{B.2}$$

Beweis. Wir setzen $u=ab*c$ und erhalten

$$\text{(A.2)}\quad ab*c=b*(a*c). \tag{1}$$

Wir setzen $b=x*x$ und erhalten

$$u(u*(a*c))=(a*c)((a*c)*u), \tag{2}$$

womit wir weiter zeigen können

(3)
$$\begin{aligned} v*v &= (v*v)(a*a) \\ &= (v*v)\big(a(v*v)*a\big) \\ &= (v*v)\big((v*v)*(a*a)\big) && (1) \\ &= (a*a)\big((a*a)*(v*v)\big) && (2) \\ &= a*a \\ &:= e. \end{aligned}$$

Hiernach folgen sukzessive die Hilfssätze:

(4)
$$\begin{aligned} (e*a)*e &= (e*a)*(e*e) \\ &= e(e*a)*e && (1) \\ &= a(a*e)*e && (2) \\ &= (a*e)*(a*e) && (1) \\ &= e, \end{aligned}$$

(5)
$$\begin{aligned} e*a &= (e*a)\big((e*a)*e\big) && (4) \\ &= e\big(e*(e*a)\big) && (2) \\ &= e(e*a), \end{aligned}$$

(6)
$$\begin{aligned} (e*a)*a &= e(e*a)*a && (5) \\ &= (e*a)*(e*a) && (2) \\ &= e, \end{aligned}$$

(7)
$$\begin{aligned} e*a &= (e*a)\big((e*a)*a\big) && (6) \\ &= a\big(a*(e*a)\big) \\ &= a(e\,a*a), \end{aligned}$$

(8)
$$\begin{aligned} e*e\,a &= (e\,a)\big(e(e\,a)*e\,a\big) && (7) \\ &= (e\,a)\big(e\,a*(e*e\,a)\big) \\ &= (e\,a)\big(a*(e*(e*e\,a))\big) \\ &= (e\,a)\big(a*(e*e\,a)\big) \\ &= (e\,a)(e\,a*e\,a) \\ &= e\,a, \end{aligned}$$

(9) $$a(a*eb)=a(a*(e*eb)) \quad (8)$$
$$=(e*eb)((e*eb)*a)$$
$$=(eb)(eb*a),$$

(10) $$e(bc)=(e(bc))((eb)c*(eb)c)$$
$$=(e(bc))(c*(eb*(eb)c))$$
$$=(e(bc))(c*(b*(e*(eb)c)))$$
$$=(e(bc))(bc*(e*(eb)c))$$
$$=(e(bc))(e(bc)*(eb)c)$$
$$=((eb)c)((eb)c*e(bc)) \quad (9)$$
$$=((eb)c)(c*(eb*e(bc)))$$
$$=((eb)c)(c*(b*(e*e(bc))))$$
$$=((eb)c)(bc*(e*e(bc)))$$
$$=((eb)c)(e(bc)*e(bc))$$
$$=(eb)c,$$

(11) $$ea=ea(ea*ea) \quad (10)$$
$$=ea(a*(e*ea))$$
$$=e(a(a*ea)) \quad (8)$$
$$=e(ea(ea*a)) \quad (9)$$
$$=(ee)(a(ea*a)) \quad (10)$$
$$=e(a(ea*a))$$
$$=e(e*a) \quad (7)$$
$$=e*a, \quad (5)$$

(12) $$ea=e*a \quad (11)$$
$$=a(ea*a) \quad (7)$$
$$=a((e*a)*a)$$
$$=ae \quad (6)$$
$$=a,$$

(13) $$e*a=ea$$
$$=a,$$

(14) $$a*e=(e*a)*e \qquad (4)$$
$$=e,$$

(15) $$(A.3) \quad a(a*b)=a\big(a*(e*b)\big)$$
$$=(e*b)\big((e*b)*a\big)$$
$$=b(b*a).$$

Aus Satz 1 ergibt sich fast unmittelbar der folgende

Satz 2. *Sei $\mathfrak{A}$ eine Klasse von Algebren, die sich mit endlich vielen Gleichungen vermöge der Operationen O_ν beschreiben läßt. Lassen sich dann die Axiome* (A.1), (A.2), (A.3) *bezüglich zweier Operationen O_1, O_2 oder aber bezüglich zweier aus den O_ν abgeleiteter Operationen verifizieren, so ist es möglich, die Klasse $\mathfrak{A}$ mit zwei Gleichungen zu charakterisieren.*

Beweis. Aus dem Beweis zu Satz 1 folgt:

(16) $$ab=e \Leftrightarrow a=e=b,$$

denn

$$ab=e \Rightarrow a=e*a=ab*a=b*(a*a)=b*e=e$$

und

$$ab=e \;\&\; a=e \Rightarrow b=eb=ab=e$$

sowie

(17) $$a=b \Leftrightarrow (a*b)(b*a)=e,$$

denn

$$a=b \Rightarrow a*b=e=b*a \Rightarrow (a*b)(b*a)=e$$

und

$$(a*b)(b*a)=e \Rightarrow a*b=e=b*a \Rightarrow a=a(a*b)=b(b*a)=b$$

und daraus resultiert, daß jedes endliche System von Gleichungen $f_\nu=g_\nu$ $(\nu=1,\dots,n)$ komprimiert werden kann zu einer einzigen Gleichung

$$p=\prod_{1}^{n}(f_\nu*g_\nu)(g_\nu*f_\nu)=e,$$

wobei diese Identität gleichwertig ist mit $x*p=e$. Daher können wir (B.2) durch Juxtaposition von $v*p$ spezialisieren zu

(B.2′) $$u\big(u*(ab*c)\big)=\big((b*(a*c))((ab*c)*u)\big)(v*p),$$

womit der Beweis abgeschlossen ist, da wir (B.2) aus (B.2′) und (B.1) durch Ersetzung von v mittels p, und $p=e$ für $u=v=a=b=c=e$ erhalten.

Es ist klar, daß in Sonderfällen (B.2′) bezüglich seiner Länge und der Anzahl seiner Variablen reduziert werden kann. Von besonderem Interesse ist diesbezüglich natürlich die kommutative komplementäre Halbgruppe. Hier können wir zeigen:

Satz 1′. *Eine Algebra $S(\bullet, *)$ ist genau dann eine kommutative komplementäre Halbgruppe, wenn sie den beiden Axiomen genügt:*

(B.1) $$a(b*b)=a,$$

(B.2k) $$(a\,b*c)\big((u(u*v))\big)=\big(v(v*u)\big)\big(b*(a*c)\big).$$

Beweis. Setzen wir $b=a*a$ und $c=a$, so erhalten wir

(18) $$(a*a)\big(u(u*v)\big)=v(v*u),$$

und es folgt für $u=w*w=v$ die Gleichung:

(19) $$a*a=w*w$$
$$:=e.$$

Setzen wir nun in (18) $u=v$, so erhalten wir

(20) $$e\,u=u,$$

woraus mit $a=b=c=e$ in (B.2k) die Gleichung

(22) $$u(u*v)=v(v*u)$$

und mit $u=v=e$ in (B.2k) die Gleichung

(23) $$a\,b*c=b*(a*c)$$

gefolgert werden kann.

Setzen wir schließlich $a=b=e$ und $u=v$ in (B.2k), so resultiert hieraus die Kommutativität.

Durch eine leichte Abänderung des Systems (B.1), (B.2k) – nämlich bei Ersetzung von $a\,b*c$ durch $a\,b*c^2$ – erhält man ein Axiomensystem für den *Brouwerschen Halbverband.* Ein System für den *abelschen Verbandsgruppenkern* liefert (B.2) zusammen mit (VK): $a*b\,a=b$. Da diesen Tatsachen jedoch keine prinzipielle Bedeutung zukommt, verzichten wir hier auf die Ausführung der Beweise.

2. Zur Grundlegung der komplementären Halbgruppe

Angeregt durch eine Beweisidee von McKenzie [11] können wir im folgenden zeigen:

Satz 3. *Sei $\mathfrak{A}$ eine Klasse komplementärer Halbgruppen, die sich mit endlich vielen Gleichungen beschreiben läßt. Dann kann man $\mathfrak{A}$ auch mittels einer einzigen Gleichung charakterisieren.*

Zunächst ein Lemma:

(24) *Es seien* $f_1, \ldots, f_m$ *eindeutige Abbildungen einer Menge M in sich selbst und es gelte:* $f_1 f_2 \cdots f_{m-1} f_m f_{m-1} \cdots f_2 f_1(a) = a$. *Dann folgt:*

$$f_n f_{n+1} \cdots f_m \cdots f_{n+1} f_n f_{n-1} f_{n-2} \cdots f_2 f_1 f_1 f_2 \cdots f_{n-2} f_{n-1}(a) = a$$

sowie

$$f_{n-1} f_{n-2} \cdots f_2 f_1 f_1 f_2 \cdots f_{n-2} f_{n-1} f_n f_{n+1} \cdots f_m \cdots f_{n+1} f_n(a) = a,$$

insbesondere also $f_\nu(b) = f_\nu(c) \Rightarrow b = c$.

Offenbar gilt $fgf(a) = a \Rightarrow fg(a) = fgfgf(a) = gf(a)$ und damit unsere Behauptung für $n = 1$. Zum Zwecke des weiteren Beweises setzen wir

$$f_{n+1} \cdots f_m \cdots f_{n+1}(a) = g(a)$$

$$f_{n-2} \cdots f_2 f_1 f_1 f_2 \cdots f_{n-2}(a) = h(a)$$

$$f_n g f_n f_{n-1} h f_{n-1}(a) = l(a)$$

$$f_{n-1} h f_{n-1} f_n g f_n(a) = r(a)$$

und erhalten wegen $r(a) = a = l(a)$

$$\begin{aligned} g f_n f_{n-1} h f_{n-1}(a) &= r g f_n f_{n-1} h f_{n-1}(a) \\ &= f_{n-1} h f_{n-1} f_n g(a), \end{aligned}$$

wie man bei Ersetzung von g und h bestätigt. Weiter geht die linke Seite in die identische Abbildung über, wenn man f_n vorschachtelt, während die rechte Seite in die identische Abbildung übergeht, wenn man $f_n(a)$ einschachtelt. Da diese Aussagen notwendig auch über Kreuz gelten, ist damit nach dem Prinzip der vollständigen Induktion alles bewiesen.

Als nächstes führen wir zwei Abkürzungen ein, und zwar setzen wir

$$a \vee b \quad \text{für } a(a*b)$$

bzw.

$$a \wedge b \quad \text{für } (b:(a*b))((b:(a*b))*(a:(b*a))),$$

um hiernach die Polynome zu bilden:

$p_1 = (x_1 \wedge x) \vee (x \wedge x)$,

$p_2 = ((x \vee x) \wedge x_2) \vee (x \wedge x)$,

$p_3 = ((x \vee x) \wedge (x \vee x)) \vee (x_3 \wedge x)$,

$p_4 = ((x \vee x) \wedge (x \vee x)) \vee ((x \vee x) \wedge x_4)$,

$p_5 = (x_5 \vee x) \wedge (x \vee x),$

$p_6 = \big((x \wedge x) \vee x_6\big) \wedge (x \vee x),$

$p_7 = \big((x \wedge x) \vee (x \wedge x)\big) \wedge (x_7 \vee x),$

$p_8 = \big((x \wedge x) \vee (x \wedge x)\big) \wedge \big((x \wedge x) \vee x_8\big),$

$p_9 = \big((x_9 \wedge x) \vee (x \wedge y_9)\big) \vee x,$

$p_{10} = x(x_{10} * x_{10}),$

$p_{11} = (x_{11} * x_{11})\, x,$

$p_{12} = \big(x\big((x_{12}\, y_{12} * z_{12}) * (y_{12} * (x_{12} * z_{12}))\big)\big)\big((v_{12} * (u_{12} * w_{12})) * (u_{12}\, v_{12} * w_{12})\big),$

$p_{13} = x f(w_{13,1} \dots w_{13,m})$ mit f gleich p aus (B.2′).

Es ist klar, daß die angeführten Polynome alle als Polynome in den Operationen $*$, :, und $\bullet$ und den Variablen $x, x_1, \dots, x_{12}, y_{12}, z_{12}, u_{12}, v_{12}, w_{13,1} \dots w_{13,m}$ aufgefaßt werden können und daß p_1 bis p_{13} übergehen in Abbildungen f_1 bis f_{13}, wenn man in p_ν jeweils $x_1 \dots\dots w_{13,m}$ mit Konstanten $u_\nu \dots\dots w_{m\nu}$ belegt. Geht man weiter hin und belegt p_ν einmal mit $\bar{u}_\nu \dots\dots \bar{w}_{m\nu}$ zu $\bar{f}_\nu$ und ein anderes Mal mit $\bar{\bar{u}}_\nu \dots\dots \bar{\bar{w}}_{m\nu}$ zu $\bar{f}_\nu$, so erhält man durch Vor- und Einschachteln aus

$$f_1 \dots \bar{f}_\nu \dots f_{13} \dots f_\nu \dots f_1(a) = a = f_1 \dots \bar{\bar{f}}_\nu \dots f_{13} \dots f_\nu \dots f_1(a)$$

die Gleichung

(25) $$\bar{f}_\nu(a) = \bar{\bar{f}}_\nu(a).$$

Dies ist der Schlüssel zum Beweis. Denn die Forderung

(A) $$\bar{f}_1 \dots \bar{f}_{12} f_{13} \bar{\bar{f}}_{12} \dots \bar{\bar{f}}_1(a) = a$$

ist offenbar notwendig. Sie reicht aber auch hin: Insofern sie die Unabhängigkeit der Polynomwerte von den indizierten Variablen impliziert, gilt zunächst

(26) $$f_1(a) = f_2(a) = f_3(a) = f_4(a) = f_1(a \vee a),$$

also nach (24)

(27) $$a \vee a = a.$$

Analog erhalten wir

(28) $$a \wedge a = a$$

und damit

(29) $$f_1(a) = \cdots\cdots\cdots\cdots = f_9(a) = a.$$

Aufgrund von (29) gewinnen wir weiter

(30)
$$\begin{aligned}(a \vee b) \vee (b \vee a) &= \big((a \wedge (b \vee a)) \vee ((b \vee a) \wedge b)\big) \vee (b \vee a)\\ &= b \vee a\\ \Rightarrow \qquad a \vee b &= (a \vee b) \wedge \big((a \vee b) \vee (b \vee a)\big)\\ &= \big((b \vee a) \vee (a \vee b)\big) \wedge (b \vee a)\\ &= b \vee a.\end{aligned}$$

Die Kommutativität bezüglich $\vee$ sichert uns die Gleichung

(31)
$$(\text{A.}1) \quad a(a*b) = b(b*a).$$

Weiter gilt wegen $\bar{f}_{10}(a) = \bar{\bar{f}}_{10}(a)$ die Gleichung $a(b*b) = a(a*a)$, also

(32)
$$(\text{A.}3) \quad a(b*b) = a.$$

Dann folgt aber

(33)
$$a*a = (a*a)(b*b) = (b*b)(b*b) = b*b$$

wegen $\bar{f}_{10}(a) = \bar{\bar{f}}_{10}(a)$ und $\bar{f}_{11}(a) = \bar{\bar{f}}_{11}(a)$, so daß mit $a*a = b*b := e$ die beiden Gleichungen

(34)
$$e \cdot e = e \quad \text{und} \quad e*e = e$$

erfüllt sind. Das wiederum bedeutet

(35)
$$\begin{aligned} &a\big((x\,y*z)*(y*(x*z))\big) = a\\ \text{und} \quad &a\big((y*(x*z))*(x\,y*z)\big) = a.\end{aligned}$$

Denn setzen wir in $f_{12}(a)$ $u_{12} = v_{12} = w_{12} = x_{12} = y_{12} = z_{12} = e$, so erhalten wir a. Setzen wir $u_{12} = v_{12} = w_{12} = e$ und $x_{12} = x$, $y_{12} = y$ sowie $z_{12} = z$, so erhalten wir die erste Zeile. Setzen wir $x_{12} = y_{12} = z_{12} = e$ und $u_{12} = x$, $v_{12} = y$ sowie $z_{12} = z$, so erhalten wir die zweite Zeile. Dann folgt aber für $f = x\,y*z$ und $g = y*(x*z)$ die Gleichung

(36)
$$f = f(f*g) = g(g*f) = g,$$

womit auch das Axiom

(37)
$$(\text{A.}2) \quad a\,b*c = b*(a*c)$$

verifiziert ist. Somit ergibt sich $(b*b)\,a = a$ und damit $f_{10}(a) = a$, so daß unsere Ausgangsgleichung zusammenschrumpft zu

(38)
$$f_{13}(a) = a,$$

womit alles gezeigt ist.

Aus Satz 3 ergeben sich mehrere Korollare. Das wesentlichste unter ihnen, das alle übrigen umfaßt, ist

Korollar 1. *Sei $\mathfrak{A}$ eine Klasse von Algebren, die sich mit endlich vielen Gleichungen beschreiben läßt. Lassen sich dann bezüglich zweier aus den definierenden Operationen ableitbarer Operationen die Axiome der komplementären Halbgruppe verifizieren, so ist es möglich, die Klasse $\mathfrak{A}$ mit einer einzigen Gleichung zu charakterisieren.*

Der Beweis verläuft analog zu dem soeben geführten.

Speziell gelten demnach:

Korollar 2. *Die kommutative komplementäre Halbgruppe läßt sich mit einer einzigen Gleichung charakterisieren.*

Korollar 3. *Der Boolesche Verband (Ring) läßt sich mit einer Gleichung charakterisieren.*

Korollar 4. *Der Brouwersche (Halb-)Verband läßt sich mit einer Gleichung charakterisieren.*

Korollar 5. *Der (abelsche) Verbandsgruppenkern läßt sich mit einer Gleichung charakterisieren.*

Korollar 6. *Die normale komplementäre Halbgruppe läßt sich mit einer Gleichung charakterisieren.*

Insbesondere ergibt sich aus Satz 2 noch der auf Grätzer zurückgehende Satz von B.H. Neumann [13]:

Korollar 7. *Gehört eine universelle Algebra A zu einer Klasse $\mathfrak{A}$, die sich mit endlich vielen Gleichungen beschreiben läßt, so läßt sich A durch Hinzunahme zweier weiterer Operationen einengen zu einer Algebra, deren zugehörige Klasse sich mit einer einzigen Gleichung definieren läßt.*

Denn man ordne A so an, daß ein minimales Element in A existiert. Dann bildet A einen Brouwerschen Halbverband bezüglich der gewählten Anordnung.

Schließlich können wir Satz 3 verschärfen zu

Satz 3'. *Eine Klasse von rechtskomplementären Halbgruppen läßt sich immer dann mit nur einer Gleichung beschreiben, wenn sich eine Operation $\wedge$ angeben läßt, bezüglich der die Polynome p_1 bis p_8 den Wert a annehmen.*

Aufgrund von Satz 2 ist im wesentlichen nur die Existenz von Fundamentalgleichungen nachgewiesen, nichts ist hingegen gesagt über Minimalwerte. Daß es sehr viel kürzere Gleichungen in weniger Variablen als die von uns konstruierten geben dürfte, scheint gewiß. So sind in der Tat die beiden folgenden Gleichungen (BR) bzw. (BV) charakteristisch für den Booleschen Ring bzw. Booleschen Verband:

$$\text{(BR)} \qquad \big((x+y)+z\big)+\big((x+z)+h(t,u,v,w)\big)=y$$

$$\text{mit } h=\big((t\,t+t(u(v\,w)+w(u\,v)))+t\big)+\big(u(v+w)+(u\,v+w\,u)\big)$$

$$\text{(BV)} \qquad \big((x+y)+z\big)+\big((x+z)+h(s,t,u,v,w)\big)=y$$

$$\text{mit } h=\big(((u*u)\,s+(u*u)(s+s))+s\big)+\big(u(v\,w+t\,t)+(w(v\,u)+u(u\,t))\big).$$

Wir begnügen uns mit diesem Hinweis und verzichten auf einen Nachweis, da das genannte Resultat hier ohne besondere Bedeutung bleibt.

3. Rechtskongruenzen in rechtskomplementären Halbgruppen

Im folgenden sollen einige Sätze bezüglich der Rechtskongruenzen rechtskomplementärer Halbgruppen vorgestellt werden, unter denen sich als wesentlichstes Resultat der Satz erweisen wird, daß in endlichen rechtskomplementären Halbgruppen jede Rechtskongruenz auch eine Kongruenz ist. Wir zeigen zunächst:

Satz 4. *Sei S eine rechtskomplementäre Halbgruppe. Dann entsprechen die Rechtskongruenzen von S eineindeutig den Rechtsidealen H aus S, die definiert sein sollen vermöge:*

(i) $$a\in S \mathbin{\&} b\in H \Rightarrow a*b\in H,$$

(ii) $$a\in H \mathbin{\&} b\in H \Rightarrow ab\in H,$$

(iii) $$ab\in H \Rightarrow a\in H,$$

aufgrund der Zuordnung

$$\Theta \to H \Leftrightarrow (a\equiv b(\Theta) \Leftrightarrow a*b\in H \ni b*a).$$

Beweis. Ist Θ eine Rechtskongruenz und 1 das Einselement aus S, so erfüllt die Klasse $\Theta 1 = \bar{1}$ die Implikationen

(i) $$a\in S \mathbin{\&} b\in \bar{1} \Rightarrow a*b\equiv a*1\equiv 1\in \bar{1},$$

(ii) $$a\in \bar{1} \mathbin{\&} b\in \bar{1} \Rightarrow ab\equiv a1\equiv a\in \bar{1},$$

(iii) $$ab\in \bar{1} \Rightarrow a\equiv a1\equiv a(a*ab)\equiv ab\in \bar{1}$$

und

$$\begin{aligned} a\equiv b(\Theta) &\Rightarrow a*b\in \bar{1}\ni b*a \\ &\Rightarrow a\equiv a(a*b)\equiv b(b*a)\equiv b(\Theta). \end{aligned}$$

Sei nun H ein Rechtsideal. Dann erzeugt die Festsetzung

$$a\equiv b(H) \Leftrightarrow a*b\in H\ni b*a$$

eine Rechtskongruenz, da

$$R\colon\ a\equiv a(H) \quad \text{und} \quad S\colon\ a\equiv b \Leftrightarrow b\equiv a(H)$$

evident sind,

$$T\colon\ a\equiv b(H) \mathbin{\&} b\equiv c(H) \Rightarrow a\equiv c(H)$$

aus

$$\begin{aligned} a*c &\leqq (a*b)\big((a*b)*(a*c)\big) \\ &= (a*b)\big(a(a*b)*c\big) \\ &= (a*b)\big(b(b*a)*c\big) \\ &= (a*b)\big((b*a)*(b*c)\big)\in H \end{aligned}$$

resultiert und $b \equiv b'$ sowohl

$$\begin{aligned} ab * ab' &= b * (a * ab') \\ &\leqq b * b' \in H \end{aligned}$$

als auch

$$\begin{aligned} (a*b)*(a*b') &= a(a*b)*b' \\ &= b(b*a)*b' \\ &= (b*a)*(b*b') \in H \end{aligned}$$

impliziert. Damit ist wegen $a \in H \Leftrightarrow 1 * a \in H \ni a * 1$ alles bewiesen.

Satz 5. *Sei S eine endliche rechtskomplementäre Halbgruppe. Dann entsprechen die Rechtskongruenzen eineindeutig den Idempotenten u mit der Eigenschaft $au*u=1$ ($a \in S$) vermöge der Abbildung $\Theta \to u \Leftrightarrow (a \equiv b(\Theta) \Leftrightarrow au = bu)$.*

Beweis. Ist u idempotent, so ist jeder Linksteiler von u auch ein Rechtsteiler von u wegen

$$\begin{aligned} xy = u &\Rightarrow u \leqq ux \leqq uxy = u \\ &\Rightarrow ux = u. \end{aligned} \tag{39}$$

Nehmen wir nun weiter an, daß $au*u=1$ für alle a aus S erfüllt ist, so erhalten wir

$$a*u \leqq u$$

wegen

$$u*(a*u) = au*u = 1$$

und

$$au = ua' \quad \text{mit} \quad a' = u*au$$

wegen

$$\begin{aligned} au &= au(au*u) \\ &= u(u*au), \end{aligned}$$

was impliziert, daß jeder Rechtsteiler y von u auch ein Linksteiler von u ist, der sogar $yu = u$ erfüllt wegen

$$\begin{aligned} xy = u \Rightarrow xu &= x(x*u)((x*u)*u) \\ &= u((x*u)*u) \\ &= u \qquad\qquad (39) \\ \Rightarrow yu &= uy' \\ &= xuy' \\ &= xyu \\ &= uu = u. \end{aligned}$$

Sei nun $u u=u$ und $a u*u=1$ für alle a aus S. Dann bildet die Menge H_u aller $h \leqq u$ ein Rechtsideal, denn es gelten die Implikationen:

(i) $$a \in S \,\&\, b \leqq u \Rightarrow a*b \leqq a*u \leqq u$$

(ii) $$a \leqq u \,\&\, b \leqq u \Rightarrow a b u = a u = u \Rightarrow a b \leqq u$$

(iii) $$a b \leqq u \Rightarrow a \leqq u.$$

Ist umgekehrt H ein Rechtsideal, so gilt $a \in H \ni b \Rightarrow a*b \in H$, weshalb mit a und b auch $a \cup b$ zu H gehört. Somit existiert in H ein Maximum, nämlich $u = \bigcup h$ $(h \in H)$, das seinerseits die Eigenschaften $u u = u$, $a*u \in H \Rightarrow a*u \leqq u \Rightarrow a u*u=1$ und daher $H=H_u$ besitzt.

Somit kann man den Beweis abschließen mittels der Implikationen

$$\begin{aligned} a u = b u \Rightarrow a*b &\leqq a*b u \\ &= a*a u \leqq u \end{aligned}$$

und

$$\begin{aligned} a*b \leqq u \geqq b*a \Rightarrow a u &= a(a*b)\, u \\ &= b(b*a)\, u = b u. \end{aligned}$$

Hiernach ergibt sich der interessante Sachverhalt

Satz 6. *Ist S eine endliche rechtskomplementäre Halbgruppe, so ist jede Rechtskongruenz von S sogar eine Kongruenz.*

Beweis. Es bleibt nach dem Bisherigen zu zeigen, daß für jedes u mit $x u*u=1$ aus $a \equiv a'$ sowohl $a b \equiv a' b$ als auch $a*b \equiv a'*b$ folgt. Dies ergibt sich aber unter Annahme von $b u = u c$ aufgrund von

$$\begin{aligned} a \equiv a' \Rightarrow\ & a u = a' u \\ \Rightarrow\ & a u c = a' u c \\ \Rightarrow\ & a b u = a' b u \end{aligned}$$

und

$$\begin{aligned} a \equiv a' \Rightarrow\ & a u = a' u \\ \Rightarrow\ & a u*b = a' u*b \\ \Rightarrow\ & u*(a*b) = u*(a'*b) \\ \Rightarrow\ & u\big(u*(a*b)\big) = u\big(u*(a'*b)\big) \\ \Rightarrow\ & (a*b)\big((a*b)*u\big) = (a'*b)\big((a'*b)*u\big) \\ \Rightarrow\ & (a*b)\big((a*b)*u\big)\, u = (a'*b)\big((a'*b)*u\big)\, u \\ \Rightarrow\ & (a*b)\, u = (a'*b)\, u. \end{aligned}$$

Vermöge des soeben bewiesenen Sachverhaltes gelingt es uns, ein Beispiel für eine rechtskomplementäre Halbgruppe zu konstruieren, deren Kongruenzklassen nicht dem Chinesischen Restsatz (vgl. [15])

(CH) $$\Theta_1 a \cap \Theta_2 b \neq \emptyset \neq \Theta_2 b \cap \Theta_3 c \neq \emptyset \neq \Theta_3 c \cap \Theta_1 a \Rightarrow \Theta_1 a \cap \Theta_2 b \cap \Theta_3 c \neq \emptyset$$

genügen und deren Kongruenzen nicht vertauschbar sind. Denn man unterwerfe die freie Halbgruppe $F(1, u, v, w, x, y, z)$ den definierenden Relationen:

$$1a = a = a1 \text{ für alle erzeugenden Elemente.}$$

$$uu = u; \quad uv = vu; \quad ux = vx = wx = x;$$

$$vv = v; \quad vw = wv; \quad uy = vy = wy = y;$$

$$ww = w; \quad wu = uw; \quad uz = vz = wz = z;$$

$$xx = yy = zz = xy = yx = yz = zy = zx = xz$$

$$= xxx = yyy = zzz = xuvw = yuvw = zuvw := 0$$

und:

$$xvw = yvw; \quad yuv = zuv; \quad zwu = xwu.$$

Dann gilt für die Klassen X von x modulo vw, Y von y modulo uv und Z von z modulo wu: $X \cap Y = \{y, yv\}$, $Y \cap Z = \{z, zu\}$, $Z \cap X = \{x, xw\}$, aber $X \cap Y \cap Z = \emptyset$.

Und setzen wir $vw = a$ und $uv = b$, so folgt $xa = ya \& yb = zb$, doch läßt sich kein Element s finden mit der Eigenschaft $xb = sb \& sa = za$, da aus $xb = sb$ resultiert, daß s gleich x, xu, xv oder xuv sein muß, wohingegen $zvw \neq xvw$ und $zvw \neq xuvw$, in jedem Falle also $xb = sb \Rightarrow sa \neq za$ erfüllt ist.

4. Polynome

In [4] wurde gezeigt, daß die Kongruenzen einer jeden komplementären Halbgruppe vertauschbar sind, einen distributiven Verband bilden und durch ihre 1-Klassen charakterisiert werden. Insbesondere folgt aus den beiden ersten Eigenschaften, daß die Kongruenzen komplementärer Halbgruppen den Chinesischen Restsatz erfüllen: „Der Durchschnitt endlich vieler Kongruenzklassen ist genau dann leer, wenn es unter diesen endlich vielen Klassen zwei disjunkte gibt" (s. oben). Somit müssen insgesamt

ein Malzew-Polynom $p(x, y, z)$ mit $p(x, x, z) = z = p(z, x, x)$,

ein Pixley-Polynom $q(x, y, z)$ mit $q(x, x, z) = q(x, z, x) = q(z, x, x) = x$,

eine Wille-Kette $p_1(x, y), \ldots, p_n(x, y), r(x, y, x_1, \ldots, x_n)$

mit $$p_\nu(x, x) = 1$$

und $$r(x, y, 1, \ldots, 1) = x, r(x, y, p_1(x, y), \ldots, p_n(x, y)) = y$$

existieren. Nichts ist indessen gesagt über die Form solcher Polynome, weshalb wir der Vollständigkeit erwähnen:

Satz 7. *Sei S eine komplementäre Halbgruppe und* $h(x, y, z) = x:(z * y)$. *Dann sind*

$$p(x, y, z) = h(x, y, z) \cup h(z, y, x) \qquad \text{ein Malzew-Polynom,}$$

$$q(x, y, z) = h(x, x, y) \cup h(y, y, z) \cup h(z, z, x) \quad \text{ein Pixley-Polynom}$$

und

$$p_1(x, y) = y * x, \qquad p_2(x, y) = (x:(y * x)) * y, \qquad p_3(w, x, y, z) = (w:y)\,z$$

eine Wille-Kette.

Nach dem Beispiel des letzten Paragraphen kann zu beliebigen rechtskomplementären Halbgruppen weder ein Malzew- noch ein Pixley-Polynom existieren, da die Existenz des einen Polynoms äquivalent ist mit der Vertauschbarkeit der Kongruenzrelationen, die Existenz des anderen äquivalent ist mit dem Chinesischen Restsatz. Es gelten jedoch zwei Abschwächungen von weittragender Bedeutung und ziemlich tiefgreifender Konsequenz (vgl. [15] und [9]), nämlich:

Satz 8. *Ist S eine rechtskomplementäre Halbgruppe, so existiert ein Polynompaar*

$$p_1(w, x, y, z), \qquad p_2(w, x, y, z)$$

mit

$$x = p_1(x, y, y, z); \qquad p_1(x, x, y, y) = p_2(x, x, y, y); \qquad p_2(x, y, y, z) = z$$

sowie ein Polynompaar

$$q_1(x, y, z), \qquad q_2(x, y, z)$$

mit

$$x = q_1(x, x, z); \qquad q_1(x, z, z) = q_2(x, z, z); \qquad q_2(x, x, z) = z$$

und

$$q_1(x, y, x) = x = q_2(x, y, x).$$

Beweis. Man betrachte die Paare

$$p_1(w, x, y, z) = w(x * y), \qquad p_2(w, x, y, z) = z(y * x)$$

und

$$q_1(x, y, z) = x\big(x(y * z) * z\big), \qquad q_2(x, y, z) = z\big(z(z * y) * x\big).$$

Daß es sich bei den aufgezeigten Gegebenheiten um Abschwächungen der Malzew- bzw. Pixley-Bedingung handelt, folgt daraus, daß bei Existenz eines Malzew-Polynoms $p(x, y, z)$ das Paar $p(w, p(x, y, z), z), z$ den Forderungen an p_1, p_2 genügt, und bei Existenz eines Pixley-Polynoms das Paar $q(x, y, z), z$ den Forderungen an q_1, q_2.

Daß für beliebige rechtskomplementäre Halbgruppen eine Wille-Kette nicht unbedingt existieren muß, liegt daran, daß in ihren Existenznachweis die Vertauschbarkeit eingeht.

Literatur

1. Birkhoff, G.: Lattice theory. Amer. Math. Soc. Colloq. Publ. 35, 3rd. ed. Providence, R.I. 1967.
2. Bosbach, B.: Komplementäre Halbgruppen. Ein Beitrag zur instruktiven Idealtheorie. Math. Ann. **161**, 279–295 (1965).
3. – Komplementäre Halbgruppen. Axiomatik und Arithmetik. Fundamenta Math. **64**, 257–287 (1969).
4. – Komplementäre Halbgruppen. Kongruenzen und Quotienten. Fundamenta Math. **69**, 1–14 (1970).
5. – Komplementäre Halbgruppen. Eine Darstellungstheorie. Math. Ann. **179**, 1–14 (1968).
6. Day, A.: A characterization of modularity for congruence lattices of algebras. Canadian Math. Bull. **12**, 167–173 (1969).
7. Grätzer, G.: Two Mal'cev type theorems in universal algebra. J. Combinat. Theory **8**, 334–342 (1970).
8. – On the spectra classes of algebras. Proc. Amer. Math. Soc. **18**, 729–735 (1967).
9. Jónsson, B.: Algebras whose congruence lattices are distributive. Math. Scandinav. **21**, 110–121 (1967).
10. Mal'cev, A.I.: On the general theory of algebraic systems. Mat. Sbornik **35**, 3–20 (1954) [russisch].
11. McKenzie, R.: On equational theories of lattices. Math. Scandinav. **27**, 24–38 (1970).
12. Mitschke, A.: Implication algebras are 3-permutable and 3-distributive (preprint).
13. Neumann, B.H.: On a problem of G. Grätzer. Publ. Math. Debrecen **14**, 325–329 (1967).
14. Pixley, A.F.: Distributivity and permutability of congruence-relations in equational classes of algebras. Proc. Amer. Math. Soc. **14**, 105–109 (1963).
15. Wille, R.: Kongruenzklassengeometrien. Lecture notes vol. 113. Berlin-Heidelberg-New York: Springer 1970.

Prof. Dr. Bruno Bosbach
Gesamthochschule Kassel

(Eingegangen am 24. Mai 1971)

Math. Z. 124, 289–314 (1972)

Sphärische Kreisebene mit dreidimensionaler nichteinfacher Automorphismengruppe

Karl Strambach

§ 1. Einleitung

Sphärische Kreisebenen sind Inzidenzstrukturen, die man erhält, indem man auf der 2-Sphäre ein System $\mathfrak{S}$ von Jordankurven, von sogenannten Kreisen, so auszeichnet, daß mit drei verschiedenen Punkten genau ein Kreis inzidiert. Existiert darüber hinaus zu jedem Kreis K und zwei Punkten p und q, von denen p mit K inzidiert, q aber nicht, genau ein Kreis L, der durch q geht und mit K nur p gemeinsam hat, so sprechen wir von sphärischen Möbiusebenen; von K und L sagen wir, sie berühren einander im Punkt p. Die Begriffe der sphärischen Kreisebene bzw. Möbiusebene verallgemeinern in natürlicher Weise den Begriff der klassischen Möbiusgeometrie der ebenen Schnitte der Kugelfläche. Sie sind zugleich die einzigen topologischen Kreis- bzw. Möbiusebenen (das sind Kreis- bzw. Möbiusebenen, deren Punkt- sowie Kreismenge topologische Räume und in denen die geometrischen Operationen des Verbindens dreier Punkte sowie des Schneidens zweier Kreise bzw. des Berührens stetig sind), die sich auf einer zweidimensionalen kompakten Mannigfaltigkeit realisieren lassen (vgl. Strambach [1970] und Wölk).

Entfernt man aus einer sphärischen Kreisebene $\mathbb{K}$ einen Punkt w und sieht die mit w inzidenten Kreise von $\mathbb{K}$ ohne den Punkt w als Geraden an, so erhält man als Unterstruktur von $\mathbb{K}$ eine sogenannte Salzmann-Ebene $\mathbb{K}_w$ (vgl. etwa Strambach [1968]). Die Salzmann-Ebenen sind eine ausführlich untersuchte Klasse zweidimensionaler topologischer Geometrien (Salzmann [1967]); die für sie vorhandenen Ergebnisse können wir auch hier vielfach verwerten. Ist $\mathbb{K}$ sogar eine Möbiusebene, so ist $\mathbb{K}_w$ eine ebene affine Ebene (Salzmann [1962]).

In einer früheren Arbeit (Strambach [1970]) haben wir gezeigt, daß jede Kreisverwandtschaft, d.h., jede eineindeutige die Konzyklität erhaltende Abbildung der Punkte einer Kreisebene $\mathbb{K}$ auf sich, ein Homöomorphismus ist und daß die Automorphismen von $\mathbb{K}$ bezüglich einer natürlichen Topologie (die zu der der gleichmäßigen Konvergenz äquivalent ist) eine höchstens sechsdimensionale Liegruppe bilden. Unter anderen Charakterisierungen der klassischen Möbiusebene haben wir dort bewiesen, daß diese Geometrie die einzige sphärische Kreisebene ist, die eine mindestens vierdimensionale Gruppe von Automorphismen (als abgeschlossene Untergruppe der vollen Automorphismengruppe) zuläßt. In einer weiteren Arbeit (Strambach [1971]) wurden alle Kreisebenen beschrieben, die eine dreidimensionale (abstrakt) einfache Gruppe von Automorphismen gestatten.

In dieser Untersuchung wollen wir die Klassifikation der sphärischen Kreisebenen mit dreidimensionaler Automorphismengruppe vervollständigen, indem wir in unserem Hauptsatz alle Kreisebenen bestimmen, die eine nicht (abstrakt) einfache dreidimensionale Gruppe von Automorphismen zulassen: Die Möbiusebenen unter ihnen, die nicht zum klassischen Modell isomorph sind, haben entweder den Heringschen Typ III.1 oder III.2; bei den echten Kreisebenen hingegen ist das Berühraxiom nur für einen einzigen Punkt verletzt, und die Zusammenhangskomponente der vollen Automorphismengruppe besteht aus Affinitäten der Form

$$(x, y) \mapsto (t x + u, t^c y + v) \quad \text{mit } t>0, u, v \in \mathbb{R}$$

und einer reellen Zahl $c>1$ bzw.

$$(x, y) \mapsto (e^t x + t e^t y + u, e^t y + v) \quad \text{mit } t, u, v \in \mathbb{R}.$$

Für Möbiusebenen ergibt sich aus dieser Arbeit zusammen mit den Ergebnissen aus Strambach [1971] der folgende Sachverhalt: *Erlaubt eine sphärische Möbiusebene eine dreidimensionale Gruppe von Kreisverwandtschaften, so hat sie den Heringschen Typ* III.1, III.2, IV.2 *bzw.* VII.2.

Bezeichnungen. Γ^1 ist die Zusammenhangskomponente der topologischen Gruppe Γ, mit Γ_p sei die Standuntergruppe der Transformationsgruppe Γ auf dem Punkt p einer Mannigfaltigkeit bezeichnet. Die von uns betrachteten Untergruppen topologischer Gruppen sind abgeschlossen. Eine Zusammenstellung der einfachsten, von uns meist ohne ausdrückliche Erwähnung benutzten Tatsachen über Liesche Transformationsgruppen auf zweidimensionalen Mannigfaltigkeiten findet man in Strambach [1967], § 1 oder [1968], § 1.

Eine Inversion ist ein Automorphismus einer Kreisebene, der einen Kreis punktweise festläßt, aber außerhalb von diesem keine Fixpunkte besitzt (Strambach [1970], S. 279). $\mathbb{R}$ bezeichnet die reellen, $\mathbb{C}$ die komplexen Zahlen.

§ 2. Vorbereitungen

Hilfssatz 1. *In der reellen affinen Ebene* $\mathbf{D}$ *seien* K_1 *und* K_2 *zwei verschiedene konvexe Kurven, die in einem Punkt* $t \in K_1 \cap K_2$ *eine gemeinsame Stützgerade* G *besitzen. Schneidet* K_1 *jedes Translat von* K_2*, das von* K_1 *und* K_2 *verschieden ist, jeweils in höchstens zwei Punkten, so können sich* K_1 *und* K_2 *(in* t*) nicht durchsetzen; liegen* K_1 *und* K_2 *in derselben von* G *bestimmten Halbebene, so umfaßt vielmehr einer der beiden von* K_i *berandeten kompakten Bereiche den anderen.*

Beweis. Es ist $K_1 \neq K_2$. Die Gerade G, die wir als x-Achse eines Koordinatensystems wählen, stütze K_1 und K_2 im Punkt $t=(0,0) \in K_1 \cap K_2$. Sowohl die Punkte von K_1 als auch die von K_2 mögen alle nichtnegative Ordinaten haben. Wir können zugleich $K_i \cap G = t$ voraussetzen, denn sonst gäbe es geeignete, der Identität hinreichend benachbarte Translationen φ mit

$$|K_1 \cap K_2^\varphi| \geqq 3.$$

Angenommen, $|K_1 \cap K_2| \geqq 2$. Dann existiert ein Punkt p von K_2, der außerhalb des von K_1 begrenzten kompakten Bereichs $\mathfrak{B}$ liegt, und ein Punkt q von K_2, der zu den inneren Punkten von $\mathfrak{B}$ zählt. Es sei nun eine Abbildung

$$(*) \qquad \rho = \{(x, y) \mapsto (x+d, y)\}$$

so gewählt, daß t^ρ negative bzw. positive Abszisse hat, je nachdem ob K_2 links bzw. rechts von t zunächst nicht außerhalb von $\mathfrak{B}$ verläuft und die so der Identität benachbart ist, daß $p^\rho \notin \mathfrak{B}$ und $q^\rho \in \mathfrak{B}$ gilt. Der Punkt t^ρ liegt außerhalb von $\mathfrak{B}$. J sei der Bogen von K_2, der von p und q begrenzt wird und t nicht enthält. Es gilt dann $|J^\rho \cap K_1| \geqq 1$. Ist L der von t^ρ und p^ρ begrenzte Bogen von K_2^ρ mit $J^\rho \cap L = p^\rho$, so gilt $|K_1 \cap L| \geqq 2$, und daher ist $|K_2^\rho \cap K_1| > 3$, was gegen die Voraussetzungen verstößt.

Wäre $|K_1 \cap K_2| = 1$ und gälte die Behauptung nicht, so wären die K_i in t nicht differenzierbar, und ein Widerspruch folgt ähnlich wie vorher.

Hilfssatz 2. *Ist Γ eine zusammenhängende, dreidimensionale, nicht abstrakt einfache Gruppe von Kreisverwandtschaften einer sphärischen Kreisebene $\mathbb{K}$, so läßt sie einen Punkt ∞ on $\mathbb{K}$ fest; die Unterebene $\mathbb{K}_\infty$ ist die reelle affine Ebene $\mathbf{D}$, und Γ induziert auf $\mathbb{K}_\infty$ eine punkttransitive Kollineationsgruppe, die wie die Erweiterung der Translationsgruppe durch eine der folgenden Gruppen wirkt:*

(a) *Eine Einparametergruppe von Drehstreckungen bzw. Drehungen.*

(b) *Die Gruppe der Affinitäten*

$$\{(x, y) \mapsto (t\,x, t^c\,y);\ 0 < t \in \mathbb{R} \text{ und } c > 0 \text{ eine feste reelle Zahl}\}.$$

(c) *Die Gruppe der Affinitäten*

$$\{(x, y) \mapsto (e^t\,x + t\,e^t\,y, e^t\,y);\ t \in \mathbb{R}\}.$$

Beweis. Daß Γ einen Punkt ∞ von $\mathbb{K}$ festläßt, auf $\mathbb{K}_\infty$ punkttransitiv operiert und daß $\mathbb{K}_\infty$ eine affine Ebene ist, folgt nach Strambach [1970], Lemma 4.10, Satz 3.12, sowie Strambach [1970a].

Angenommen, die affine Ebene $\mathbb{K}_\infty$ sei nicht desarguessch: dann läßt sie sich nach der eben zitierten Klassifikation (vgl. auch Salzmann [1965]) über der Punktmenge der reellen affinen Ebene $\mathbf{D}$ so darstellen, daß die Geraden von $\mathbb{K}_\infty$ genau die Kurven $x = \text{const}$ und $y - n = c_m(x)\,|x-m|^d$ mit $m, n \in \mathbb{R}$ sowie

$$c_m(x) = \begin{cases} 1 & \text{für } x \geqq m \\ c & \text{für } x < m \end{cases}$$

sind; die festen reellen Zahlen c und d genügen der Ungleichung $0 < c \leqq 1 < d$ und sind Invarianten der Ebene $\mathbb{K}_\infty$. Die Gruppe Γ besteht dann aus den Abbildungen $(x, y) \mapsto (p\,x + m, p^d\,y + n)$ mit $p > 0$, $m, n \in \mathbb{R}$. Die nicht durch ∞ gehenden Kreise von $\mathbb{K}$ bilden ein System $\mathfrak{S}$ von Jordankurven in $\mathbb{K}_\infty$. Da die Gruppe Γ alle (gewöhnlichen) Translationen enthält, ist jede Kurve aus $\mathfrak{S}$

nicht nur bezüglich der Geraden von $\mathbb{K}_\infty$, sondern auch bezüglich der gewöhnlichen Geraden konvex (Rosenthal). Die durch $y=dx+1-d$ beschriebene euklidische Gerade G ist euklidische Tangente an die durch $y=c_0(x)\,|x|^d$ beschriebene Kurve F im Punkt $q=(1,1)$. Es gibt einen Kreis L von $\mathbb{K}$, für den (∗) $L\cap F=q$ gilt und der durch einen Punkt $r=(0,t)$ mit $1-d<t<0$ geht (Strambach [1970], 2.4). L gehört zu $\mathfrak{S}$. Wegen (∗) ist dann p entweder ein Wendepunkt von L, oder L wird von G in p gestützt und die zu q hinreichend benachbarten Punkte von L sind durch G von r getrennt; beides widerspricht aber der Konvexität von L (Haupt-Künneth, S. 70), und $\mathbb{K}_\infty$ ist die reelle affine Ebene $\mathbf{D}$.

Nach Strambach [1970a] ist dann die Gruppe Γ eine der in der Behauptung genannten Gruppen.

§ 3. Die Klassifikation

Hauptsatz. *Ist $\mathbb{K}$ eine sphärische Kreisebene, die eine dreidimensionale zusammenhängende nicht (abstrakt) einfache Gruppe Γ von Automorphismen gestattet, so läßt Γ einen Punkt ∞ fest und operiert auf der zur reellen affinen Ebene isomorphen Unterebene $\mathbb{K}_\infty$ als Erweiterung der Translationsgruppe* T *durch eine der folgenden Einparametergruppen:*

$$\Delta_d=\{(x,y)\mapsto(x\,d^t\cos t-y\,d^t\sin t,\ x\,d^t\sin t+y\,d^t\cos t);\ t\in\mathbb{R}\ \textit{und}\ d\geqq 1\ \textit{eine feste reelle Zahl}\},$$

$$\Xi_c=\{(x,y)\mapsto(t\,x,\ t^c\,y);\ 0<t\in\mathbb{R}\ \textit{und}\ c\geqq 1\ \textit{eine feste reelle Zahl}\},$$

$$\Pi=\{(x,y)\mapsto(e^t\,x+t\,e^t\,y,\ e^t\,y);\ t\in\mathbb{R}\ \textit{und}\ e\ \textit{die Basis der natürlichen Logarithmen}\}.$$

Die mit ∞ inzidenten Kreise von $\mathbb{K}$ sind in jedem Fall genau die gewöhnlichen durch ∞ abgeschlossenen Geraden von $\mathbb{K}_\infty$. Die nicht mit ∞ inzidenten Kreise von $\mathbb{K}$ sind Bilder eines einzigen Kreises K unter Γ; seine Gestalt bestimmt also die Struktur der Kreisebene $\mathbb{K}_K$. Zwei über derselben Punktmenge realisierte Kreisebenen $\mathbb{K}_{K_1}$ und $\mathbb{K}_{K_2}$ sind genau dann isomorph, wenn auf ihnen die gleiche Gruppe Γ operieren kann und die Jordankurve K_1 sich auf die Kurve K_2 unter solchen Affinitäten abbilden läßt, die Γ normalisieren.

(A) *Gilt $\Gamma=\mathsf{T}\cdot\Delta_1$, so ist $\mathbb{K}$ die klassische Möbiusgeometrie der ebenen Schnitte der Kugelfläche.*

Gilt $\Gamma=\mathsf{T}\cdot\Delta_d$ mit $d>1$, so ist $\mathbb{K}$ eine Möbiusebene $\mathbb{M}_K$. Ein ihre Struktur bestimmender Kreis K ist eine streng konvexe Jordankurve; diese wird durch eine auf $\mathbb{C}\setminus 0$ definierte, reellwertige Stützfunktion $h\geqq 0$ mit $h(r\,e^{-i\frac{\pi}{2}})=0$ beschrieben, so daß für $-\frac{\pi}{2}<t\leqq\frac{3\pi}{2}$ und alle z die Beziehung

$$(1)\qquad h(z)\leqq d^{(\frac{3}{2}\pi-t)}\,h(z\,e^{i(t-\frac{3}{2}\pi)})-\Re z\,\Re(w_t\,z_t)-\Im z\,\Im(w_t\,z_t)$$

erfüllt ist, wobei

$$z_t = \frac{\partial h}{\partial \mathfrak{Re}\, z}(e^{it}) + i\frac{\partial h}{\partial \mathfrak{Im}\, z}(e^{it}) \quad \textit{und} \quad w_t = d^{(\frac{3}{2}\pi - t)}\, e^{i(\frac{3}{2}\pi - t)}$$

gilt, die Gleichheit nur für $t = \frac{3}{2}\pi$ *eintritt und in* z *stets eine der beiden Funktionen*

$$\frac{d\,\dfrac{\partial h(z)}{\partial \mathfrak{Im}\, z}}{d\,\dfrac{\partial h(z)}{\partial \mathfrak{Re}\, z}} \quad \textit{bzw.} \quad \frac{d\,\dfrac{\partial h(z)}{\partial \mathfrak{Re}\, z}}{d\,\dfrac{\partial h(z)}{\partial \mathfrak{Im}\, z}}$$

existiert.

Es gibt zu jedem $d > 1$ *überabzählbar viele nichtisomorphe Möbiusebenen. Zwei zur klassischen Möbiusgeometrie nichtisomorphe Möbiusebenen, die zu verschiedenen Werten von* d *gehören, sind stets nichtisomorph; sie sind aber alle vom Heringschen Typ* III.1.

Die volle Automorphismengruppe Θ *einer nicht zur klassischen Möbiusgeometrie isomorphen Möbiusebene* $\mathbb{M}_K$ *zerfällt über* Γ *mit einer orientierungstreuen, endlichen zyklischen Gruppe* Ψ, *mit einer Inversion* φ *bzw. mit einem semidirekten Produkt von* Ψ *mit* φ. *Es gilt* $|\Psi| = l > 1$ *genau dann, wenn* h *für alle* z *und* $1 \leqq l \leqq n-1$ *zusätzlich der Bedingung*

$$h(z) = h\left(e^{-i\frac{2\pi l}{n}} z\right) - \mathfrak{Re}\, z\; \mathfrak{Re}\left(e^{i\frac{2\pi l}{n}} z_l\right) - \mathfrak{Im}\, z\; \mathfrak{Im}\left(e^{i\frac{2\pi l}{n}} z_l\right) \tag{2}$$

genügt, wobei

$$z_l = \frac{\partial h}{\partial \mathfrak{Re}\, z}\left(e^{i\pi\left(\frac{2l}{n} - \frac{1}{2}\right)}\right) + i\frac{\partial h}{\partial \mathfrak{Im}\, z}\left(e^{i\pi\left(\frac{2l}{n} - \frac{1}{2}\right)}\right)$$

ist. Die Gruppe Ψ *enthält genau dann Inversionen, die auf* $\mathbb{K}_\infty$ *als Geradenspiegelungen wirken, wenn* K *bezüglich der* y*-Achse oder bezüglich einer Geraden* $y = \text{const}$ *symmetrisch ist.*

(B_1) *Gilt* $\Gamma = \mathrm{T} \cdot \Xi_1$, *so ist* $\mathbb{K}$ *eine Möbiusebene* $\mathbb{M}_K$. *Ein ihre Struktur bestimmender Kreis* K *ist eine beliebige, streng konvexe, differenzierbare Jordankurve; daher gibt es überabzählbar viele nichtisomorphe Möbiusebenen, die* $\mathrm{T} \cdot \Xi_1$ *als eine Gruppe von Automorphismen zulassen.*

Ist eine Möbiusebene $\mathbb{M}_K$ *nicht die klassische Möbiusgeometrie, so hat sie den Heringschen Typ* III.1 *bzw.* III.2; *vom Typ* III.2 *ist* $\mathbb{M}_K$ *genau dann, wenn* K *zentralsymmetrisch ist. Ist die volle Automorphismengruppe* Θ *einer Möbiusebene* $\mathbb{M}_K$, *die den Typ* III *hat, von* Γ *verschieden, so zerfällt sie über* Γ *mit einer endlichen Gruppe* Ψ, *die zyklisch, die Kleinsche Vierergruppe oder eine Diedergruppe ist und nur affine Drehungen oder Spiegelungen enthalten kann.*

(B_2) *Gilt* $\Gamma = \mathrm{T} \cdot \Xi_c$ *mit* $c > 1$, *so wird die Struktur der Kreisebene* $\mathbb{K}_K$ *durch eine Jordankurve* K *bestimmt, die durch ein Quadrupel von stetig differenzierbaren Funktionen* g_j *beschrieben wird, so daß die folgenden, für* g_j *notwendigen und hinreichenden Bedingungen erfüllt sind;*

(3) *g_1 ist für $a_1 \leqq x \leqq 0$, die Funktion g_4 ist für $0 \leqq x \leqq a_2$ definiert; dabei ist $a_1 < 0 < a_2$. Weiter sei $g_1(0) = g_4(0) = g_1'(0) = g_4'(0) = 0$ und $g_1(a_1) = b_1 < 1$ sowie $g_4(a_2) = b_2 < 1$.*

(4) *g_2 ist für $a_1 \leqq x \leqq \tilde{e}$, die Funktion g_3 ist für $\tilde{e} \leqq x \leqq a_2$ definiert; dabei ist $a_1 < \tilde{e} < a_2$. Weiter sei $g_2(\tilde{e}) = g_3(\tilde{e}) = 1$ und $g_2'(\tilde{e}) = g_3'(\tilde{e}) = 0$ sowie $g_2(a_1) = b_1$ und $g_3(a_2) = b_2$.*

(5) *Die Ableitungen g_2' und g_3' sind streng monoton fallende, die Ableitungen g_1' und g_4' sind streng monoton wachsende Funktionen.*

(6) *Für $j = 1, 4$ sei $k_j(x) = x[g_j(x)]^{-c^{-1}}$, für $j = 2, 3$ sei $k_j(x) = x[1 - g_j(x + \tilde{e})]^{-c^{-1}}$. Die Funktionen k_j sind streng monoton fallend mit $|k_j(0)| = \infty$.*

(7) *Es seien $t_1 > t_2$ zwei reelle Zahlen, für die entweder $t_j \geqq 1$ oder $t_j \leqq 1$ gilt, und $x_l^{(t_j)}$ zwei solche Stellen aus dem Definitionsbereich von g_l, daß*

$$|t_j^{c-1} g_l'(t_j^{-1} x_l^{(t_j)})| = \begin{cases} |g_l'(a_i)|, & \textit{falls } |g_l'(a_i)| < \infty \\ 1, & \textit{falls } |g_l'(a_i)| = \infty \end{cases}$$

ist; dabei gilt $i = 1$ für $l = 1, 2$ und $i = 2$ für $l = 3, 4$.

Bezeichnet $f_l^{t_j}(x) = t_j^c g_l(t_j^{-1} x + t_j^{-1} x_l^{(t_j)}) - t_j^c g_l(t_j^{-1} x_l^{(t_j)})$, so hat jede Funktion $F_l^{t_1, t_2} = f_l^{t_2} - f_l^{t_1}$, die auf dem Durchschnitt der Definitionsbereiche von $f_l^{t_1}$ und $f_l^{t_2}$ definiert ist, genau eine Nullstelle, und zwar für $x = 0$. Für $l = 1, 4$ ist $F_l^{t_1, t_2}$ nirgendwo negativ, für $l = 2, 3$ ist $F_l^{t_1, t_2}$ nirgendwo positiv.

Die Kreisebene $\mathbb{K}$ ist genau dann sogar eine Möbiusebene, wenn $g_l'(a_1) = g_s'(a_2) = \infty$ für $l = 1, 2$ und $s = 3, 4$ gilt. Eine Möbiusebene $\mathbb{K}$, die $\Gamma_c = \mathrm{T} \cdot \Xi_c$ als eine Gruppe von Automorphismen zuläßt, hat stets den Heringschen Typ III.1. *Ist $\mathbb{K}$ keine Möbiusebene, so ist ∞ der einzige Punkt, von dem aus es mehrere Berührkreise an Kreise von $\mathbb{K}$ gibt. Zu jedem $c > 1$ existieren überabzählbar viele nichtisomorphe Kreis- sowie Möbiusebenen mit Γ_c als Automorphismengruppe. Zwei Kreisebenen $\mathbb{K}_i$, auf denen jeweils als Automorphismengruppen Γ_{c_i} operieren, sind für $c_1 \neq c_2$ niemals isomorph.*

Die volle Automorphismengruppe Θ einer Kreisebene $\mathbb{K}$ zerfällt über Γ_c mit einem Komplement Φ, das in der Gruppe der Affinitäten $\{(x, y) \mapsto (\pm x, \pm y)\}$ enthalten ist. Φ enthält genau dann die Inversion $(x, y) \mapsto (-x, y)$, wenn K bezüglich der y-Achse symmetrisch ist; die Inversion $(x, y) \mapsto (x, -y)$ gehört zu Φ genau dann, wenn $g_2 = 1 - g_1$ und $g_3 = 1 - g_4$ gilt. Die Spiegelung $(x, y) \mapsto (-x, -y)$ liegt in Φ genau dann, wenn K zentralsymmetrisch ist.

(C) *Gilt $\Gamma = \mathrm{T} \cdot \Pi$, so wird die Struktur der Kreisebene $\mathbb{K}_K$ durch eine Jordankurve K bestimmt, die durch ein Paar von stetig differenzierbaren Funktionen h_j beschrieben wird (sie drücken die Abhängigkeit der Abszissen von den Ordinaten aus), so daß die folgenden, für h_j notwendigen und hinreichenden Bedingungen erfüllt sind:*

(8) *Die Funktionen h_j sind für $0 \leqq y \leqq 1$ definiert; es gilt $h_j(0) = h_j(1) = 0$ sowie $h_2 \geqq 0$ bzw. $h_1 \leqq 0$.*

(9) *h_1' ist streng monoton wachsend, h_2' dagegen streng monoton fallend.*

(10) *Die Funktionen*

$$k_1(y) = y\, e^{\frac{-h_1(y)}{y}} \quad \text{und} \quad k_2(y) = y \cdot e^{\frac{-h_2(y+1)}{y}},$$

die für $0 < y \leqq 1$ bzw. $-1 \leqq y < 0$ definiert sind, sind streng monoton fallend, und es gilt $\lim\limits_{|y| \mapsto 0} |k_j(y)| = \infty$.

(11) *Es seien $t_2 > t_1$ nichtnegative bzw. nichtpositive reelle Zahlen, zu denen es jeweils Stellen $y_{t_i}^{(j)}$ gibt, die durch die Beziehungen $t_i + h_j'(e^{-t_i} y_{t_i}^{(j)}) = 0$ bestimmt sind.*

Bezeichnet

$$\tilde{f}_j^{t_i}(y) = t_i(y + y_{t_i}^{(j)}) + e^{t_i} h_j[e^{-t_i}(y + y_{t_i}^{(j)})] - t_i\, y_{t_i}^{(j)} - e^{t_i} h_j(e^{-t_i} y_{t_i}^{(j)}),$$

so hat jede Funktion $\tilde{F}_j^{t_1, t_2} = \tilde{f}_j^{t_2} - \tilde{f}_j^{t_1}$, die auf dem Durchschnitt der Definitionsbereiche von $\tilde{f}_j^{t_1}$ und $\tilde{f}_j^{t_2}$ definiert ist, genau eine Nullstelle, und zwar für $y = 0$. Die Funktion $\tilde{F}_1^{t_1, t_2}$ ist nirgendwo positiv, $\tilde{F}_2^{t_1, t_2}$ ist dagegen nirgendwo negativ.

Die Kreisebene $\mathbb{K}_K$ ist genau dann sogar eine Möbiusebene, wenn $h_1'(1) = h_2'(0) = \infty$ gilt; ist sie eine Möbiusebene so hat sie stets den Heringschen Typ III.1. *Ist $\mathbb{K}_K$ keine Möbiusebene, so ist ∞ der einzige Punkt, von dem aus es mehrere Berührkreise an Kreise von $\mathbb{K}_K$ gibt. Es existieren überabzählbar viele nichtisomorphe Kreis- sowie Möbiusebenen $\mathbb{K}_K$, auf denen Γ operiert und deren strukturbestimmender Kreis K sogar unendlich oft differenzierbar ist. Die volle Automorphismengruppe Θ von $\mathbb{K}_K$ ist eine zerfallende Erweiterung von Γ durch eine Gruppe der Ordnung $\leqq 2$. Ist $\Theta \neq \Gamma$, so enthält Θ Involutionen, die auf der zum Punkt ∞ gehörenden Unterebene wie Punktspiegelungen wirken, und dies trifft genau dann zu, wenn $h_2(y) = -h_1(1-y)$ gilt.*

Beweis. Der erste Teil des Satzes folgt nach Hilfssatz 2. Es seien $\mathbb{K}_{K_1}$ und $\mathbb{K}_{K_2}$ zwei über derselben Punktmenge realisierte, isomorphen Kreisebenen, auf denen die gleiche zusammenhängende Gruppe Γ operiere. Der Isomorphismus ρ zwischen $\mathbb{K}_{K_1}$ und $\mathbb{K}_{K_2}$ ist dann auf $\mathbb{K}_\infty$ eine Affinität; sie kann so gewählt werden, daß $K_2 = K_1^\rho$ gilt. $\mathbb{K}_{K_2}$ erlaubt als Automorphismengruppe das Erzeugnis Φ von Γ und $\rho^{-1}\Gamma\rho$. Ist $\mathbb{K}_{K_2}$ nicht die klassische Möbiusgeometrie, so ist Φ dreidimensional (Strambach [1970], 4.7), und es gilt daher $\rho^{-1}\Gamma\rho = \Gamma$. Ist $\mathbb{K}_{K_2}$ isomorph zur klassischen Möbiusebene, so gilt entweder $\Gamma = \mathrm{T} \cdot \Delta_d$ oder $\Gamma = \mathrm{T} \cdot \Xi_1$, und in beiden Fällen ist Γ ein Normalteiler der Gruppe $\{z \mapsto az + b;\ 0 \neq a, b \in \mathbb{C}\}$.

(A) Zuerst betrachten wir den Fall, daß $\Gamma = \mathrm{T} \cdot \Delta_d$ gilt.

(a) Ist $d = 1$, so ist Γ die Gruppe der orientierungstreuen euklidischen Bewegungen, und $\mathbb{K}$ ist nach Buckel und van Heemert die klassische Möbiusgeometrie. (Dasselbe folgt nach Strambach [1970], Lemma 4.1, denn die Punktbahnen der maximalen kompakten Untergruppen von Γ sind dann euklidische Kreise.)

Es sei also von nun an stets $d > 1$.

(b) Der Kreis $\tilde{K}$ von $\mathbb{K}$ gehe nicht durch ∞ und sei in einem Punkt $\tilde{t} \in \tilde{K}$ nicht differenzierbar. Dann existiert auch ein nicht mit ∞ inzidenter Kreis $K = \tilde{K}^{\gamma}$ mit einem geeigneten Element $\gamma \in \Gamma$, der im Punkt $t = (0, 0)$ nicht differenzierbar ist, so daß überdies die linke Tangente T_1 an K in t die nichtpositive Abszissenhalbachse $X^- = \{(x, 0);\ 0 \geqq x \in \mathbb{R}\}$ enthält (vgl. Haupt-Künneth, S. 63) und K nur Punkte mit nichtnegativen Ordinaten besitzt; die rechte Tangente T_2 schließt dann mit der Halbachse $X^+ = \{(x, 0);\ 0 \leqq x \in \mathbb{R}\}$ einen Winkel $0 < \beta < \pi$. In der Standgruppe Γ_t gibt es Elemente γ, für die der in der oberen Halbebene liegende Halbstrahl von T_1^{γ} mit X^+ den (positiv gemessenen) Winkel $0 < \kappa < \beta$ einschließt und $\beta + \kappa < \pi$ gilt. Dann liegen aber hinreichend zu t benachbarte Punkte von K^{γ} außerhalb bzw. innerhalb des von K begrenzten kompakten Bereichs, je nachdem ob sie sich beim Durchlaufen von K^{γ} im positiven Sinn links bzw. rechts von t befinden. Dies widerspricht Hilfssatz 1, denn K und K^{γ} liegen in derselben Halbebene bezüglich einer Geraden $y = k x$ mit $\kappa < \operatorname{arc\,tg} k < \beta$. Daher ist jeder Kreis von $\mathbb{K}$, der nicht mit ∞ inzidiert, eine differenzierbare Kurve. Da diese Kreise streng konvex sind, ist jeder von ihnen sogar eine stetig differenzierbare Kurve (Bonnesen-Fenchel, S. 15).

(c) Wir betrachten die Standgruppe Γ_K eines nicht durch ∞ gehenden Kreises K. Da die Kurve K endlichen (euklidischen) Durchmesser hat, ist die Standgruppe Γ_K in einem Komplement des Translationsnormalteilers N enthalten. Die abelsche Gruppe Γ_K läßt dann im Innern von K einen Punkt c fest; wir können sogar $c = (0, 0)$ annehmen. Die Punkte von $\mathbb{K}_{\infty}$ können durch komplexe Zahlen repräsentiert werden; die Zahl z liege auf K. Wäre die Gruppe $\Gamma_K \neq 1$, so enthielte sie Elemente $\varphi^n = (z \mapsto z\, d^{nt}\, e^{int})$ mit einem festen t und $n \in \mathbb{N}$. Der Abstand zwischen z^{φ^n} und c würde für große n beliebig groß, was der Endlichkeit des Durchmessers von K widerspricht und $\Gamma_K = 1$ beweist.

(d) Jeder Punkt $z = u_1 + i u_2 = r e^{i\varphi}$ von $\mathbb{K}_{\infty}$ kann auch als ein Vektor aufgefaßt werden. Die Stützfunktion der Kurve $K \not\ni \infty$, die mit $c = (0, 0)$ inzidiert, die reelle Achse als Tangente in c besitzt und nur Punkte mit nichtnegativen Ordinaten hat, sei die über $\mathbb{C} \setminus 0$ definierte reellwertige Funktion $h(z)$; sie ordnet dem Vektor $z = r e^{i\varphi}$ den mit r multiplizierten Abstand des Punktes 0 von derjenigen orientierten Tangente an K, die zum Vektor $e^{i\left(\varphi + \frac{\pi}{2}\right)}$ parallel ist (Bonnesen-Fenchel, S. 24). Da K nur reguläre Stützgeraden besitzt (s. (b)), hat h stetige partielle Ableitungen und für die Punkte (x, y) von K gelten die Beziehungen $x = \dfrac{\partial h}{\partial u_1}$, $y = \dfrac{\partial h}{\partial u_2}$ (Bonnesen-Fenchel, S. 26). Für den Berührungspunkt z_t der zu $e^{i\left(t + \frac{\pi}{2}\right)}$ parallelen, orientierten Tangenten an K gilt $z_t = \dfrac{\partial h}{\partial u_1}(e^{it}) + i \dfrac{\partial h}{\partial u_2}(e^{it})$. Es sei $w_t = d^{(\frac{3}{2}\pi - t)} \cdot e^{i(\frac{3}{2}\pi - t)}$ mit $-\dfrac{\pi}{2} < t \leqq \dfrac{3}{2}\pi$. Wenden wir auf K den Automorphismus

$$\gamma_t = (z \mapsto w_t\, z - w_t\, z_t) \in \Gamma, \tag{12}$$

so geht K^{γ_t} durch den Punkt 0, hat die reelle Achse als Tangente und besitzt nur Punkte mit nichtnegativen Ordinaten. Für die Stützfunktion h_t von K^{γ_t} gilt dann

$$h_t(z) = d^{(\frac{3}{2}\pi - t)}\, h(z\, e^{i(t-\frac{3}{2}\pi)}) - \mathfrak{Re}\, z\, \mathfrak{Re}(w_t\, z_t) - \mathfrak{Im}\, z\, \mathfrak{Im}(w_t\, z_t)$$

(vgl. Bonnesen-Fenchel, S. 24 und 30). Der Durchmesser von K^{γ_t} ist für $t \neq \frac{3}{2}\pi$ größer als der von K. Nach Hilfssatz 1 umfaßt dann K^{γ_t} den Kreis K; der Durchschnitt $K^{\gamma_t} \cap K$ besteht nur aus dem Punkt 0, weil es sonst in der Unterebene $\mathbb{K}_0$ einander berührende Geraden gäbe (Salzmann [1967], 2.8). Für die Stützfunktionen h und h_t gilt dann

$$h \leqq h_t, \tag{13}$$

wobei für $t \neq \frac{3}{2}\pi$ die Gleichheit nur für $z = r\, e^{-i\frac{\pi}{2}}$ angenommen wird.

(e) Sei umgekehrt $h(z)$ eine auf $\mathbb{C} \setminus 0$ definierte, reellwertige nichtnegative Stützfunktion (vgl. Bonnesen-Fenchel, S. 24) mit $h(r\, e^{-i\frac{\pi}{2}}) = 0$, so daß für $-\dfrac{\pi}{2} < t \leqq \dfrac{3}{2}\pi$ und alle z die Beziehung (1) gilt, dort die Gleichheit nur für $t = \frac{3}{2}\pi$ eintritt und für jedes z stets eine der beiden Funktionen

$$\frac{d\,\dfrac{\partial h}{\partial \mathfrak{Im}\, z}}{d\,\dfrac{\partial h}{\partial \mathfrak{Re}\, z}} \quad \text{bzw.} \quad \frac{d\,\dfrac{\partial h}{\partial \mathfrak{Re}\, z}}{d\,\dfrac{\partial h}{\partial \mathfrak{Im}\, z}}$$

existiert. Dann beschreibt h eine konvexe, stetig differenzierbare Kurve K, die durch 0 geht und dort als Tangente die reelle Achse X hat (Bonnesen-Fenchel, S. 26 und Ostrowski, II, S. 296).

Es sei $\mathfrak{X} = \{K^{\gamma_t}\}$ mit γ_t aus (12) und $\mathfrak{Y} = \{\mathfrak{X}^{\Sigma}\}$ die Menge der Bilder von Kurven aus $\mathfrak{X}$ unter der Gruppe $\Sigma = \{z \mapsto d^{2\pi k} z;\, k \in \mathbb{N}\}$. Dann überdeckt $\mathfrak{Y} \cup X$ die in 0 punktierte reelle affine Ebene $\mathbf{D}$ schlicht, denn je zwei Kurven aus $\mathfrak{X}$ haben nur den Punkt 0 gemeinsam (Bonnesen-Fenchel, S. 24), und jeder mit 0 inzidente Kreis, der X in 0 als Tangente hat, gehört zu $\mathfrak{Y} \cup X$.

Ist $\delta \in \Gamma$, so überdeckt $\mathfrak{Y}^{\delta} \cup X^{\delta}$ die in 0^{δ} punktierte Ebene $\mathbf{D}$ schlicht. Da Γ auf $\mathbf{D}$ fahnentransitiv operiert, genügt $\mathbb{K} = \mathbf{D} \cup \infty$ dem Berühraxiom. Ist $b \neq \infty$ ein von 0 verschiedener Punkt von $\mathbb{K}$, so existiert in jedem Büschel $\mathfrak{Y}^{\lambda} \cup X^{\lambda}$ mit dem Trägerpunkt 0 genau eine Kurve, die durch b geht. Aus Stetigkeitsgründen überdecken nun die durch 0 und b gehenden Kurven aus K^{Γ} zusammen mit der durch ∞ abgeschlossenen Geraden $0 \cup b$, die in 0 und b doppelt punktierte 2-Sphäre $\mathbf{D} + \infty$ schlicht, und daher bilden die kompaktifizierten euklidischen Geraden zusammen mit den Kurven K^{Γ} das System der Kreise der sphärischen Möbiusebene $\mathbb{M}_K$.

(f) Ist $\mathbb{M}_K$ nicht die klassische Möbiusgeometrie, so ist Γ die Zusammenhangskomponente der vollen Automorphismengruppe Θ (Strambach [1970], 4.8). Es gilt $\Theta=\Gamma\cdot\Theta_K$ und $\Theta_K\cap\Gamma=1$ (vgl. (c)). Da die Gruppe Θ_K keine nichtinvolutorischen axialen Kollineationen besitzt (Strambach [1970], 3.12), ist sie sogar als Permutationsgruppe isomorph zu einer diskreten Untergruppe der von $\{z\mapsto k\,e^{it}z;\ k\neq 0,\ t\in\mathbb{R}\}$ und $z\mapsto\bar{z}$ erzeugten Gruppe. Da K kompakt ist, enthält Θ_K lauter Elemente endlicher Ordnung. Ist nun $\Theta_K\neq 1$, so ist Θ_K als Permutationsgruppe isomorph zu einer der folgenden Gruppen:

(14) $\Lambda_n=\{z\mapsto e^{i\frac{2\pi}{n}l}z;\ l=1,\dots,n-1$ und $n\neq 1$ eine feste natürliche Zahl$\}$,

(15) $\Phi=\{(x,y)\mapsto(x,\pm y)\}$ oder $\Psi=\{(x,y)\mapsto(\pm x,y)\}$,

(16) das semidirekte Produkt von Λ_n mit Φ bzw. Ψ.

Ist Θ_K isomorph zu einer Gruppe Λ_n, so erfüllt die Stützfunktion h der Kurve K außer (1) wegen $K=K^{\gamma_l}$ mit

$$\gamma_l=\left(z\mapsto e^{\frac{2\pi i l}{n}}z-e^{\frac{2\pi i l}{n}}\left[\frac{\partial h}{\partial\,\mathfrak{Re}\,z}\left(e^{i\pi\left(\frac{2l}{n}-\frac{1}{2}\right)}\right)+\frac{\partial h}{\partial\,\mathfrak{Im}\,z}\left(e^{i\pi\left(\frac{2l}{n}-\frac{1}{2}\right)}\right)\right]\right)$$

zusätzlich für alle z und $1\leqq l\leqq n-1$ die Beziehungen (2). Enthält Θ_K die Gruppe Φ bzw. ist Θ_K isomorph zu Ψ, so ist K bezüglich der y-Achse bzw. bezüglich einer Geraden $y=\text{const}$ symmetrisch.

Da die Gruppen Δ_d für verschiedene Werte von d in der vollen Affinitätengruppe nicht konjugiert sind, können zu verschiedenen d konstruierte Möbiusebenen nicht isomorph sein. Daß jede Möbiusebene $\mathbb{M}_K$, die Γ als Automorphismengruppe zuläßt, entweder die klassische Möbiusgeometrie oder vom Heringschen Typ III.1 ist, folgt unmittelbar (vgl. Hering).

(g) Jetzt zeigen wir, daß es zu jeder Zahl $d>1$ überabzählbar viele Möbiusebenen $\mathbb{M}_K$ gibt, die nicht isomorph sind und deren nicht durch ∞ gehende Kreise Eilinien mit differenzierbarer Krümmung sind.

Sei K eine durch $p=(0,0)$ gehende Eilinie (vgl. Strubecker, S. 48), deren Tangente in p die reelle Achse X sei. Man kann K durch eine Stützfunktion $f(r)$ beschreiben, die dem Neigungswinkel v der inneren Normalen von K gegen die nichtnegative reelle Halbachse X^+ den (mit Vorzeichen versehenen) Abstand des Punktes p von derjenigen orientierten Tangente an K zugeordnet, die mit X^+ den Winkel $v+\frac{\pi}{2}$ einschließt. Nach Strubecker, S. 52, ist $f(v)=f(v+2\pi)$ eine nichtnegative, viermal stetig differenzierbare Funktion, und $f(v)=0$ gilt genau dann, wenn $v=\frac{\pi}{2}(1+4k)$ mit $k\in\mathbb{N}$ ist. Wir nehmen zusätzlich an, daß in jedem Punkt von K die Krümmung $\kappa\neq 0$ ist. Wegen der Konvexität von K gilt dann sogar stets $\kappa>0$. Nach Strubecker, S. 53, gilt für die Krümmungsradien $g=1/\kappa$ von K die Beziehung

$$g=f+f''>0. \tag{17}$$

Ist $\gamma_t=(z\mapsto w_t z-w_t a_t)$ ein Element aus Γ mit

$$a_t=-f(-t-\pi)\,e^{i(-t-\pi)}-f'(-t-\pi)\,e^{-i(-t-\pi)}\in K$$

(vgl. Strubecker, S. 52) und $w_t = d^{(t+\frac{3}{2}\pi)} e^{i(t+\frac{3}{2}\pi)}$, so hängt für alle v und

$$-\frac{3\pi}{2} \leqq t < \frac{\pi}{2}$$

die Funktion g_t der Krümmungsradien von K^{γ_t} mit g durch die Beziehung

$$g_t(v) = d^{(t+\frac{3}{2}\pi)} \cdot g(v-t-\tfrac{3}{2}\pi) \tag{18}$$

zusammen, weil g unter Translationen invariant bleibt.

Die Kurven K und K^{γ_t} liegen auf derselben Seite ihrer gemeinsamen Tangente im Punkt 0.

Nach Blaschke, S. 115, ist die Beziehung

$$g \leqq g_t \tag{19}$$

dafür hinreichend, daß K^{γ_t} die Kurve K umfaßt; gilt für $t \neq -\frac{3}{2}\pi$ in (19) stets die Ungleichheit, so hat K mit K^{γ_t} nur den Punkt 0 gemeinsam. Erfüllt also die Funktion g der Krümmungsradien von K die Beziehung (19), so existiert eine Möbiusebene $\mathbb{M}_K$.

(h) Mit (19) ist wegen (18) die für alle v geltende Ungleichung

$$d^{(t+\frac{3}{2}\pi)} \geqq 1 + \frac{g(v) - g(v-t-\frac{3}{2}\pi)}{g(v-t-\frac{3}{2}\pi)} \tag{20}$$

gleichwertig. Beschränken wir uns von nun an auf die Betrachtung solcher Kurven K, daß für die Funktion g die Beziehung

$$g \geqq C > 0 \tag{21}$$

mit einer festen reellen Zahl C gilt, so wird (20) von jeder Funktion g erfüllt, die für alle v der Bedingung

$$C(d^{(t+\frac{3}{2}\pi)} - 1) \geqq g(v) - g(v-t-\tfrac{3}{2}\pi) = (t+\tfrac{3}{2}\pi) \cdot g'(w) \tag{22}$$

mit $v - t - \frac{3}{2}\pi \leqq w \leqq v$ genügt.

Die Ungleichung (22) erfüllen insbesondere solche Funktionen, für die

$$C(d^{(t+\frac{3}{2}\pi)} - 1) - (t+\tfrac{3}{2}\pi) \cdot g' \geqq 0 \tag{23}$$

richtig ist und die Gleichheit dort nur für $t = -\frac{3}{2}\pi$ eintritt. Eine hinreichende Bedingung für die Gültigkeit von (23) ist

$$\frac{\partial}{\partial t}\left[C(d^{(t+\frac{3}{2}\pi)} - 1) - (t+\tfrac{3}{2}\pi)\, g'\right] = C d^{(t+\frac{3}{2}\pi)} \ln d - g' \geqq 0, \tag{24}$$

wobei die Gleichheit nur für $t = -\frac{3}{2}\pi$ eintreten kann; dann ist nämlich

$$C(d^{(t+\frac{3}{2}\pi)} - 1) - (t+\tfrac{3}{2}\pi) \sup g'(v) \quad \text{für } -\tfrac{3}{2}\pi \leqq t < \frac{\pi}{2} \tag{25}$$

eine streng monoton wachsende Funktion. Daher wird (19) insbesondere von allen Funktionen g erfüllt, für die

$$\sup g' \leqq C \ln d \tag{26}$$

gilt.

(i) Es sei $g \geqq C$ eine Funktion, die (26) erfüllt und die Periode 2π hat. Die Stützfunktion f einer Kurve K ist dann die periodische Lösungsfunktion der Differentialgleichung (17) mit den Anfangsbedingungen $f(-\frac{3}{2}\pi) = f(\pi/2) = f'(\pi/2) = 0$ (Blaschke S. 116 oder Kamke S. 257). Die Bedingung für die Geschlossenheit von f ist

$$\int_{-\frac{3}{2}\pi}^{\pi/2} g(v) \cos v\, dv = \int_{-\frac{3}{2}\pi}^{\pi/2} g(v) \sin v\, dv = 0 \quad \text{(Blaschke S. 116).} \tag{27}$$

Die stetig differenzierbaren reellen Funktionen mit der Periode 2π bilden einen Vektorraum $\mathfrak{V}$ der Dimension des Kontinuums über $\mathbb{R}$. Es sei $\mathfrak{L}$ der von $\cos x$ und $\sin x$ erzeugte zweidimensionale Unterraum von $\mathfrak{V}$ und $\mathfrak{L}^\perp$ sein orthogonales Komplement (vgl. Natanson, S. 193, Def. 1). Eine Funktion aus $\mathfrak{V}$ gehört genau dann zu $\mathfrak{L}^\perp$, wenn ihre Fourier-Reihe die Form

$$a_0 + \sum_{\nu=2}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x) \tag{28}$$

hat. $\mathfrak{L}_1^\perp$ bezeichne den Unterraum von $\mathfrak{L}^\perp$, der zu dem von den Konstanten aufgespannten Unterraum orthogonal ist. Ist g' eine beliebige Funktion aus $\mathfrak{L}_1^\perp$, so kann sie so normiert werden, daß (26) erfüllt ist. Für die Fourier-Reihe von g' gilt $a_0 = 0$. Es gibt nun Stammfunktionen g von g', für die $g \geqq C$ gilt, weil g durch gliedweises Aufintegrieren der Fourier-Reihe von g' erhalten wird. $g \in \mathfrak{L}^\perp$ und erfüllt (27). In $\mathfrak{L}^\perp$ liegen also überabzählbare viele Funktionen g, die verschieden sind, aber auf einem einzigen offenen Intervall I übereinstimmen. Daher existieren auch überabzählbar viele verschiedene, dreimal stetig differenzierbare Stützfunktionen f als Lösungen von (17), die ebenfalls nur auf I übereinstimmen; je zwei solche Stützfunktionen sind affin nicht äquivalent und führen daher zu nichtisomorphen Möbiusebenen.

(B) Von nun an sei $\Gamma = \mathrm{T} \cdot \Xi_c$ vorausgesetzt.

(a′) Wäre $c < 0$, so bestände die Bahn von $(1, 1)$ unter der Standgruppe $\Gamma_{(0,0)}$ aus genau derjenigen Punkten der reellen Ebene $\mathbf{D}$, die auf der durch $y = t^c$ mit $t > 0$ dargestellten Kurve H liegen. In der Kreisebene $\mathbb{K}$ gibt es einen mit $(2, 1)$ inzidenten Kreis K, der mit der durch ∞ abgeschlossenen Geraden $x = 0$ genau den Punkt $(0, 0)$ gemeinsam hat (Strambach [1970], 2.4). Da $(0, 0)$ und $(2, 1)$ auf verschiedenen Seiten der Kurve H liegen, existieren in $H \cap K$ zwei verschiedene Punkte u und v, so daß auf H außerhalb des abgeschlossenen Intervalls $J = [u, v]$ keine Punkte von K liegen. In $\Gamma_{(0,0)}$ gibt es ein Element γ mit $u^\gamma = v$; da γ auf H wie eine Translation auf der reellen Zahlengeraden wirkt, gilt $v^\gamma \notin J$ und $v^\gamma \notin K$. Die verschiedenen Kreise K und K^γ müßten sich dann in ihren Schnittpunkten jeweils durchsetzen (Salzmann [1967], 2.8), was Hilfssatz 1 widerspricht, weil die Jordankurven K und K^γ in $(0,0)$ als gemeinsame Stützgerade $x=0$ haben. Dieser Widerspruch beweist $c>0$.

(b′) Da die Gruppe Ξ_c unter der Affinität $(x,y)\mapsto(y,x)$ zur Gruppe $\Xi_{c^{-1}}$ konjugiert ist, können wir uns auf den Fall $c\geqq 1$ beschränken.

(c′) Sei K ein nicht durch ∞ gehender Kreis einer Kreisebene $\mathbb{K}$, die Γ als eine Gruppe von Automorphismen zuläßt. Die Standgruppe Γ_K enthält keine Translationen $\neq 1$ und operiert auf der Jordankurve K effektiv; sie ist dann eine Untergruppe der Standgruppe eines Punktes. Da K endlichen Durchmesser hat, muß sogar $\Gamma_K=1$ gelten.

(d′) Nun sei $c=1$ vorausgesetzt. Wir nehmen weiter an, der Kreis K sei in einem Punkt t nicht differenzierbar. Da K eine konvexe Jordankurve ist, existiert in t eine eindeutig bestimmte „vordere" Tangente T_v und eine eindeutig bestimmte „hintere" Tangente T_h (vgl. Haupt-Künneth, S. 63); es ist $T_v\neq T_h$. Die von T_h bestimmte offene Halbebene, die Punkte von K enthält, sei mit $\mathfrak{H}$ bezeichnet. Ist nun p ein Punkt aus $\mathfrak{H}\cap T_v$, so existiert ein mit p inzidenter Kreis L, der die durch ∞ abgeschlossene Gerade T_h in t berührt, d.h., sie in t als Stützgerade hat (Strambach [1970], 2.4). Wegen der Konvexität der Kurve L sind sowohl die vordere als auch die hintere Tangente in t an L von T_v verschieden. Würde L die Kurve K umfassen, so gäbe es Bilder L^τ mit geeigneten Elementen $\tau\in\Gamma_t$, die zwar beliebig kleinen Durchmesser hätten, aber nicht im Innern von K lägen, weil sich p^τ stets außerhalb von K befindet. Daher muß L in t die Kurve K durchsetzen, und dies widerspricht Hilfssatz 1. Jeder nicht mit ∞ inzidente Kreis von $\mathbb{K}$ ist also eine streng konvexe, stetig differenzierbare Jordankurve der reellen Ebene $\mathbf{D}$ (Bonnesen-Fenchel, S. 15).

(e′) Gibt man umgekehrt eine streng konvexe, differenzierbare Jordankurve K der reellen Ebene $\mathbf{D}$ vor, so bildet die Menge der Kurven K^γ mit $\gamma\in\Gamma=\mathrm{T}\cdot\Xi_1$ zusammen mit den durch ∞ abgeschlossenen euklidischen Geraden eine Möbiusebene $\mathbb{M}_K$ (Ewald). Zwei Möbiusebenen $\mathbb{M}_{K_1}$ und $\mathbb{M}_{K_2}$ sind genau dann isomorph, wenn die Kurven K_j affin äquivalent sind. Da es überabzählbar viele nicht affin äquivalente Jordankurven gibt, die konvex und differenzierbar sind, kann Γ auf überabzählbar vielen nichtisomorphen Möbiusebenen operieren.

(f′) Θ sei die volle Automorphismengruppe einer Möbiusebene $\mathbb{M}_K$, die nicht zur klassischen Möbiusgeometrie der ebenen Schnitte der Kugelfläche isomorph sei. Dann gilt $\infty^\Theta=\infty$, denn sonst wäre Θ punkttransitiv, was 4.3 aus Strambach [1970] widerspricht. Da Γ auf den mit ∞ nichtinzidenten Kreisen scharf transitiv operiert, ist $\Theta=\Gamma\cdot\Theta_K$ und $\Gamma\cap\Theta_K=1$. Da $\dim\Theta=3$ gilt (Strambach [1970], Satz 4.7) und Θ keine nichtinvolutorischen axialen Kollineationen enthält (Strambach [1970], 3.12), ist Θ_K (vgl. (c)) eine diskrete Untergruppe einer kompakten Gruppe von Affinitäten und daher isomorph zu einer der in (B_1) angeführten Gruppen. Die Aussagen über den Heringschen Typ von $\mathbb{M}_K$ folgen sofort nach Hering und Ewald.

(g′) Von nun an werden wir stets $\Gamma = \mathrm{T} \cdot \Xi_c$ mit $c > 1$ voraussetzen. K sei wiederum ein mit ∞ nicht inzidenter Kreis einer Kreisebene $\mathbb{K}$, auf der Γ als Gruppe von Automorphismen operiert. Die streng konvexe Jordankurve K gehe durch den Punkt $(0,0)$ und habe dort $y=0$ als Stützgerade. Weiter darf man annehmen, daß K mit $(\tilde{e}, 1)$ inzidiert und dort $y=1$ als Stützgerade besitzt, denn die Streckungen normalisieren Γ.

(h′) Die Bahnen $B_{(x,y)}$ der Punkte (x, y) mit $x \neq 0$ unter der Standgruppe $\Gamma_{(0,0)}$ sind Kurven der Form $w = y\,(v/x)^c$, wobei v alle positiven bzw. negativen reellen Zahlen durchläuft, je nachdem ob $x > 0$ bzw. $x < 0$ gilt.

Angenommen, eine Bahn $B_{(x,y)}$ treffe K in verschiedenen Punkten; l_1 bzw. l_2 seien die Punkte, die unter den gemeinsamen Punkten von K und $B_{(x,y)}$ die größte bzw. kleinste Ordinate haben. In der Standgruppe $\Gamma_{(0,0)}$ existiert nun ein der Identität so benachbartes Element γ, daß l_1^γ außerhalb, aber l_2^γ im Innern von K liegt. Die Kreise K und K^γ durchsetzen sich also in $(0,0)$ und haben dort $y=0$ als gemeinsame Stützgerade, was Hilfssatz 1 widerspricht. Daher hat jede Bahn $B_{(x,y)}$ mit K höchstens einen Punkt gemeinsam.

Angenommen, K enthalte nur Punkte mit lauter nichtnegativen bzw. nichtpositiven Abszissen. Dann liegt ein Punkt $r = (0, a)$ mit $0 < a \in \mathbb{R}$ außerhalb des von K begrenzten kompakten Bereichs $\mathfrak{B}$. Mit r inzidiert ein Kreis L, der die Gerade $y=0$ im Punkt $(0,0)$ berührt (Strambach [1970], 2.4). In der Menge L^{Ξ_c} befinden sich Kreise mit beliebig kleinem Durchmesser. Da r^δ mit $\delta \in \Xi_c$ stets außerhalb von $\mathfrak{B}$ liegt, gibt es geeignete Kreise L^δ, die weder K umfassen noch in $\mathfrak{B}$ enthalten sind. Dies widerspricht Hilfssatz 1, und K besitzt sowohl Punkte mit positiven als auch solche mit negativen Abszissen. Da K die Kurve K^ρ für kein $1 \neq \rho \in \Gamma$ in $(0,0)$ durchsetzen kann und $|K \cap K^\rho| \leqq 2$ gilt, ist K in $(0,0)$ differenzierbar und hat dort waagrechte Tangente.

Analog folgt, daß auch die Jordankurve K^ε mit $\varepsilon = \{(x, y) \mapsto (x - \tilde{e}, y - 1)\}$ jede Bahn $B_{(x,y)}$ in höchstens einem Punkt trifft und im Punkt $(0,0)$ differenzierbar ist. Dann ist aber K in $(\tilde{e}, 1)$ differenzierbar und hat dort waagrechte Tangente.

(i′) Angenommen, die Kurve K sei in einem Punkt t nicht differenzierbar, in dem sie keine senkrechte Stützgerade besitzt. In t besitzt dann aber K eine eindeutig bestimmte vordere Tangente T_1 und eine eindeutig bestimmte hintere Tangente T_2, so daß $T_1 \neq T_2$ gilt (Haupt-Künneth, S. 63). Wegen $(0,0) \neq t \neq (\tilde{e}, 1)$ sind die Steigungsfaktoren k_i von T_i von Null verschieden. Es gibt eine Translation τ mit $t^\tau = (0,0)$. Die Stützgeraden T_i^τ von K^τ haben wiederum k_i als Steigungsfaktoren. Es sei etwa $|k_1| > |k_2|$. Wegen $\operatorname{sign} k_1 = \operatorname{sign} k_2$ existiert in der Standgruppe $\Gamma_{(0,0)}$ ein Element ω mit $T_1^{\tau\omega} = T_2^\tau$. Dann gilt $|K^\tau \cap T_2^{\tau\omega}| = 2$ und $|K^{\tau\omega} \cap T_1| = 2$; daher durchsetzen sich die Kurven K^τ und $K^{\tau\omega}$ in $(0,0)$. Da $T_2^\tau = T_1^{\tau\omega}$ sowohl K^τ als auch $K^{\tau\omega}$ stützt, erhält man einen Widerspruch zum Hilfssatz 1. Also ist die Kurve K in jedem Punkt stetig differenzierbar (Bonnesen-Fenchel, S. 15), in dem sie keine senkrechte Stützgerade besitzt.

(j′) K läßt sich in vier „konvexe" Bögen C_j zerlegen, die durch stetig differenzierbare Funktionen $g_j(x)$ mit den Eigenschaften (3) und (4) beschrieben werden. Die Bedingung (5) folgt aus der Konvexität von K (Ostrowski I, S. 253). Umgekehrt folgt aus (5), indem man die Anfangswerte von g_j' berücksichtigt, daß g_1 und g_3 streng monton fallende, g_2 und g_4 dagegen streng monoton wachsende Funktionen sind.

(k′) Da g_2 und g_3 von unten konkav (Bonnesen-Fenchel, § 4), aber die Bahnen $B_{(x,y)}$, die die beiden durch g_2 und g_3 dargestellten Bögen treffen, von unten konvex sind, gilt $|B_{(x,y)} \cap C_k| \leqq 1$ für $k=2,3$. Für $j=1,4$ und

$$\varepsilon=\{(x,y)\mapsto(x-\tilde{e},y-1)\}$$

ist analog $|B_{(x,y)} \cap C_j^\varepsilon| \leqq 1$, weil C_j^ε von unten konvex ist, nur Punkte mit negativen Ordinaten enthält und die C_j^ε treffenden Bahnen $B_{(x,y)}$ konkav sind.

(l′) Für $h=1,4$ ist die Bedingung

$$|B_{(x,y)} \cap C_h| \leqq 1 \tag{29}$$

genau dann erfüllt, wenn die in (6) angeführten Funktionen k_h streng monoton sind, denn die Bahn $B_{(x_0,g_h(x_0))}$ des Punktes $(x_0, g_h(x_0))$ wird durch

$$w_{x_0}^h(v)=g_h(x_0)(v/x_0)^c$$

beschrieben, und die Funktionswerte von $k_h(x)$ sind Abszissen der Schnittpunkte von $w_x^h(v)$ mit der Geraden $y=1$. Analog ist für $h=2,3$ die Bedingung

$$|B_{(x,y)} \cap C_h^\varepsilon| \leqq 1 \quad \text{mit} \quad \varepsilon=\{(x,y)\mapsto(x-\tilde{e},y-1)\} \tag{30}$$

genau dann erfüllt, wenn die Funktionen k_h aus (6) für $h=2,3$ streng monoton fallend sind, denn die Funktionswerte von $k_h(x)$ sind Abszissen der Schnittpunkte der Bahn $B_{(x,1-g_h(x+\tilde{e}))}$ mit $y=-1$.

(m′) Angenommen, der Kreis K der Kreisebene $\mathbb{K}$ habe mit der Bahn Q eines Punktes (x_0, y_0) mit $y_0>0$ keinen Punkt gemeinsam; dann ist Q sogar punktfremd zu allen Kreisen aus $K^{\Gamma(0,0)}$. Der Punkt z habe positive Ordinate und werde durch Q von den von $(0,0)$ verschiedenen Punkten jedes Kreises aus $K^{\Gamma(0,0)}$ getrennt. In $\mathbb{K}$ existiert nun ein mit z inzidenter Kreis L, der die (durch ∞ abgeschlossene) Gerade $y=0$ berührt (Strambach [1970], 2.4). Nach Hilfssatz 1 umfaßt dann L den von K begrenzten kompakten Bereich. Daher schneidet L die Halbgerade $\{(0,y);\ y>0\}$ in einem Punkt t. Da t und z in der oberen Halbebene durch Q getrennt werden, gilt $|L\cap Q|=1$, was der Tatsache $L\notin K^{\Gamma(0,0)}$ widerspricht. Daher trifft K die Bahn $B_{(x,y)}$ jedes Punktes (x,y) mit $y>0$ in genau einem Punkt (vgl. (h′)). Analog zeigt man, daß der Kreis K^ε mit $\varepsilon=\{(x,y)\mapsto(x-\tilde{e},y-1)\}$ jede Bahn $B_{(x,y)}$ mit $y<0$ in genau einem Punkt trifft. Die beiden vorstehenden Sachverhalte sind aber genau dann erfüllt, wenn für die streng monotonen Funktionen k_j aus (6) die Beziehung $|k_j(0)|=\infty$ gilt.

(n′) Der Kreis $K^{\gamma_t} = K_t$ mit

(31) $$\gamma_t = [(x, y) \mapsto (t\,x, t^c\,y);\ 0 < t \in \mathbb{R}] \in \Gamma_{(0,0)}$$

wird durch das Quadrupel der Funktionen $g_l^t(x) = t^c g_l(t^{-1} x)$ beschrieben, wobei $t\,a_1 \leqq x \leqq 0$ bzw. $t\,a_1 \leqq x \leqq t\,\tilde{e}$ bzw. $t\,\tilde{e} \leqq x \leqq t\,a_2$ bzw. $0 \leqq x \leqq t\,a_2$ gilt, je nachdem ob $l=1$ oder $l=2$ oder $l=3$ oder $l=4$ ist.

d_l sei $|g_l'(a_i)|$ bzw. 1, je nachdem ob $|g_l'(a_i)| < \infty$ bzw. $|g_l'(a_i)| = \infty$ gilt. Zu jedem $t \geqq 1$ existiert dann stets genau ein $x_l^{(t)}$, so daß $|t^{c-1} g_l'(t^{-1} x_l^{(t)})| = d_l$ erfüllt ist. Sind nun $t_1 > t_2 \geqq 1$ zwei reelle Zahlen, so haben die Kurven K_{t_1} und K_{t_2} in $(x_l^{t_1}, g_l^{t_1}(x_l^{t_1}))$ und $(x_l^{t_2}, g_l^{t_2}(x_l^{t_2}))$ parallele Stützgeraden. Dann haben die Kurven $K_{t_i}^{\rho_i}$ mit $\rho_i = [(x, y) \mapsto (x - x_l^{(t_i)}), y - g_l^{(t_i)}(x_l^{(t_i)})]$ in $(0,0)$ eine gemeinsame Stützgerade und außer $(0,0)$ keinen weiteren Punkt gemeinsam. Da der mit $(0,0)$ inzidente Bogen von $K_{t_i}^{\rho_i}$ durch die Funktion

$$f_l^{t_i}(x) = g_l^{t_i}(x + x_l^{(t_i)}) - g_l^{t_i}(x_l^{(t_i)})$$

mit $t_i\,a_1 - x_l^{(t_i)} \leqq x \leqq -x_l^{(t_i)}$ bzw. $t_i\,a_1 - x_l^{(t_i)} \leqq x \leqq t_i\,\tilde{e} - x_l^{(t_i)}$ bzw. $t\,\tilde{e} - x_l^{(t_i)} \leqq x \leqq t_i\,a_2 - x_l^{(t_i)}$ bzw. $-x_l^{(t_i)} \leqq x \leqq t_i\,a_2 - x_l^{(t_i)}$ beschrieben wird, je nachdem ob $l=1$ oder $l=2$ oder $l=3$ oder $l=4$ gilt, hat jede Funktion $F_l^{t_1, t_2} = f_l^{t_2} - f_l^{t_1}$ eine einzige Nullstelle. Für $l = 1, 4$ ist $F_l^{t_1, t_2} \geqq 0$, für $l = 2, 3$ gilt $F_l^{t_1, t_2} \leqq 0$.

(o′) Nun gehen wir umgekehrt von einer vorgegebenen Kurve K aus, die durch die Funktionen g_l, $l = 1, \ldots, 4$, mit den Eigenschaften (3) bis (7) beschrieben wird. Wir zeigen, daß die durch ∞ abgeschlossenen gewöhnlichen Geraden zusammen mit dem Kurvensystem K^Γ die Kreise einer Kreisebene $\mathbb{K}_K$ bilden, auf der $\Xi_c \cdot \mathrm{T}$ als Gruppe von Automorphismen operiert. Die Beziehungen aus (6) sichern (vgl. (k′) und (l′)), daß die Menge der Kurven $K^{\Xi_c} \cup K^{\varepsilon \Xi_c}$, wobei $\varepsilon = [(x, y) \mapsto (x - \tilde{e}, y - 1)]$ ist, zusammen mit der Geraden $y = 0$ die in $(0,0)$ punktierte affine Ebene $\mathbf{D}$ schlicht überdeckt.

(p′) Die Kurve K^{τ_i} mit $\tau_i = [(x, y) \mapsto (x - a_i, y - b_i)]$ wird durch die Funktionen $g_l^{\tau_i}(x) = g_l(x + a_i) - b_i$ beschrieben; dabei wird x jeweils den Intervallen $\mathfrak{I}_l$ entnommen, die von $a_1 - a_i$ und $-a_i$ bzw. $a_1 - a_i$ und $\tilde{e} - a_i$ bzw. $\tilde{e} - a_i$ und $a_2 - a_i$ bzw. $-a_i$ und $a_2 - a_i$ begrenzt werden, je nachdem ob $l=1$ oder $l=2$ oder $l=3$ oder $l=4$ gilt.

Von den durch die Funktionen $g_l^{\tau_i}$ beschriebenen Bögen $C_l^{\tau_i}$ von K^{τ_i} sind $C_1^{\tau_i}$ und $C_4^{\tau_i}$ von unten konvex, $C_2^{\tau_i}$ und $C_3^{\tau_i}$ von unten konkav. Die Bahnen $B_{(x,y)}$ von Punkten (x, y) mit $y \neq 0$ und $x \neq 0$ unter der Gruppe Ξ_c sind (von unten) streng konvex bzw. konkav, je nachdem ob $y > 0$ bzw. $y < 0$ gilt. Für $k = 1, 2$ gilt daher

(32) $$|C_k^{\tau_1} \cap B_{(x,y)}| \leqq 1.$$

Jede Bahn $B_{(x,y)}$ für $x > 0$, $y \neq 0$ wird durch eine Funktion $w(v) = y(v/x)^c$ mit $0 < v \in \mathbb{R}$ beschrieben; die Ableitung von $B_{(x,y)}$ ist dann für $y > 0$ stets positiv, für $y < 0$ stets negativ. Wegen der Beziehungen (3) bis (5) gilt $g_3^{\tau_1\prime} \leqq 0$ und $g_4^{\tau_1\prime} \geqq 0$ Daher ist auch für $k = 3, 4$ die Beziehung (32) richtig. Da $B_{(x,y)} \cup (0,0)$ in $(0,0)$ die Gerade $y = 0$ als Tangente besitzt, gilt für $x > 0$ sogar stets

$|B_{(x,y)} \cap C_l^{\tau_1}| = 1$. Analog beweist man $|B_{(x,y)} \cap C_l^{\tau_2}| = 1$, sofern $x < 0$ ist. Daher überdeckt die Menge der Kurven $K^{\tau_1 \Xi_c} \cup K^{\tau_2 \Xi_c}$ zusammen mit der Geraden $x = 0$ die in $(0, 0)$ punktierte affine Ebene $\mathbf{D}$ schlicht.

(q′) K umfaßt die Kurve K^γ, falls $\gamma \in \{(x, y) \mapsto (t\,x, t^c\,y);\ 0 < t < 1\}$ gilt. Es sei $(\tilde{e}, 1)^\gamma = (e_1, e_2)$ und L die Gerade $y - e_2 = \dfrac{1 - e_2}{\tilde{e} - e_1}(x - e_1)$. Gehört nun λ zu der Menge $\mathfrak{Y}$ der Translationen, die (e_1, e_2) auf einen inneren Punkt des von (e_1, e_2) und $(\tilde{e}, 1)$ begrenzten Intervalls $\mathfrak{S}$ von L abbilden, so gilt $K \cap K^{\gamma\lambda} = \emptyset$, und $K^{\gamma\lambda}$ liegt ganz im Innern des von K begrenzten kompakten Bereichs $\mathfrak{B}$, denn sonst erhielte man einen Widerspruch zu (o′).

(r′) Wir wollen zeigen, daß von den zwei Kurven $K = \bigcup_{j=1}^{4} C_j$ und $K^\varphi = \bigcup_{j=1}^{4} C_j^\varphi$ mit $\varphi \in \Gamma$, die im Punkt t eine gemeinsame Stützgerade G haben, die eine die andere umfaßt. Der Punkt t gehöre zu C_k; dann liegt er auch auf C_k^φ.

Angenommen, es wäre $|C_k^\varphi \cap K| > 1$. Sei nun ρ dasjenige Element aus Γ, das t auf $(0, 0)$ so abbildet, daß der Absolutbetrag der Steigung der Geraden G^ρ gerade $|t_j\, g_k'(t_j^{-1} x_k^{(j)})|$ beträgt. Wegen $\Gamma = \Xi_c \cdot \mathrm{T}$ sichert uns aber die Bedingung (7) zu, daß $|C_k^{\varphi\rho} \cap K^\rho| = 1$ gilt, denn die Translationen lassen die Steigungen von Geraden invariant. Aus diesem Widerspruch folgt nun $C_k^\varphi \cap K = t$. Wegen der Transitivität von Γ auf K^Γ darf man sogar voraussetzen, daß C_k^φ in dem von K begrenzten kompakten Bereich enthalten ist.

Ist nun

$$\varphi\,\rho = [(x, y) \mapsto (w_1\,x + u_1,\ w_1^c\,y + v_1)]$$

und

$$\rho = [(x, y) \mapsto (w_2\,x + u_2,\ x_2^c\,y + v_2)],$$

so gilt wegen (7) die Beziehung $w_1 < w_2$ und $s = w_1\,w_2^{-1} < 1$. Da dann

$$\varphi = [(x, y) \mapsto (s\,x + u_3,\ s^c\,y + u_3)]$$

gilt, umfaßt K die Kurve K^λ mit $\lambda = [(x, y) \mapsto (s\,x, s^c\,y)]$. Sind α und β_i, $i = 1, 2$ diejenigen Translationen, die $(\tilde{e}, 1)^\lambda$ auf $(\tilde{e}, 1)$ bzw. $(a_i, b_i)^\lambda$ auf (a_i, b_i) abbilden, so umfaßt K nach (p′) und (o′) die Kurven $K^{\lambda\alpha}$ und $K^{\lambda\beta_i}$.

Angenommen, es sei $|K \cap K^\varphi| \geqq 2$. Der Punkt z sei ein zu t auf K direkt benachbarter Punkt aus $K \cap K^\varphi$; er liege auf dem Teilbogen C_h.

Ist C_h zu C_k benachbart, d.h. gilt $|C_h \cap C_k| = 1$, so zerlegt das Paar $\{t, z\}$ die Kurve K in zwei Bögen, von denen einer genau einen der vier Punkte $(0, 0)$, $(\tilde{e}, 1)$, (a_i, b_i) enthält; dieser Punkt heiße d. Mit L sei diejenige mit d inzidente, waagrechte bzw. senkrechte Gerade bezeichnet, die K^φ in zwei verschiedenen Punkten trifft. Ist nun σ die Translation, für die $d^{\varphi\sigma} = d$ gilt, so wird wegen der scharfen Transitivität von Γ auf der Menge K^Γ die Kurve K^φ auf eine der Kurven K^λ, $K^{\lambda\alpha}$, $K^{\lambda\beta_i}$ abgebildet. Dies ist jedoch ein Widerspruch, weil entweder z^σ oder t^σ außerhalb des von K begrenzten kompakten Bereichs liegt, je nachdem ob z bzw. t von d^φ durch L getrennt wird.

Ist C_h nicht zu C_k benachbart, d.h., gilt $|C_h \cap C_k| = \emptyset$, so liegt $(1, \tilde{e})^{\varphi} = \hat{d}$ im Innern von K. Es sei $\tau = [(x, y) \mapsto (x+a, y)]$ diejenige Translation, die $\hat{d}$ auf einen inneren Punkt des Intervalls $\mathfrak{S}$ auf der Geraden L aus (q') abbildet. Da Γ auf der Menge K^{Γ} scharf transitiv wirkt und $(0,0)^{\varphi}$ im Innern von K liegt, gehört einerseits $K^{\varphi\tau}$ zur Menge $K^{\mathfrak{P}}$, andererseits befindet sich aber einer von den Punkten t^{τ} und z^{τ} außerhalb des von K begrenzten kompakten Bereichs, was (q') widerspricht.

(s') Ist G eine beliebige Gerade durch $(0,0)$, so existieren an jede Kurve L aus K^{Ξ_c} zwei verschiedene Stützgeraden S_L^+ und S_L^-, die zu G parallel sind; s_L^+ bzw. s_L^- seien die Berührpunkte von S_L^+ bzw. S_L^- mit L. Es gibt zwei Translationen τ_L^+ und τ_L^- mit $(s_L^+)^{\tau_L^+} = (0,0) = (s_L^-)^{\tau_L^-}$. Die Kurven $L^{\tau_L^+}$ und $L^{\tau_L^-}$ liegen in verschiedenen der beiden von G bestimmten Halbebenen. Die Mengen $\mathfrak{X}^+ = \{L^{\tau_L^+};\ L \in K^{\Xi_c}\}$ und $\mathfrak{X}^- = \{L^{\tau_L^-};\ L \in K^{\Xi_c}\}$ überdecken zusammen mit G die in $(0,0)$ punktierte Ebene $\mathbf{D}$ schlicht. Die Menge $\mathfrak{X}^+ \cup \mathfrak{X}^- \cup G$ heiße das Büschel $\mathfrak{C}_G$ zur Geraden G. Zu jeder mit $(0,0)$ inzidenten Geraden H existiert genau ein Büschel $\mathfrak{C}_H$, das $\mathbf{D} \setminus (0,0)$ schlicht überdeckt. Ist z ein Punkt aus $\mathbf{D} \setminus (0,0)$, so sei C_H die Kurve aus $\mathfrak{C}_H$, die mit z inzidiert. Die Menge $\{C_H;\ H \ni (0,0)\}$ überdeckt die in $(0,0)$ und z doppelt punktierte Ebene D schlicht, denn konvergieren die Geraden $H_n \ni (0,0)$ gegen eine Gerade J, so gilt aus Stetigkeitsgründen $\lim\limits_{n \to \infty} C_{H_n} = C_J$.

Die Kurven aus K^{Γ} bilden also zusammen mit den gewöhnlichen Geraden das System der Kreise einer sphärischen Kreisebene $\mathbb{K}_K$, die genau dann sogar eine Möbiusebene ist, wenn K in den Punkten (a_i, b_i) differenzierbar ist.

(t') Die volle Automorphismengruppe Θ von $\mathbb{K}_K$ zerfällt über $\Gamma_c = \mathrm{T} \cdot \Xi_c$ mit der Standgruppe Θ_K des Kreises K. Da sich Γ_c nicht in die Gruppe $\{z \mapsto hz + d;\ h \neq 0,\ d \in \mathbb{C}\}$ einbetten läßt, ist $\mathbb{K}_K$ nicht isomorph zur klassischen Möbiusgeometrie, und es gilt $\dim \Theta = 3$ (Strambach [1970], § 4). Γ_c ist also die Zusammenhangskomponente und daher insbesondere ein Normalteiler von Θ. Dann läßt aber Θ die Gesamtheit der Geraden $y = \text{const}$ und $x = \text{const}$ als Ganzes invariant und die Bahn $(0,0)^{\Theta_K}$ ist eine Teilmenge des Quadrupels $\{(0,0), (\tilde{e}, 1), (a_1, b_1), (a_2, b_2)\}$. Daher gilt $|\Theta_K| \leq 8$. Da jede nichtinvolutorische Affinität endlicher Ordnung eine Drehung ist, ist Θ_K eine Untergruppe der Diedergruppe der Ordnung 8. Enthielte Θ_K ein Element der Ordnung 4, so läge die Affinität $\delta = [(x, y) \mapsto (y, -x)]$ in Θ, was nicht möglich ist, weil δ die Gruppe Γ_c nicht normalisieren kann. Daher gilt $\Theta = \Gamma_c \cdot \Phi$, wobei die Gruppe Φ in der Menge der Affinitäten $\{(x, y) \mapsto (\pm x, \pm y)\}$ enthalten ist. Die übrigen Behauptungen über Θ ergeben sich durch leichte Rechnungen.

(u') Ist $k = 1, 4$, so sei

$$g_k^{\varepsilon_k}(x) = b_i \left(\frac{x}{a_i}\right)^{\varepsilon_k} \tag{33}$$

mit festem $c < \varepsilon_k \in \mathbb{R}$ und $i = 1$ bzw. $i = 2$, je nachdem ob $k = 1$ oder $k = 4$ gilt. Ist $h = 2, 3$, so sei

$$g_h(x) = 1 - (1 - b_i)\left(\frac{x}{a_i}\right)^{\varepsilon_h} \tag{34}$$

mit festem $c<\varepsilon_h\in\mathbb{R}$ und $i=1$ bzw. $i=2$, je nachdem ob $h=2$ oder $h=3$ gilt. Durch Nachrechnen überzeugt man sich, daß das Quadrupel der Funktionen g_l den Bedingungen (3) bis (6) des Hauptsatzes genügt. Um (7) zu verifizieren, reicht es wegen $F_l^{t_1,t_2}(0)=0$ zu zeigen, daß jede Funktion $F_l^{t_1,t_2}$ nur an der Stelle $x=0$ ihr einziges Minimum bzw. Maximum annimmt, je nachdem ob $l=1,4$ oder $l=2,3$ gilt. Verschwindet an einer Stelle x_0 die Ableitung von $F_l^{t_1,t_2}$, so ist

$$\left(\frac{t_2^{-1}(x_0+x_l^{(t_2)})}{t_1^{-1}(x_0+x_l^{(t_1)})}\right)^{\varepsilon_l-1}=\left(\frac{t_1}{t_2}\right)^{c-1}=\left(\frac{t_2^{-1}x_l^{(t_2)}}{t_1^{-1}x_l^{(t_1)}}\right)^{\varepsilon_l-1}, \tag{35}$$

woraus $x_0=0$ folgt. Man stellt fest, daß

$$\left(\frac{t_2^{-1}x_l^{(t_2)}}{t_1^{-1}x_l^{(t_1)}}\right)^{\varepsilon_l-2}-\left(\frac{t_1}{t_2}\right)^{c-2}>0 \tag{36}$$

genau dann gilt, wenn die zweite Ableitung von $F_l^{t_1,t_2}$ an der Stelle 0 positiv bzw. negativ ist, je nachdem ob $l=1,4$ oder $l=2,3$ ist. Wegen (35) folgt aus (36) die Beziehung

$$\left(\frac{t_1}{t_2}\right)^{\frac{(c-1)(\varepsilon_l-2)}{\varepsilon_l-1}}-\left(\frac{t_1}{t_2}\right)^{c-2}>0,$$

woraus sich $\varepsilon_l>c$ ergibt. Daher beschreibt das Quadrupel der Funktionen $g_l^{\varepsilon_l}$ eine Kurve K, die zu einer echten Kreisebene $\mathbb{K}_K$ führt. Hält man die Zahlen a_i, b_i fest und variiert die Parameter ε_l, so erhält man zu jedem $c>1$ überabzählbar viele nichtisomorphe Modelle.

(v′) Es sei $a_1<0<a_2$ und $0<b_1, b_2, \delta<1$ sowie $\varepsilon>c-1$.

Ist $l=1,4$, so setzen wir

$$g_l^{\varepsilon,\delta}(x)=(\operatorname{sign} a_l)\, d_l \int_0^x \frac{|\xi|^\varepsilon}{|a_l-\xi|^\delta}\, d\xi \quad \text{mit} \quad a_1\leqq x\leqq 0 \quad \text{bzw.} \quad 0\leqq x\leqq a_2,$$

je nachdem ob $l=1$ oder $l=4$ ist; die Zahl $d_l>0$ ist dabei so zu wählen, daß $b_l=g_l^{\varepsilon,\delta}(a_l)$ gilt. (Wegen der Konvergenz der auftretenden Integrale vgl. Courant, S. 216.)

Ist $l=2,3$, so setzen wir

$$g_l^{\varepsilon,\delta}(x)=d_l \int_{a_l}^x \frac{|\xi|^\varepsilon}{|a_l-\xi|}\, d\xi+b_l \quad \text{mit} \quad a_1\leqq x\leqq 0 \quad \text{bzw.} \quad 0\leqq x\leqq a_2,$$

je nachdem ob $l=2$ oder $l=3$ ist; $d_l>0$ werde so gewählt, daß $g_l^{\varepsilon,\delta}(0)=1$ gilt. Man stellt fest, daß

$$\frac{dg_l^{\varepsilon,\delta}}{dx}=g_l'=\frac{r(l)\,|x|^\varepsilon}{|a_i-x|^\delta}$$

gilt, wobei $r(l)=-1$ für $l=1,3$ und $r(l)=1$ für $l=2,4$ ist und $i=1$ für $l=1,2$ sowie $i=2$ für $l=3,4$ gilt. Daraus folgt

$$g_l''=\frac{d^2 g_l^{\varepsilon,\delta}}{dx^2}=m(l)\left[\frac{\varepsilon\,|x|^{\varepsilon-1}}{|a_i-x|^\delta}+\frac{\delta\,|x|^\varepsilon}{|a_i-x|^{1+\delta}}\right]$$

mit $m(l)=1$ für $l=1,4$ und $m(l)=-1$ für $l=2,3$ sowie $i=1$ für $l=1,2$ und $i=2$ für $l=3,4$. Es gilt $g_l''\geqq 0$ bzw. $g_l''\leqq 0$, je nachdem ob $m(l)$ positiv oder negativ ist; der Wert 0 kann von g_l'' höchstens für $x=0$ angenommen werden.

Die Bedingung $|k_j(0)|=\infty$ aus (6) sagt aus, daß die Kurve K bzw. K^ε mit $\varepsilon=[(x,y)\mapsto(x-\tilde{e},y-1)]$ jede Bahn $B_{(x,y)}$ mit $y>0$ bzw. $y<0$ trifft. Dies folgt aber daraus, daß für $l=1,4$ die Beziehung $\lim\limits_{x\mapsto 0}\dfrac{g_l(x)}{x^c}=0$, für $l=2,3$ dagegen $\lim\limits_{x\mapsto 0}\dfrac{g_l(x+\tilde{e})-1}{x^c}=0$ gilt. Diese beiden Bedingungen sind in unserem Fall wegen $\varepsilon>c-1$ erfüllt, wovon man sich etwa mit Hilfe der l'Hospitalschen Regel überzeugt. Das strenge Abnehmen der Funktionen k_j ist damit äquivalent, daß $k_j'<0$ ist für alle inneren Punkte der jeweiligen Definitionsbereiche. Für $l=1,4$ ist $k_l'=c^{-1}(g_l)^{-c^{-1}-1}(c\,g_l-x\,g_l')$, für $l=2,3$ gilt, wenn man $\tilde{e}=0$ bedenkt, $k_l'=c^{-1}(1-g_l)^{-c^{-1}-1}[c(1-g_l)+x\,g_l']$.

Die Ableitung der Funktion $s_l=c\,g_l-x\,g_l'$ ist in unserem Fall wegen $\varepsilon>c-1$ für $l=1$ und $x\neq 0$ positiv, für $l=4$ und $x\neq 0$ negativ. Ebenso leicht ist nachzurechnen, daß die Ableitung der Funktion $s_l=c(1-g_l)+x\,g_l'$ für $l=2$ und $x\neq 0$ positiv, für $l=3$ und $x\neq 0$ negativ ausfällt. Wegen $s_l(0)=0$ für alle l ist daher (6) erfüllt. Um (7) zu verifizieren, reicht es wegen $F_l^{t_1,t_2}(0)=0$ zu zeigen, daß jede Funktion $F_l^{t_1,t_2}$ nur für $x=0$ ihr einziges Minimum bzw. Maximum annimmt, je nachdem ob $l=1,4$ oder $l=2,3$ gilt. Bezeichnet

$$K_l^{t_j}(x)=t_j^{c-1}\,\frac{|t_j^{-1}x+t_j^{-1}x_l^{(t_j)}|^\varepsilon}{|a_i-t_j^{-1}x-t_j^{-1}x_l^{(t_j)}|^\delta}$$

mit $i=1$ für $l=1,2$ und $i=2$ für $l=3,4$, so ist

$$\Delta F_l^{t_1,t_2}=\frac{dF_l^{t_1,t_2}}{dx}=r(l)[K_l^{t_2}-K_l^{t_1}],$$

wobei $r(l)=1$ für $l=2,4$ und $r(l)=-1$ für $l=1,3$ gilt. Wegen $t_1^{-1}<t_2^{-1}$ ist $t_1^{-1}x<t_2^{-1}x$ bzw. $t_1^{-1}x>t_2^{-1}x$, je nachdem ob $x>0$ oder $x<0$ gilt. Da $\Delta F_l^{t_1,t_2}(0)=0$ ist (vgl. (7)), folgt $K_l^{t_2}(x)>K_l^{t_1}(x)$ für $x>0$ und $K_l^{t_2}(x)<K_l^{t_1}(x)$ für $x<0$; daher hat $\Delta F_l^{t_1,t_2}$ nur eine Nullstelle.

Bedenkt man, daß wegen (7) die Beziehung

$$|a_i-t_j^{-1}x_l^{(t_j)}|^{-\delta}=t_j^{1-c}\,|t_j^{-1}x_l^{(t_j)}|^{-\varepsilon}$$

mit $i=1$ für $l=1,2$ und $i=2$ für $l=3,4$ besteht, so erhält man

$$\begin{aligned}\Delta^2 F_l^{t_1,t_2}(0)=\frac{d^2F_l^{t_1,t_2}}{dx^2}(0)=m(l)\big[&\varepsilon\big(t_2^{-1}|t_2^{-1}x_l^{(t_2)}|^{-1}-t_1^{-1}|t_1^{-1}x_l^{(t_1)}|^{-1}\big)\\&+\delta\big(t_2^{(1-c)\delta^{-1}}|t_2^{-1}x_l^{(t_2)}|^{-\varepsilon\delta^{-1}}-t_1^{(1-c)\delta^{-1}}|t_1^{-1}x_l^{(t_1)}|^{-\varepsilon\delta^{-1}}\big)\big]\end{aligned}$$

mit $m(l)=1$ für $l=1,4$ und $m(l)=-1$ für $l=2,3$.

Da $t_1 > t_2$ ist, umfaßt $K^{\gamma_{t_1}}$ die Kurve $K^{\gamma_{t_2}}$, wenn die γ_{t_i} von der Form (31) sind. Daher ist

$$|t_2^{-1} x_l^{(t_2)}| < |t^{-1} x_l^{(t_1)}| \quad \text{und} \quad t_2^{-1} |t_2^{-1} x_l^{(t_2)}|^{-1} > t_1^{-1} |x_l^{(t_1)}|^{-1}.$$

Daraus folgt $\Delta^2 F_l^{t_1, t_2}(0) > 0$ für $l = 1, 4$ und $\Delta^2 F_l^{t_1, t_2}(0) < 0$ für $l = 2, 3$, und es gilt (7).

Hält man etwa die Zahlen a_i fest und variiert die Parameter ε_l und δ_l in zulässigen Grenzen, so erhält man zu jedem c überabzählbar viele nichtisomorphe Modelle von Möbiusebenen.

(C) Zum Schluß behandeln wir den Fall $\Gamma = \mathrm{T} \cdot \Pi$. Sei K ein nicht durch ∞ gehender Kreis einer Kreisebene $\mathbb{K}$, die Γ als eine Automorphismengruppe zuläßt. Es gilt dann $\Gamma_K = 1$, weil K endlichen Durchmesser hat.

(a″) Die streng konvexe Jordankurve K gehe durch den Punkt $(0, 0)$ und habe dort $y = 0$ als Stützgerade. Weiter darf man annehmen, daß K mit $(0, 1)$ inzidiert und dort $y = 1$ als Stützgerade besitzt, denn $\Gamma_{(0,0)}$ operiert auf den von der x-Achse verschiedenen Geraden transitiv und die Streckungen normalisieren die Gruppe Γ.

(b″) Die Bahnen $B_{(x,y)} = \{(w, v)\}$ von Punkten (x, y) mit $y \neq 0$ unter der Standgruppe $\Gamma_{(0,0)}$ lassen sich durch $w(v) = \frac{x}{y} v + v \ln \frac{v}{y}$ beschreiben, wobei v die positiven bzw. negativen reellen Zahlen durchläuft, je nachdem ob $y > 0$ bzw. $y < 0$ ist.

Nun läßt sich genau wie in (h′) des Teils (B) zeigen, daß

$$|K \cap B_{(x,y)}| \leqq 1 \quad \text{und} \quad |K^\varepsilon \cap B_{(x,y)}| \leqq 1 \quad \text{mit} \quad \varepsilon = [(x, y) \mapsto (x, y-1)]$$

gilt.

(c″) Angenommen, die Kurve K sei in einem Punkt t nicht differenzierbar, in dem sie keine waagrechte Stützgerade besitzt. Dann besitzt aber K in t eine eindeutig bestimmte vordere Tangente T_1 und eine eindeutig bestimmte hintere Tangente T_2 mit $T_1 \neq T_2$ (Haupt-Künneth, S. 63). Ist nun τ die Translation, die t auf $(0, 0)$ abbildet, so kommt man wie in (i′) des Teils (B) zu einem Widerspruch, und die Kurve K ist in jedem Punkt stetig differenzierbar, in welchem sie keine waagrechte Stützgerade besitzt.

(d″) K läßt sich in zwei „konvexe" Bögen C_j zerlegen, die durch die stetig differenzierbaren Funktionen h_j mit den Eigenschaften (8) beschrieben werden. Die Bedingung (9) ist damit äquivalent, daß C_1 von links konvex, C_2 dagegen von links konkav ist (Ostrowski I, S. 253).

(e″) Die Gültigkeit von $|B_{(x,y)} \cap K| \leqq 1$ und $|B_{(x,y)} \cap K^\varepsilon| \leqq 1$, wobei $\varepsilon = [(x, y) \mapsto (x, y-1)]$ bedeutet, ist zu der strengen Monotonie der Funktionen k_j aus (10) äquivalent, denn jede der Bahnen $B_{(x,y)}$ hat mit der Ordinatenachse genau den Punkt $(0, y\, e^{-x/y})$ gemeinsam.

(f″) Wie in (m′) aus (B) läßt sich nun zeigen, daß für $y > 0$ jede Bahn $B_{(x,y)}$ mit K, für $y < 0$ dagegen jede Bahn $B_{(x,y)}$ mit K^ε, wobei $\varepsilon = [(x, y) \mapsto (x, y-1)]$

ist, genau einen Punkt gemeinsam hat, und dies ist, in analytischer Sprechweise ausgedrückt, mit $|k_j(0)|=\infty$ äquivalent.

(g″) Der Kreis $K^{\gamma_t}=K_t$ mit

$$\gamma_t=[(x,y)\mapsto(e^t x+t\,e^t y,\, e^t y)]\in\Gamma_{(0,0)} \tag{37}$$

wird durch das Paar der Funktionen $h_j^t(y)=t\,y+e^t h_j(e^{-t}y)$ beschrieben, wobei $0\leqq y\leqq e^t$ gilt.

Wegen der Konvexität von h_1 und der Konkavität von h_2 existiert genau eine Stelle $y_t^{(j)}$, an der h_1 ihr Minimum bzw. h_2 ihr Maximum annimmt; in den Punkten $(h_j(y_t^{(j)}), y_t^{(j)})$ hat dann K_t eine zur y-Achse parallele Stützgerade. Es seien t_i zwei reelle Zahlen, die den in (11) an sie gestellten Forderungen genügen. Dann haben die Kurven $K_{t_i}^{\rho_i}$ mit

$$\rho_i=\left[(x,y)\mapsto\left(x-h_j^{t_i}(y_{t_i}^{(j)}),\, y-y_{t_i}^{(j)}\right)\right]$$

in $(0,0)$ die y-Achse als Stützgerade, und es gilt $K_{t_2}^{\rho_2}\cap K_{t_1}^{\rho_1}=(0,0)$. Da der mit $(0,0)$ inzidente Bogen von $K_{t_i}^{\rho_i}$ durch die Funktion

$$\tilde f_j^{t_i}(y)=t_i(y+y_{t_i}^{(j)})+e^{t_i}h_j[e^{-t_i}(y+y_{t_i}^{(j)})]-h_j^{t_i}(y_{t_i}^{(j)})\quad\text{mit}\quad -y_{t_i}^{(j)}\leqq y\leqq e^{t_i}-y_{t_i}^{(j)}$$

beschrieben wird, hat jede Funktion $\tilde F_j^{t_1,t_2}=\tilde f_j^{t_2}-\tilde f_j^{t_1}$, die auf dem Durchschnitt der Definitionsbereiche von $\tilde f_j^{t_i}$ erklärt ist, eine einzige Nullstelle, und es gilt $\tilde F_1^{t_1,t_2}\leqq 0$ sowie $\tilde F_2^{t_1,t_2}\geqq 0$ (vgl. Hilfssatz 1).

(h″) Sei nun umgekehrt eine Kurve K vorgegeben, die durch die Funktionen h_j, $j=1,2$, mit den Eigenschaften (8) bis (11) beschrieben wird. Wir zeigen, daß K ein strukturbestimmender Kreis einer Kreisebene $\mathbb{K}_K$ ist, auf der $\mathrm{T}\cdot\Pi$ als Zusammenhangskomponente der vollen Automorphismengruppe wirkt. Die Beziehung (10) sichert, daß die Menge der Kurven $K^\Pi\cup K^{\varepsilon\Pi}$ wobei $\varepsilon=[(x,y)\mapsto(x,y-1)]$ ist, zusammen mit der Geraden $y=0$ die in $(0,0)$ punktierte affine Ebene $\mathbf{D}$ schlicht überdeckt.

(i″) K umfaßt die Kurve K^γ, falls $\gamma\in\{(x,y)\mapsto(e^t x+t\,e^t y,\, e^t y);\ t<0\}$ gilt. Es sei $(0,1)^\gamma=(e_1,e_2)$ und L die Gerade $y-e_2=\dfrac{e_2-1}{e_1}(x-e_1)$. Gehört nun λ zur Menge $\mathfrak{Y}$ der Translationen, die (e_1,e_2) auf einen Punkt des von (e_1,e_2) und $(0,1)$ begrenzten offenen Intervalls von L abbilden, so gilt $K\cap K^{\gamma\lambda}=\emptyset$, und $K^{\gamma\lambda}$ gehört wegen (h″) ganz zu dem von K begrenzten kompakten Bereich.

(j″) Wir wollen zeigen, daß von den zwei Kurven $K=C_1\cup C_2$ und $K^\varphi=C_1^\varphi\cup C_2^\varphi$ mit $\varphi\in\Gamma$, die im Punkt t mit $(0,0)\neq t\neq(0,1)$ gemeinsame Stützgerade S haben, die eine die andere umfaßt. Der Punkt t gehöre zu C_k; dann liegt er auch auf C_k^φ. Sei nun ρ dasjenige Element aus Γ, das t auf $(0,0)$ so abbildet, daß S^ρ die y-Achse ist. Da $\Gamma=\Pi\cdot\mathrm{T}$ ist, sichert (11), daß $|C_k^{\varphi\rho}\cap C_k^\rho|=1$ und daher auch $|C_k^\varphi\cap C_k|=1$ gilt. Wegen der Transitivität von Γ auf K^Γ darf man voraussetzen, daß

$$\varphi\,\rho=[(x,y)\mapsto(e^{s_1}x+s_1 e^{s_1}y+u_1,\, e^{s_1}y+v_1)]$$

und

$$\rho = [(x, y) \mapsto (e^{s_2} x + s_2 e^{s_2} y + u_2, e^{s_2} y + v_2)]$$

mit (*) $s_1 < s_2$ wählbar sind.

Hätte $C_k^{\varphi\rho}$ mit $C_j^\rho \neq C_k^\rho$ Punkte gemeinsam, so läge einer von den Punkten $(0,0)^{\varphi\rho}$ und $(0,1)^{\varphi\rho}$ — er sei mit $r^{\varphi\rho}$ bezeichnet — außerhalb des von K^ρ begrenzten kompakten Bereichs. Der Absolutbetrag der Abszisse von $r^{\varphi\rho}$ wäre dann größer als der von r^ρ.

Ist nun τ die Translation, die $r^{\varphi\rho}$ auf r^ρ abbildet, so läge $(0,0)^\tau \in K^{\varphi\rho\tau}$ außerhalb des von K^ρ begrenzten kompakten Bereichs $\mathfrak{B}$. Da wegen (*) die Kurve $K^{\varphi\rho\tau}$ in der Nähe von r^ρ innerhalb von $\mathfrak{B}$ verläuft, erhalten wir einen Widerspruch zu (h″), und es gilt $C_k^\varphi \cap K = t$.

Angenommen, es sei $|K \cap K^\varphi| \geqq 2$. Ein zu t auf K direkt benachbarter Punkt aus $K \cap K^\varphi$ sei z. Mit ρ bezeichnen wir diejenige Translation $(x, y) \mapsto (x+a, y)$, die $(0,1)^\varphi$ auf einen Punkt der Geraden L aus (i″) abbildet. Da $(0,0)$ im Innern von K liegt, gehört $K^{\varphi\rho}$ einerseits zu $K^{\mathfrak{D}}$, andererseits befindet sich aber einer von den Punkten t^ρ und z^ρ außerhalb des von K begrenzten kompakten Bereichs, was (i″) widerspricht.

(k″) Nun beweist man weiter wie in (s′), daß die Kurven von $K^{\Pi\mathrm{T}}$ zusammen mit den gewöhnlichen Geraden ein System der Kreise einer Kreisebene $\mathbb{K}_K$ bilden. $\mathbb{K}_K$ ist sogar genau dann eine Möbiusebene, wenn K in den Punkten $(0,0)$ und $(0,1)$ differenzierbar ist. Dies trifft aber genau dann zu, wenn $h_1'(1) = h_2'(0) = \infty$ gilt, denn $h_1'(0) = h_2'(1) = -\infty$ ist wegen (10) stets erfüllt.

(l″) Die volle Automorphismengruppe Θ von $\mathbb{K}_K$ zerfällt über $\Gamma = \mathrm{T} \cdot \Pi$ mit der Standgruppe Θ_K des Kreises K. Da sich Γ nicht in die Gruppe $PSL_2(\mathbb{C})$ einbetten läßt, ist $\mathbb{K}_K$ nie isomorph zur klassischen Möbiusgeometrie, und es gilt $\dim \Theta = 3$ (Strambach [1970], § 4). Die restlichen Behauptungen über Θ ergeben sich durch Berechnung des Normalisators von Γ in der Gruppe der Affinitäten.

(m″) Nun wollen wir Beispiele von Möbiusebenen konstruieren, auf denen Γ jeweils als Gruppe von Automorphismen operiert. Es seien b_j zwei reelle Zahlen mit $-1 < b_2 < 0 < b_1 < 1$. Wir setzen

$$h_j = b_j^{-1}[y \ln y + (1-y) \ln (1-y)] \quad \text{mit} \quad 0 \leqq y \leqq 1 \tag{38}$$

und rechnen nach, daß $h_j'' = b_j^{-1} y^{-1} \dfrac{1}{1-y}$ sowie

$$k_1' = y^{-1} e^{-b_1^{-1}[\ln y + (1-y) y^{-1} \ln (1-y)]} \{y + b_1^{-1} \ln (1-y)\} \quad \text{mit} \quad 0 \leqq y \leqq 1$$

und

$$k_2' = y^{-1} e^{-b_2^{-1}[y^{-1} \ln (1+y) - \ln(-y)]} \{y + b_2^{-1} \ln (1+y)\} \quad \text{mit} \quad -1 \leqq y \leqq 0$$

gilt. Wegen $|b_j| < 1$ folgt $k_j'(y) < 0$ für $y \neq 0$ und mit Hilfe der l'Hospitalschen Regel auch $\ln |k_j(0)| = \infty$. Daher erfüllen die Funktionen (38) die Bedingungen (8) bis (10).

Hätte

$$\frac{d^2 \tilde{F}_j^{t_1,t_2}}{dy^2} = b_j^{-1}\{(y+y_{t_2}^{(j)})^{-1}[1-e^{-t_2}(y+y_{t_2}^{(j)})]^{-1} - (y+y_{t_1}^{(j)})^{-1}[1-e^{-t_1}(y+y_{t_1}^{(j)})]^{-1}\}$$

eine Nullstelle, so dürfte die Diskriminante D der Gleichung

$$y^2(e^{-t_1}-e^{-t_2})+2y(y_{t_1}^{(j)}e^{-t_1}-y_{t_2}^{(j)}e^{-t_2})+y_{t_2}^{(j)}-y_{t_1}^{(j)}+e^{-t_1}[y_{t_1}^{(j)}]^2-e^{-t_2}[y_{t_2}^{(j)}]^2=0$$

nicht negativ sein. Es gilt aber

$$D=(y_{t_1}^{(j)}e^{-t_1}-y_{t_2}^{(j)}e^{-t_2})^2-(e^{-t_1}-e^{-t_2})[y_{t_2}^{(j)}-y_{t_1}^{(j)}-e^{-t_2}(y_{t_2}^{(j)})^2+e^{-t_1}(y_{t_1}^{(j)})^2],$$

woraus sich durch elementare Umformungen

$$D=(y_{t_2}^{(j)}-y_{t_1}^{(j)})[e^{-t_1-t_2}(y_{t_2}^{(j)}-y_{t_1}^{(j)})-e^{-t_1}+e^{-t_2}]$$

ergibt.

Da $\frac{e^{t(1-b_j)}}{(1+e^{-b_j t})^2}(1-b_j+e^{-b_j t})>0$ die Ableitung der Funktion $y_t^{(j)}=\frac{e^{t(1-b_j)}}{1+e^{-b_j t}}$ ist (vgl. (11)), gilt $D<0$ genau dann, wenn

$$\left[\frac{e^{t_2(1-b_j)}}{1+e^{-b_j t_2}}-\frac{e^{t_1(1-b_j)}}{1+e^{-b_j t_1}}\right]\cdot e^{-t_1-t_2}-e^{-t_1}+e^{-t_2}<0$$

erfüllt ist, und dies ist, wie elementare Umformungen zeigen, mit

$$e^{-t_1}(1+e^{-b_j t_1})-e^{-t_2}(1+e^{-b_j t_2})>0$$

gleichwertig. Daher ist die Diskriminante D insbesondere dann negativ, wenn die Funktion $l(t)=e^{-t}+e^{-t(1+b_j)}$ monoton fallend ist, und dies trifft wegen $l'=e^{-t}[-2-b_j]<0$ zu. Also gilt für die für $y \geqq -y_{t_1}^{(j)}$ definierten Funktionen $\frac{d^2 \tilde{F}_j^{t_1,t_2}}{dy^2}$ sogar $\frac{d^2 \tilde{F}_1^{t_1,t_2}}{dy^2}<0$ und $\frac{d^2 \tilde{F}_2^{t_1,t_2}}{dy^2}>0$. Die Funktionen $\frac{\tilde{F}_j^{t_1,t_2}}{dy}$ sind dann streng monoton und verschwinden jeweils nur für $y=0$. Daher hat $\tilde{F}_1^{t_1,t_2}$ bzw. $\tilde{F}_2^{t_1,t_2}$ nur ein Maximum bzw. Minimum; dieses wird jeweils für $y=0$ angenommen, woraus $\tilde{F}_1^{t_1,t_2} \leqq 0$ und $\tilde{F}_2^{t_1,t_2} \geqq 0$ folgt. Damit ist auch (11) nachgeprüft, und wir erhalten zu verschiedenen Paaren von zulässigen Werten b_j nichtisomorphe Möbiusebenen, auf denen $\mathrm{T}\cdot\Pi$ operiert.

(n″) Zum Schluß wollen wir zeigen, daß $\Gamma=\mathrm{T}\cdot\Pi$ auch auf Kreisebenen operiert, in denen das Berühraxiom verletzt ist. Es seien $a_2<0<a_1$ zwei reelle Zahlen mit $|a_j|>1$. Wir setzen $h_1=a_1 y \ln y$ sowie $h_2=a_2(1-y)\ln(1-y)$ und rechnen nach, daß diese Funktionen (8) bis (10) erfüllen, denn es ist

$$h_1''=a_1 y^{-1} \quad \text{und} \quad h_2''=a_2(1-y)^{-1}$$

sowie

$$k_1'=y^{(1-a_1)}[1-a_1]<0 \quad \text{mit} \quad 0<y\leqq 1$$

und

$$k_2'=(-y)^{a_2}(1+a_2)<0 \quad \text{mit} \quad -1\leqq y<0.$$

Nun wollen wir (11) verifizieren. Da die y-Achse die Tangente an die durch $\tilde{f}_j^{t_i}$ beschriebenen Bogenstücke im Punkt $(0,0)$ ist, reicht es also für unsere Zwecke zu zeigen, daß jede Funktion $\tilde{F}_1^{t_1,t_2}$ bzw. $\tilde{F}_2^{t_1,t_2}$ nur für $y=0$ ihr einziges Maximum bzw. Minimum annimmt. Es ist

$$\text{(39)} \qquad \ln y_{t_i}^{(1)} = t_i(1-a_1^{-1})-1$$

sowie

$$\text{(40)} \qquad y_{t_i}^{(2)} = e^{t_i} - e^{t_i(1+a_2)-1}.$$

Man rechnet nach, daß

$$\frac{d^2 F_1^{t_1,t_2}}{dy^2} = a_1 (y+y_{t_2}^{(1)})^{-1} (y+y_{t_1}^{(1)})^{-1} [y_{t_1}^{(1)} - y_{t_2}^{(1)}]$$

gilt. Da $e^{-t_i}(y+y_{t_i}^{(1)})$ zum Definitionsbereich von h_1 gehört, ist $(y+y_{t_i}^{(1)})>0$, und $\frac{d^2 F_1^{t_1,t_2}}{dy^2}$ ist genau dann negativ, wenn (41) $y_{t_1}^{(1)} < y_{t_2}^{(1)}$ gilt. Daß aber die Beziehung (41) stets erfüllt ist, folgt aus (39). Es gilt weiter

$$\frac{d^2 F_2^{t_1,t_2}}{dy^2} = a_2 e^{-t_1-t_2} [1-e^{-t_2}(y+y_{t_2}^{(2)})]^{-1} [1-e^{-t_1}(y+y_{t_1}^{(2)})]^{-1} \cdot (e^{t_1} - e^{t_2} + y_{t_2}^{(2)} - y_{t_1}^{(2)}).$$

Da $1-e^{-t_i}(y+y_{t_i}^{(2)})$ zum Definitionsbereich von h_2 gehört, ist

$$[1-e^{-t_i}(y+y_{t_i}^{(2)})]>0, \quad \text{und} \quad \frac{d^2 F_2^{t_1,t_2}}{dy^2}$$

ist wegen $a_2<0$ genau dann positiv, wenn (42) $e^{t_1}-e^{t_2}+y_{t_2}^{(2)}-y_{t_1}^{(2)}<0$ gilt. Daß die Ungleichung (42) erfüllt ist, folgt sofort, wenn man (40) berücksichtigt. Also ist $\frac{d^2 F_1^{t_1,t_2}}{dy^2}<0$ und $\frac{d^2 F_2^{t_1,t_2}}{dy^2}>0$, und $\frac{dF_j^{t_1,t_2}}{dy}$ sind streng monotone Funktionen mit genau einer Nullstelle für $y=0$; an dieser nimmt $F_1^{t_1,t_2} \leqq 0$ bzw. $F_2^{t_1,t_2} \geqq 0$ ihr ein einziges Maximum bzw. Minimum an, und daher ist (11) erfüllt. Die Funktionen h_j beschreiben also zusammen eine geschlossene Kurve K, und diese führt zu einer echten Kreisebene $\mathbb{K}_K$, auf der $\mathrm{T} \cdot \Pi$ operieren kann. Durch Variation der Parameter a_j bekommt man so überabzählbar viele nichtisomorphe Modelle.

Weitere Kreisebenen, in denen das Berühraxiom verletzt ist und die Γ als Gruppe von Automorphismen zulassen, erhält man, indem man die Funktionen h_j so wählt, daß h_m die in (m″), aber h_n mit $n \neq m$ die in (n″) beschriebene Gestalt haben.

Literatur

Blaschke, W.: Kreis und Kugel. Leipzig: Verlag von Veit 1916.

Bonnesen, F., Fenchel, W.: Theorie der konvexen Körper. Berlin: Springer 1935.

Buckel, W.: Eine Kennzeichnung des Systems aller Kreise mit nicht verschwindendem Radius. J. Reine Angew. Math. **191**, 13–30 (1953).

Courant, R.: Vorlesungen über Differential- und Integralrechnung I. Berlin-Göttingen-Heidelberg: Springer 1961.
Ewald, G.: Aus konvexen Kurven bestehende Möbiusebenen. Abh. Math Sem. Univ. Hamburg **30**, 179–187 (1967).
Haupt, O., Künneth, H.: Geometrische Ordnungen. Grundlehren-Bd. 133. Berlin-Heidelberg-New-York: Springer 1967.
Heemert, van A.: Zur Kennzeichnung der Systeme der Kreise und der Kegelschnitte. J. Reine Angew. Math. **194**, 183–189 (1955).
Hering, Ch.: Eine Klassifikation der Möbius-Ebenen. Math. Z. **87**, 252–262 (1965).
Kamke, E.: Differentialgleichungen I. Lösungsmethoden und Lösungen, (6. Aufl.) Leipzig: Akademische Verlagsgesellschaft 1959.
Natanson, B. A.: Theorie der Funktionen einer reellen Veränderlichen. Berlin: Akademie Verlag 1961.
Ostrowski, A.: Vorlesungen über Differential- und Integralrechnung I. Funktionen einer Variablen. Basel-Stuttgart: Birkhäuser 1960.
– Vorlesungen über Differential- und Integralrechnung II. Differentialrechnung auf dem Gebiete mehrer Variablen. Basel-Stuttgart: Birkhäuser 1961.
Rosenthal, A.: Die Translationsordnung ebener Kurven. Monatsh. Math. Phys. **45**, 76–91 (1937).
Salzmann, H.: Kompakte zweidimensionale projektive Ebenen. Math. Ann. **145**, 401–428 (1962).
– Zur Klassifikation topologischer Ebenen. III. Abh. Math. Sem. Univ. Hamburg **28**, 250–261 (1965).
– Topological planes. Advances Math. **2**, 1–60 (1967).
Strambach, K.: Eine Charakterisierung der klassischen Geometrien. Arch. der Math. **18**, 539–544 (1967).
– Zur Klassifikation von Salzmann-Ebenen mit dreidimensionaler Kollineationsgruppe. Math. Ann. **179**, 15–30 (1968).
– Sphärische Kreisebenen. Math. Z. **113**, 266–292 (1970).
– Salzmann-Ebenen mit punkttransitiver dreidimensionaler Kollineationsgruppe. Indagationes Math. **32**, 254–267 (1970a).
– Sphärische Kreisebenen mit einfacher Automorphismengruppe. Erscheint in Geometriae dedicata, Vol. 1 (1971).
Strubecker, K.: Differentialgeometrie I. Kurventheorie der Ebene und des Raumes. Sammlung Göschen 1113/1113a. Berlin 1955.
Wölk, R. D.: Topologische Möbiusebenen. Math. Z. **93**, 311–333 (1966).

Dr. Karl Strambach
Math. Inst. der Universität Tübingen
D-7400 Tübingen, Wilhelmstraße 7

(Eingegangen am 25. Juni 1971)

Math. Z. 124, 315–318 (1972)

On Incidence Matrices of Finite Projective and Affine Spaces

William M. Kantor

It is well-known that the rank of each incidence matrix of all points *vs.* all e-spaces of a finite d-dimensional projective or affine space is the number of points of the geometry, where $1 \leqq e \leqq d-1$ (see [1], p. 20). In this note we shall generalize this fact:

Theorem. *Let $0 \leqq e < f \leqq d-e-1$, and let $M_{e,f}$ be an incidence matrix of all e-spaces vs. all f-spaces of $PG(d,q)$ or $AG(d,q)$. Then the rank of $M_{e,f}$ is the number of e-spaces of the geometry.*

We shall only prove the theorem in the case of $AG(d,q)$. The projective case is similar and simpler. We note that the same proof shows that, *if $1 \leqq e < f \leqq d-e$, an incidence matrix of all e-sets vs. all f-sets of a set of d points has rank* $\binom{d}{e}$.

The relevant definitions are found in [1], §§ 1.3 and 1.4. The dimension of a subspace X of a projective space will be denoted $\dim(X)$. The empty subspace has dimension -1.

Proof. Let E and F denote any e-space and f-space, respectively. Set

$$(E,F)=\begin{cases}1 & \text{if } E\subset F\\ 0 & \text{if } E\not\subset F.\end{cases}$$

For a suitable ordering of e-spaces and f-spaces we have $M_{e,f}=((E,F))$. Let $R(E)$ denote the row of $M_{e,f}$ corresponding to the e-space E.

Assume that there is a nontrivial dependence relation among the rows of $M_{e,f}$. This may be assumed to have the form

$$R(E^*)=\sum_{E\neq E^*} a(E)\,R(E) \tag{1}$$

for some e-space E^*, where each $a(E)$ is a real number. Let Γ be the group of all collineations of $AG(d,q)$ taking E^* to itself. If $\alpha\in\Gamma$ then $(E^\alpha, F^\alpha)=(E,F)$.

Then (1) implies that, for all F,

$$(E^*, F) = (E^*, F^\alpha) = \sum_E a(E^\alpha)(E^\alpha, F^\alpha)$$
$$= \sum_E a(E^\alpha)(E, F),$$

so that $R(E^*) = \sum_{E \neq E^*} a(E^\alpha) R(E)$. Thus,

$$\begin{aligned} |\Gamma|\, R(E^*) &= \sum_\Gamma \sum_{E \neq E^*} a(E^\alpha) R(E) \\ &= \sum_{E \neq E^*} R(E) \sum_\Gamma a(E^\alpha). \end{aligned} \tag{2}$$

Let $\mathfrak{S}$ be the set of all ordered pairs (i, j) of integers satisfying the conditions $-1 \leqq i, j \leqq e-1$ and $j = i$ or $i+1$. Order $\mathfrak{S}$ lexicographically: $(i_1, j_1) < (i_2, j_2)$ if either $i_1 < i_2$ or $i_1 = i_2$ and $j_1 < j_2$.

We now embed $AG(d, q)$ in $PG(d, q)$. Let H_∞ be the hyperplane at infinity. If $(i, j) \in \mathfrak{S}$, let $\mathscr{E}(i, j)$ be the set of e-spaces $E \not\subset H_\infty$ such that $\dim(E \cap E^* \cap H_\infty) = i$ and $\dim(E \cap E^*) = j$. The $\mathscr{E}(i, j)$ are precisely the orbits of Γ of e-spaces other than E^*. Fix $E_{ij} \in \mathscr{E}(i, j)$. Then (2) implies that

$$\begin{aligned} |\Gamma|\, R(E^*) &= \sum_{\mathfrak{S}} \sum_{\mathscr{E}(i,j)} R(E) \sum_\Gamma a(E^\alpha) \\ &= \sum_{\mathfrak{S}} \sum_{\mathscr{E}(i,j)} R(E)\, |\mathscr{E}(i,j)| \sum_\Gamma a(E_{ij}^\alpha) \\ &= \sum_{\mathfrak{S}} \Big[|\mathscr{E}(i,j)| \sum_\Gamma a(E_{ij}^\alpha) \Big] \sum_{\mathscr{E}(i,j)} R(E). \end{aligned}$$

There is thus a dependence relation of the form

$$R(E^*) = \sum_{\mathfrak{S}} b_{ij} \sum_{\mathscr{E}(i,j)} R(E) \tag{3}$$

with b_{ij} real.

Let $(m, n) \in \mathfrak{S}$. Since $e + f \leqq d - 1$ there is an f-space F_{mn} such that

$$\dim(F_{mn} \cap E^* \cap H_\infty) = m \quad \text{and} \quad \dim(F_{mn} \cap E^*) = n.$$

As $n \leqq e - 1$, $F_{mn} \not\supset E^*$. By (3),

$$0 = \sum_{\mathfrak{S}} b_{ij} \sum_{\mathscr{E}(i,j)} (E, F_{mn}) \tag{4}$$

whenever $(m, n) \in \mathfrak{S}$.

Let (i, j), $(m, n) \in \mathfrak{S}$. If $(E, F_{mn}) = 1$ for some $E \in \mathscr{E}(i, j)$ then $E \subset F_{mn}$ implies that $i \leqq m, j \leqq n$. In particular, if $(i, j) > (m, n)$ then

$$\sum_{\mathscr{E}(i,j)} (E, F_{mn}) = 0. \tag{5}$$

Also, there is an e-space $E \subset F_{ij}$ such that $E \cap E^* \cap H_\infty = F_{ij} \cap E^* \cap H_\infty$ and $E \cap E^* = F_{ij} \cap E^*$. Thus,

$$\sum_{\mathscr{E}(i,j)} (E, F_{ij}) \neq 0. \tag{6}$$

By (4) and (5),

$$0 = \sum_{(i,j) \leqq (m,n)} b_{ij} \sum_{\mathscr{E}(i,j)} (E, F_{mn}) \tag{7}$$

for all $(m, n) \in \mathfrak{S}$. This is a system of $|\mathfrak{S}|$ equations in the $|\mathfrak{S}|$ unknowns b_{ij}, $(i,j) \in \mathfrak{S}$. If we use the ordering of $\mathfrak{S}$, the coefficient matrix is triangular, with nonzero diagonal entries by (6). It follows that $b_{ij} = 0$ for all $(i,j) \in \mathfrak{S}$, contradicting (3). This completes the proof.

We note incidentally that the matrices $M_{e,f}$ have the following property: if

$$0 \leqq e < f < g \leqq d-1,$$

then $M_{e,f} M_{f,g} = c_{e,f,g} M_{e,g}$, where $c_{e,f,g}$ is the number of f-spaces contained in a g-space G and containing an e-space contained in G.

Corollary 1. *If* $0 \leqq e \leqq d-e-1$, *then a collineation of* $PG(d,q)$ *induces similar permutations on the sets of e-spaces and* $(d-e-1)$*-spaces.*

Proof. [1], p. 22.

Corollary 2. *Let* Γ *be a collineation group of* $PG(d,q)$ *or* $AG(d,q)$. *If*

$$0 \leqq e < f \leqq d-e-1,$$

then Γ *has at least as many orbits of f-spaces as e-spaces. Moreover, these quantities are equal in the projective case if* $f = d-e-1$.

Proof. [1], pp. 20–22.

Corollary 3. *Let* Γ *be a collineation group of* $PG(d,q)$ *or* $AG(d,q)$ *transitive on f-spaces. If* $0 \leqq e < f \leqq d-e-1$ *then* Γ *is transitive on e-spaces, and its rank on f-spaces is at least as large as that on e-spaces. Moreover, these ranks are equal in the projective case if* $f = d-e-1$.

Proof. See the proof of [2], Theorem 4.4.

We remark that, by using [1], p. 21, together with the analogue of our theorem for incidence matrices of subsets of a finite set, we also obtain an elementary proof of Theorem 1 of Livingstone and Wagner [3].

Parts of these corollaries are due to Wagner [5] in the projective case.

Added in Proofs. The corollaries can also be easily deduced from the following facts. Let θ_e be the permutation character obtained from the action of the full collineation group of $PG(d,q)$ or $AG(d,q)$ on the e-spaces of the geometry, where $0 \leqq e \leqq d-1$. If $0 \leqq e \leqq d/2$, then $\theta_{e-1} \subset \theta_e = \theta_{d-1-e}$ for $PG(d,q)$, and $\theta_{e-1} \subset \theta_{d-e} \subset \theta_e$ for $AG(d,q)$. These inclusions can be proved by calculating (θ_e, θ_f) whenever $0 \leqq e \leqq f \leqq d-1$. In the projective case, this result and this calculation are essentially contained in [4], pp. 276–278.

References

1. Dembowski, P.: Finite geometries. Berlin-Heidelberg-New York: Springer 1968.
2. Kantor, W. M.: Automorphism groups of designs. Math. Z. **109**, 246–252 (1969).
3. Livingstone, D., Wagner, A.: Transitivity of finite permutation groups on unordered sets. Math. Z. **90**, 393–403 (1965).
4. Steinberg: R.: A geometric approach to the representations of the full linear group over a Galois field. Trans. Amer. Math. Soc. **71**, 274–282 (1951).
5. Wagner, A.: Collineations of finite projective spaces as permutations on the sets of dual subspaces. Math. Z. **111**, 249–254 (1969).

Prof. William M. Kantor
Department of Mathematics
University of Oregon
Eugene, Oregon 97403
USA

(Received September 30, 1970)

Math. Z. 124, 319–342 (1972)

Über die Approximation von linearen Gleichungen zweiter Art und Eigenwertproblemen in Banach-Räumen

HANSGEORG JEGGLE

Einleitung

In dieser Arbeit wird die konstruktive Lösung von allgemeinen Gleichungen zweiter Art und Eigenwertaufgaben der Gestalt

$$(zL-C)\,u=v$$

mit Hilfe einer Folge approximierender Probleme

$$(zL_\iota-C_\iota)\,u_\iota=v_\iota, \qquad \iota\in\Lambda_0,$$

untersucht. Für die inhomogenen Aufgaben wird ein Äquivalenzsatz bewiesen, der eine konstruktive Existenzaussage unter gleichzeitiger Behandlung der Existenz- und Konvergenzfrage für die Näherungsprobleme gibt. Auf seiner Grundlage kann die Konvergenz von Spektralmengen, insbesondere von Eigenwerten, von Eigen- und Hauptvektoren zu diesen Eigenwerten und von Spektralprojektionen und ihren Bildern studiert werden.

Die erwähnten Resultate können für eine weite Klasse von Operatoren bewiesen werden, nämlich für jene, bei denen L linear und in einem Banachschen Raum E dicht definiert, C überdies abgeschlossen und L in gewissem Sinn klein gegen C ist (vgl. 1.(6)). Im ersten Abschnitt werden die Spektraleigenschaften dieser Paare (L, C) untersucht. Dies betrifft insbesondere die Charakterisierung der algebraischen Vielfachheit, der spektralen Teilräume und die Reduzierung der Operatoren auf ihnen.

Zahlreiche für die Anwendungen wichtige Näherungsverfahren (z.B. Quadraturformel- oder Differenzenverfahren) führen zu Approximationen in Banachschen Räumen E_ι, $\iota\in\Lambda_0$ (Λ_0 eine unendliche Indexmenge), außerhalb des Ausgangsraumes E. Für die dabei benötigte Definition einer Konvergenz von Elementen $u_\iota\in E_\iota$, $\iota\in\Lambda$, gegen ein Element $u\in E$ benützen wir den von Stummel [18] eingeführten diskreten Konvergenzbegriff. Mit ihm beschäftigt sich unter anderem der zweite Abschnitt.

Von Petryshyn wurde in [16] die Klasse von linearen, beschränkten Operatoren studiert, die bei den Projektionsverfahren die wesentliche Rolle spielt. Diese Approximationen und Operatoren sind für die Anwendungen nicht ausreichend, man denke an Verfahren, wie die oben erwähnten, die außerhalb des Ausgangsraumes führen. Vor allem solche Fälle, bei denen simultane Störungen von Operatoren, Grundgebieten und gegebenenfalls Daten vorliegen, lassen es wünschenswert erscheinen, auf die Einschränkungen, denen die Approximation – proper mappings [16] unterliegen, zu verzichten. Dies wird im dritten Abschnitt durch eine geeignete Erweiterung der Bedingung (H) von Petryshyn [16] auf der Grundlage der diskreten Konvergenz möglich. Auf diese Weise können wir diskrete Approximation und Störungsprobleme gleichzeitig untersuchen. Beim Studium der Approximierbarkeit parameterunabhängiger Gleichungen erster Art hat Grigorieff [10] die entsprechende Verallgemeinerung von (H) vorgenommen.

Zentrales Resultat im vierten Abschnitt ist ein Satz über die Konvergenz der Eigenwerte ihrer algebraischen Vielfachheit nach. Diese Behauptung wurde für verschiedene Konstellationen mit beschränkten Operatoren bereits von Anselone-Palmer [1], Atkinson [2], Grigorieff [8] und Jeggle [13] bewiesen. Umfassende Resultate für Operatoren $zA-B$ mit B beschränkt und A vollstetig erzielt Stummel in den Arbeiten [19, 20], die vorerwähnten Ergebnisse lassen sich den seinen unterordnen. Soweit dies aufgrund der in dieser Arbeit veränderten Voraussetzungen überhaupt möglich ist, können die Resultate aus [19] in unsere eingeordnet werden, ebenso die von Grigorieff in den Arbeiten [7] und [9].

Diese Arbeit enthält Verallgemeinerungen eines Teils der Habilitationsschrift [14] des Verfassers. Dort wird vor allem die Approximation von Eigenwertaufgaben für nichtlinear vom Parameter abhängige Operatoren untersucht. Den Herren R. D. Grigorieff und W. Wendland danke ich herzlich für diese Arbeit fördernde Diskussionen und Hinweise.

1. Einige Sätze aus der Spektraltheorie

Es seien E und F Banachsche Räume über demselben Körper $\mathbb{K}$, wobei $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$. Mit $L(E,F)$ bezeichnen wir die Menge der in E dicht definierten linearen Operatoren in F, mit $C(E,F)$ die Teilmenge der abgeschlossenen Operatoren aus $L(E,F)$ und mit $B(E,F)$ den Banachschen Raum der beschränkten linearen Operatoren von E in F. Die Definitionsmenge eines Operators T schreiben wir $D(T)$, das Bild $R(T)$ und den Nullraum $N(T)$. $\mathscr{U}(K)$ bedeutet den Umgebungsfilter einer kompakten Menge $K\subset\mathbb{K}$ im Sinne der üblichen Topologie von $\mathbb{K}$.

Zu dem Paar $(L_1,L_2)\in L(E,F)\times L(E,F)$ definieren wir in $\mathbb{K}$ die $L(E,F)$-wertige Funktion

$$A(z):=zL_1-L_2,\qquad z\in\mathbb{K}, \tag{1}$$

wobei $D(L_2)\subset D(L_1)$ gelte. Bei Eigenschaften, die die gesamte Operatorschar $\{A(z)|z\in\mathbb{K}\}$ betreffen, werden wir auch sagen, A besitze diese Eigenschaft.

Als *Resolventenmenge* $\rho(A)$ von A bezeichnen wir die Menge der Zahlen $z\in\mathbb{K}$, für die $A^{-1}(z)\in B(F,E)$ existiert. Die Menge $\mathscr{C}_{\mathbb{K}}\rho(A)$ heißt *Spektrum* $\sigma(A)$ von A. Die in $\rho(A)$ definierte, $B(F,E)$-wertige Funktion $A^{-1}(.)$ heißt *Resolventenfunktion*, ihre Werte *Resolventen*.

Jede Zahl $z_0\in\mathbb{K}$, zu der Elemente $u\neq 0$ aus $D(A)$ existieren mit

$$A(z_0)\,u=0 \tag{2}$$

heißt *Eigenwert* von A, das zugehörige Element u *Eigenvektor* von A zu z_0. Die $\dim N(A(z_0))$ nennen wir die *geometrische Vielfachheit* von z_0. Ferner definieren wir zu jedem Eigenwert z_0 von A *Hauptvektoren k-ter Stufe* $u^k \in D(A)$, $k=1,2,\ldots$, durch die Pakete $(u^1,\ldots,u^k)$, deren Elemente mittels der Rekursion

$$(3) \qquad A(z_0)\,u^1=0, \qquad A(z_0)\,u^j+L_1\,u^{j-1}=0, \qquad j=2,3,\ldots,k,$$

erklärt sind. Ist m die höchste Stufe, die zu z_0 vorkommt, so nennen wir m den *Anstieg* $\alpha(z_0)$ dieses Eigenwerts. Die lineare Hülle aller Hauptvektoren heißt *algebraische Eigenmannigfaltigkeit* $M(z_0, A)$ von A zu z_0, ihre Dimension *algebraische Vielfachheit* von z_0. $M(\sigma, A)$ bezeichnet die lineare Hülle aller Hauptvektoren zu einer aus endlich vielen Eigenwerten bestehenden Menge σ.

Die wichtigsten Spektraleigenschaften von Paaren $(L_1, L_2)\in B(E,F)\times C(E,F)$ sind wohlbekannt [7, 18]. Wir werden eine allgemeinere Klasse von Paaren untersuchen, die wir durch die folgende Bedingung beschreiben:

(*B*) *Sei* $(L, C)\in L(E,F)\times C(E,F)$ *mit* $D(C)\subset D(L)$. *Die in* $\mathbb{K}$ *durch* $T(z):=zL-C$ *definierte Operatorschar besitze die folgenden Eigenschaften:*

$$(4) \qquad \bigwedge_{z\in\mathbb{K}} T(z)\in C(E,F), \qquad \bigwedge_{z\in\rho(T)} LT^{-1}(z)\in B(F,F).$$

Eine hinreichende Bedingung für (*B*) lautet folgendermaßen:

(5) *Wenn* $(L, C)\in L(E,F)\times C(E,F)$ *mit* $D(C)\subset D(L)$ *die Eigenschaft*

$$(6) \qquad \bigwedge_{\varepsilon>0}\ \bigvee_{K(\varepsilon)>0}\ \bigwedge_{u\in D(C)} \|Lu\|\leqq\varepsilon\,\|Cu\|+K(\varepsilon)\,\|u\|$$

besitzt, dann erfüllt $T(z)=zL-C$ *die Bedingung* (4).

Beweis. Die Zugehörigkeit von $T(z)$ zu $C(E,F)$ für $z\in\mathbb{K}$ ist aus der Störungstheorie bekannt [6]. Ferner haben wir wegen (6) und $C=zL-T(z)$ für $v\in F$ und $z\in\rho(T)$

$$\|LT^{-1}(z)\,v\|\leqq\varepsilon\big(|z|\,\|LT^{-1}(z)\,v\|+\|v\|\big)+K(\varepsilon)\,\|T^{-1}(z)\,v\|.$$

Für $\varepsilon|z|\leqq p<1$ folgt die Behauptung:

$$\|LT^{-1}(z)\,v\|\leqq\big((1-p)^{-1}(\varepsilon+K(\varepsilon)\,\|T^{-1}(z)\|)\big)\,\|v\|.$$

Wir bemerken, daß (6) und damit (*B*) für alle Paare $(L, C)\in B(E,F)\times C(E,F)$ trivialerweise erfüllt ist. Ein weitergehendes Beispiel für einen Operator, der (*B*) erfüllt, erhält man, wenn man im Sobolewschen Raum $W^{m,p}(\mathbb{R}^+)$, $1<p<\infty$, die folgenden Differentialoperatoren betrachtet:

$$C:=\sum_{k=0}^{m} a_k D^k, \qquad L:=\sum_{k=0}^{m-1} b_k D^k$$

mit $a_m^{-1}\in L^\infty(\mathbb{R}^+)$ und $a_k\in L^\infty(\mathbb{R}^+)$ für $k=0,\ldots,m$, sowie meßbaren und lokal zu $L^p(\mathbb{R}^+)$ gehörigen b_k, $k=0,\ldots,m-1$, für die gilt:

$$\sup_{s\in\mathbb{R}^+}\int_s^{s+1}|b_k(t)|^p\,dt<\infty, \qquad k=0,\ldots,m-1.$$

Das auf diese Weise definierte Paar (L, C) erfüllt die Bedingung (6) für alle $u \in W^{m,p}(\mathbb{R}^+)$ in der $L^p(\mathbb{R}^+)$-Norm [6].

Die Bedingung (5) ist z.B. immer dann erfüllt, wenn $D(C) \subset D(L)$ und L abschließbar ist. Wir beweisen nun einige Eigenschaften der Paare (L, C):

(7) *Das Paar (L, C) erfülle (B). Wenn die Resolventenmenge $\rho(T) \neq \emptyset$ ist, so ist sie offen und die Resolventenfunktion $T^{-1}(.)$ ist in ihr analytisch.*

Beweis. Sei $z_0 \in \rho(T)$. Dann haben wir:

$$T(z) = \big(I_F - (z_0 - z)\, LT^{-1}(z_0)\big)\, T(z_0).$$

Aufgrund von (B) existiert zu jedem $\varepsilon < 1$ eine Zahl $\delta > 0$, so daß

$$|z - z_0|\, \|LT^{-1}(z_0)\| \leqq \varepsilon$$

für alle z mit $|z - z_0| < \delta$. Das Lemma von Banach [21] erbringt damit $R\big(T(z)\big) = F$ für $|z - z_0| < \delta$ und die Abschätzung

(8) $$\|T^{-1}(z)\| \leqq (1-\varepsilon)^{-1}\, \|T^{-1}(z_0)\|, \qquad |z - z_0| < \delta,$$

d.h., $\{z \,|\, |z - z_0| < \delta\} \subset \rho(T)$. Also ist $\rho(T)$ offen. Weiterhin gibt es zu jedem $z_0 \in \rho(T)$ einen Kreis $K \in \mathscr{U}(\{z_0\})$, in dem die Resolventenfunktion in die konvergente Neumannsche Reihe

$$T^{-1}(z) = \sum_{j=0}^{\infty} \big(T^{-1}(z_0)(LT^{-1}(z_0))^j\big)(z - z_0)^j$$

entwickelt werden kann. Damit ist auch die Analytizität nachgewiesen.

(9) *Wenn (L, C) die Bedingung (B) erfüllt, dann gilt für $z_1, z_2 \in \rho(T)$ die Resolventengleichung*

$$T^{-1}(z_1) - T^{-1}(z_2) = -(z_1 - z_2)\, T^{-1}(z_1)\, LT^{-1}(z_2).$$

Zum weiteren Studium der Spektraleigenschaften von (L, C) benötigen wir noch einige Begriffe. Unter einer *Spektralmenge* von A verstehen wir eine Teilmenge σ' von $\sigma(A)$, die abgeschlossen in $\mathbb{K}$ ist. Im Falle $\mathbb{K} = \mathbb{C}$ bezeichnen wir als *Cauchy-Gebiet* Δ zu einer Spektralmenge σ' eine Umgebung $\Delta \in \mathscr{U}(\sigma')$, die aus höchstens endlich vielen Komponenten besteht, die paarweise disjunkte abgeschlossene Hüllen besitzen. Dabei sei der Rand $\partial\Delta$ die Vereinigung endlich vieler geschlossener, rektifizierbarer Jordan-Kurven mit der üblichen Orientierung. Ferner sei die folgende Bedingung erfüllt:

$$\sigma' = \sigma(A) \cap \Delta, \qquad \partial\Delta \subset \rho(A).$$

Schließlich definieren wir noch Kurvenintegrale. Sei J eine rektifizierbare Jordan-Kurve in $\mathbb{C}$, f eine komplexwertige stetige Funktion auf J und φ eine stetige beschränkte Abbildung von J in einen komplexen Banachschen Raum E. Dann existiert in E das Integral

$$I(\varphi) := \int_J f(z)\, \varphi(z)\, dz,$$

wenn $I(\varphi)$ wie üblich als Grenzwert von Näherungssummen in der Normtopologie von E definiert wird. Nach dem vorstehenden existiert insbesondere zu jeder Spektralmenge σ' von T das Resolventenintegral

$$\int_{\partial\Delta} f(z)\, A^{-1}(z)\, dz \tag{10}$$

in $B(F, E)$, wenn f auf dem Rand eines Cauchy-Gebiets Δ zu σ' stetig ist.

Die Definition (10) zeigt auch, daß

$$F(\sigma', T) := \frac{1}{2\pi i} \int_{\partial\Delta} LT^{-1}(z)\, dz \tag{11}$$

ein Operator aus $B(F, F)$ ist, falls (L, C) die Bedingung (B) erfüllt. Aus (9) erkennt man, daß $F(\sigma', T)$ sogar eine Projektion in F ist. Falls L *beschränkt* ist, erhalten wir auf dieselbe Weise, daß

$$E(\sigma', T) := \frac{1}{2\pi i} \int_{\partial\Delta} T^{-1}(z)\, L\, dz \tag{12}$$

eine stetige Projektion in E ist. In diesem Fall kann man mit Hilfe der Spektralprojektionen (11) und (12) Teilräume E_1, E_2 und F_1, F_2 von E bzw. F charakterisieren und auf ihnen den Operator T reduzieren [7, 19, 21].

Wir beschäftigen uns hier mit dem Fall, daß L im allgemeinen unbeschränkt ist und (B) erfüllt, also (s. (12)) nur die Spektralprojektion $F(\sigma', T)$ zu σ' existiert. Bei der Charakterisierung der spektralen Teilräume E_1, E_2 müssen wir daher anders vorgehen als es üblich ist.

Für den Rest dieses Abschnitts setzen wir voraus, daß (B) durch (L, C) erfüllt wird und $\rho(T) \neq \emptyset$ ist. Ferner sei stets $\mathbb{K} = \mathbb{C}$.

Wenn $L, C \in B(E, F)$ und $L^{-1} \in B(F, E)$ existiert, dann bezeichnen wir als Resolventenmenge $\rho(T)$ die ursprünglich so definierte Menge, vereinigt mit dem Punkt ∞.

Zunächst erklären wir wie in [7] mit Hilfe von $y \in \rho(T)$, $y \neq \infty$, die bijektive Abbildung $\varphi\colon \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ durch

$$\varphi(z) := (z-y)^{-1} \quad \text{für } z \in \overline{\mathbb{C}} \setminus \{\infty, y\}, \qquad \varphi(\infty) := 0, \ \varphi(y) := \infty. \tag{13}$$

Ferner definieren wir die Abbildung $S\colon F \to F$ und die Operatorschar $V(z)$ durch

$$S := -LT^{-1}(y); \qquad V(z) := z\, I_F - S, \qquad z \in \mathbb{C}. \tag{14}$$

Nun behaupten wir folgenden Hilfssatz:

(15) *Die Operatorschar $V(z)$ ist aus $B(F, F)$ und es gilt $\rho(V) = \varphi(\rho(T))$.*

Beweis. Der erste Teil der Behauptung ist aus (14) und (B) unmittelbar ersichtlich. Damit haben wir auch $\varphi(y) = \infty \in \rho(V)$. Weiter besteht die Identität

$$V(\varphi(z)) = \varphi(z)\, T(z)\, T^{-1}(y), \qquad z \neq y. \tag{16}$$

Ihr entnimmt man, daß $V(\varphi(z))$ für $z \in \rho(T) \cap \mathbb{C} \setminus \{y\}$ eine bijektive Abbildung von F auf sich ist. Denn

$$\varphi(z)\, V^{-1}(\varphi(z)) = T(y)\, T^{-1}(z) = I_F + (y-z)\, LT^{-1}(z).$$

Also haben wir $\varphi(\rho(T) \cap \mathbb{C}) \subset \rho(V)$.

Sei umgekehrt $\zeta := \varphi(z) \in \rho(V) \cap \mathbb{C} \setminus \{0\}$. Aufgrund von (16) ist dann $z \in \rho(T)$. Wegen (14) und (B) zählt auch ∞ zu $\rho(V)$. Aufgrund (13) haben wir aber $\varphi(y) = \infty$ und $y \in \rho(T)$. Somit ist $\rho(V) \cap (\mathbb{C} \setminus \{0\}) \subset \varphi(\rho(T) \cap \mathbb{C})$.

Falls $\infty \in \rho(T)$, haben wir auch $0 \in \rho(V)$, denn S ist in diesem Falle eine beschränkte, bijektive Abbildung von F auf sich. $0 \in \rho(V)$ ist auch nur dann möglich, wenn $\infty \in \rho(T)$: Wenn L^{-1} nicht existiert, dann existiert auch S^{-1} nicht. Existiert zwar L^{-1} auf $R(L)$, ist jedoch $R(L) \neq F$, dann folgt $R(S) \neq F$. Existiert schließlich sogar $L^{-1} \in B(F, E)$, so heißt das $S^{-1} = -T(y)\, L^{-1} = y\, I_F + CL^{-1}$. Da $CL^{-1} \notin B(F, F)$, ist auch $S^{-1} \notin B(F, F)$.

Wir nehmen nun spektrale Zerlegungen von E und F vor und untersuchen die Restriktionen der Operatoren auf ihnen. Zunächst vereinbaren wir: Es seien σ_1 eine Spektralmenge von T mit $\infty \notin \sigma_1$, $\sigma_2 = \mathscr{C}_{\sigma(T)}\, \sigma_1$ und Δ ein Cauchy-Gebiet zu σ_1. Ferner bedeute:

$$F_1 := R(F(\sigma_1, T)), \qquad F_2 := N(F(\sigma_1, T)),$$
$$E_1 := T^{-1}(y)\, F_1, \qquad E_2 := T^{-1}(y)\, F_2,$$

und

$$T_j(z) := T(z)|E_j, \quad L_j := L|E_j, \quad C_j := C|E_j, \quad j = 1, 2.$$

Nun haben wir den folgenden Satz:

(17) *Die Restriktionen $T_1(z)$ und $T_2(z)$ sind Abbildungen von E_1 in F_1 bzw. E_2 in F_2. Es existieren $T_1^{-1}(z) \in B(F_1, E_1)$ und $T_2^{-1}(z) \in B(F_2, E_2)$ genau dann, wenn $z \in \mathscr{C}_{\mathbb{C}}\, \sigma_1$ bzw. $z \in \mathscr{C}_{\mathbb{C}}\, \sigma_2$. Ferner ist $R(L_1) = F_1$ und es existiert $L_1^{-1} \in B(F_1, E_1)$.*

Beweis. $V(z)$ kann in F bezüglich F_1 und F_2 in Teile $V_1(z)$ und $V_2(z)$ reduziert werden [21]. Aufgrund von (15) gilt für ihre Spektren $\sigma(V_1) = \varphi(\sigma_1)$ und $\sigma(V_2) = \varphi(\sigma_2)$. Aus (16) fließt

$$V_1(\varphi(z)) = \varphi(z)\, T_1(z)\, T^{-1}(y) \colon F_1 \to F_1, \qquad z \in \mathbb{C},\ z \neq y.$$

Also haben wir $R(T_1(z)) \subset F_1$ und $T_1^{-1}(z) \in B(F_1, E_1)$ genau, wenn $\varphi(z) \notin \sigma(V_1)$, damit ist äquivalent $z \notin \sigma_1$. Wieder gilt $0 \in \rho(V_1)$ genau, wenn $\infty \in \rho(T_1)$. Mit derselben Argumentation erschließt man aus (15) und (16) auch den $T_2(z)$ betreffenden Teil der Behauptung. Schließlich bedeutet $\infty \notin \sigma_1$, daß $0 \in \rho(V_1)$, d.h. für $S_1 := -LT^{-1}(y)$ gilt $S_1^{-1} \in B(F_1, F_1)$. Daher haben wir $L_1 \in L(E_1, F_1)$ mit $L_1^{-1} = -T^{-1}(y)\, S_1^{-1} \in B(F_1, E_1)$.

Nun werden wir noch auf den Zusammenhang von Polen der Resolventenfunktion und isolierten Eigenwerten eingehen. Zunächst zeigen wir:

(18) *Wenn z_0 ein Pol der Ordnung n der Resolventenfunktion $T^{-1}(.)$ ist, so ist auch $\varphi(z_0)$ ein Pol der Ordnung n von $V^{-1}(.)$.*

Beweis. Aus (16) erhalten wir die Identität

(19) $$T^{-1}(z)=(\varphi(z)\,T^{-1}(y))\,V^{-1}(\varphi(z)), \qquad z\in\rho(T)\setminus\{y\}.$$

Weiter gibt es einen Kreis $U\in\mathcal{U}(\{z_0\})$, so daß $\varphi(.)\,T^{-1}(y)$ in U, $T^{-1}(.)$ und $\tilde{V}^{-1}(.):=V^{-1}(\varphi(.))$ in $U\setminus\{z_0\}$ holomorph sind. Insbesondere besitze $\varphi(.)$ in U eine Taylor-Entwicklung und $T^{-1}(.)$ in $U\setminus\{z_0\}$ eine Laurent-Entwicklung. Aus (19) ist nun ersichtlich, daß z_0 ein Pol von $\tilde{V}^{-1}(.)$ ist. Wäre er von der Ordnung m, so folgte aus (19) für $z\in U\setminus\{z_0\}$

$$T^{-1}(z)=\sum_{j=1}^{m}(T^{-1}(y)\,\tilde{V}_{-j})(z-z_0)^{-j}+\tilde{t}(z),$$

wobei $\tilde{t}$ holomorph in U wäre. Wegen $T^{-1}(y)\,\tilde{V}_{-m}\neq 0$ muß dann $m=n$ sein.

Nun zeigen wir noch, daß $\tilde{V}^{-1}(.)$ genau dann einen Pol der Ordnung n in z_0 besitzt, wenn $\zeta_0=\varphi(z_0)$ ein Pol der Ordnung n von $V^{-1}(.)$ ist. Denn

$$V^{-1}(\zeta)=\tilde{V}^{-1}(z)=\sum_{j=1}^{n}\tilde{V}_{-j}(z-z_0)^{-j}+\tilde{v}(z)$$

in $U\setminus\{z_0\}$ mit $U\in\mathcal{U}(\{z_0\})$. Mit (13) gewinnt man hieraus

$$V^{-1}(\zeta)=\sum_{j=1}^{n}V_{-j}(\zeta-\zeta_0)^{-j}+v(\zeta)$$

in $\varphi(U\setminus\{z_0\})$. $\tilde{v}$ und v sind jeweils holomorph in U bzw. $\varphi(U)$. Dabei sind V_{-j} Linearkombinationen von $\tilde{V}_{-i}$, $j\leqq i\leqq m$, $j=1,\dots,m$. Umgekehrt ist aus (13) unmittelbar ersichtlich, daß $\tilde{V}^{-1}(.)$ in z_0 einen Pol der Ordnung n besitzt, wenn $V^{-1}(.)$ in $\zeta_0=\varphi(z_0)$ einen solchen hat.

Jeder Pol ζ_0 der Ordnung n von $V(.)$ ist ein Eigenwert von V, für den man

$$G_1:=R\left(\frac{1}{2\pi i}\int_{\partial D}V^{-1}(\zeta)\,d\zeta\right)=\{u\in F\,|\,V(\zeta_0)^n u=0\}$$

hat [21]. Nun kann man die Gleichung

$$G_1=\{u\in F\,|\,V(\zeta_0)\,u^k=u^{k-1},\ u^0=0,\ 1\leqq k\leqq r,\ 1\leqq r\leqq n,\ u^r=u\}$$

leicht überprüfen. Also können wir G_1 wegen (3) als algebraische Eigenmannigfaltigkeit $M(\zeta_0, V)$ von V zu ζ_0 charakterisieren. Nun behaupten wir:

(20) *Wenn z_0 ein Pol der Ordnung n von* $T^{-1}(.)$ ist, dann gilt

$$M(\varphi(z_0), V)=R(F(z_0, T))=F_1.$$

Beweis. Aufgrund von (18) ist $\zeta_0=\varphi(z_0)$ Pol der Ordnung n von $V^{-1}(.)$. Bekanntlich ist [21]

$$M(\zeta_0, V)=R\left(\frac{1}{2\pi i}\int_{\partial D}V^{-1}(\zeta)\,d\zeta\right),$$

wobei D ein Cauchy-Gebiet zu $\{\zeta_0\}=\{\varphi(z_0)\}$ ist. Mit Hilfe von (16) und einer Substitution der Integrationsveränderlichen erhalten wir

$$\int_{\partial D} V^{-1}(\zeta)\, d\zeta = \int_{\partial\Delta} L T^{-1}(z)\, dz.$$

Dabei ist D so zu wählen, daß Δ mit $\varphi(\Delta)=D$ Cauchy-Gebiet zu $\{z_0\}$ ist. Dies ist wegen (15) stets möglich.

Damit haben wir eine Charakterisierung von F_1 durch die Hauptvektoren von V zu $\varphi(z_0)$. Wir benötigen noch den folgenden Hilfssatz:

(21) *$T(y)$ vermittelt eine bijektive Abbildung der Hauptvektoren k-ter Stufe von T zum Eigenwert z_0 auf die von V zum Eigenwert $\varphi(z_0)$, $k=1,2,\ldots$.*

Beweis. Sei $(u^1,\ldots,u^k)\in D(T)^k$ ein Paket zu z_0. Dann existieren $v^j\in F$ mit $u^j=T^{-1}(y)\,v^j$, $j=1,\ldots,k$. Aus (3) und (16) ersehen wir

$$V(\varphi(z_0))\,v^j=-\varphi(z_0)\,S\,v^{j-1},\qquad j=1,\ldots,k,$$

mit $v^0=0$. Ein Induktionsschluß führt zu

$$V^i(\varphi(z_0))\,v^j=(-1)^i\varphi^i(z_0)\,S^i v^{j-i},\qquad i=0,\ldots,j;\ j=1,\ldots,k.$$

Wir haben zu zeigen: $V^k(\varphi(z_0))\,v^k=0$ und $V^i(\varphi(z_0))\,v^j\neq 0$ für $i=0,\ldots,j-1$ und $j=1,\ldots,k$. Die erste Bedingung ist klar, wäre die zweite falsch, so müßten gleichzeitig

$$S^i v^{j-i}=0,\qquad V^{j-i}(\varphi(z_0))\,v^{j-i}=0$$

wahr sein. Dies bedeutete, daß $v^{j-i}\neq 0$ zugleich Hauptvektor zu 0 und $\varphi(z_0)$ wäre. Das ist jedoch wegen (3) und (17) unmöglich. Also ist $v=T(y)\,u$ Hauptvektor k-ter Stufe von V zu $\varphi(z_0)$.

Umgekehrt nehmen wir an, v sei Hauptvektor k-ter Stufe von V zu $\varphi(z_0)$. Dann existiert $(v^1,\ldots,v^k)\in F_1^k$ mit $v^k=v$ und

$$V(\varphi(z_0))\,v^j=v^{j-1},\qquad j=1,\ldots,k,\ v^0=0.$$

Hierzu gibt es $(u^1,\ldots,u^k)\in E_1^k$ mit $u^j=T^{-1}(y)\,v^j$, $j=1,\ldots,k$. Aufgrund von (16) haben wir

$$\varphi(z_0)\,T(z_0)\,u^j=T(y)\,u^{j-1},\qquad j=1,\ldots,k,\ u^0=0.$$

Induktiv folgt aus dieser Beziehung

$$\varphi^2(z_0)\,T(z_0)\,u^j=-\sum_{i=1}^{j-1}(z_0-y)^{j-i-1}\,L\,u^i,\qquad j=1,\ldots,k.$$

Daraus erschließt man sukzessive, daß $(u^1,\ldots,u^k)$ ein Paket von T zu z_0 ist, d.h., daß $u=T^{-1}(y)\,v$ Hauptvektor k-ter Stufe von T zu z_0 ist.

Die Sätze (20) und (21) haben das folgende wichtige Korollar:

(22) *Wenn z_0 ein Pol der Ordnung n von $T^{-1}(.)$ ist, dann ist z_0 ein Eigenwert von T mit dem Anstieg $\alpha(z_0)=n$. Der Spektralraum E_1 von T zu z_0 ist dann identisch mit der algebraischen Eigenmannigfaltigkeit $M(z_0,T)$ von T zu z_0.*

In einem wichtigen Spezialfall haben wir ferner:

(23) *Wenn z_0 ein Pol der Ordnung n von $T^{-1}(.)$ und* $\dim N(T(z_0))<\infty$ *ist, dann gilt*

$$\dim M(z_0, T)=\dim R(F(z_0, T))<\infty.$$

Aufgrund von (3) und (22) ist nämlich $\dim M(z_0, T)<\infty$. Der Rest folgt aus (21).

Aus Referenzgründen führen wir noch das folgende Resultat über die Meromorphie von Resolventenfunktionen an. Der Beweis ist im wesentlichen in [12] enthalten. Dabei bedeutet $\Phi^l(E, F)$ die Menge der Semifredholm-Operatoren mit stetig projiziertem Bild.

(24) *Das Paar (L, C) erfülle die Bedingung (B), ferner sei $T(z)\in\Phi^l(E, F)$ für alle z aus einem Gebiet $G\subset\mathbb{C}$ und es existiere $z_0\in G$ mit $T^{-1}(z_0)\in B(F, E)$. Dann ist $T^{-1}(z)\in B(F, E)$ mit Ausnahme von höchstens abzählbar vielen Punkten, die sich nicht in G häufen und sämtlich Pole der Resolventenfunktion $T^{-1}(.)$ sind.*

Bemerkung. Wenn σ eine Menge ist, die endlich viele Pole von $T^{-1}(.)$ enthält und in einem Cauchy-Gebiet $\dim N(T(z))<\infty$ gilt, dann kann (23) auch für diese Situation formuliert werden:

$$\dim M(\sigma, T)=\dim R(F(\sigma, T))<\infty.$$

2. Diskrete Konvergenz

Wir stellen nun die wichtigsten Begriffe und Eigenschaften der diskreten Konvergenztheorie von Stummel zusammen und geben erste Anwendungen auf den hier untersuchten Problemkreis. Zum Teil halten wir uns eng an die Arbeit [18].

Im folgenden bedeutet Λ_0 eine abzählbare, linear geordnete Indexmenge. $\Lambda, \Lambda_i, i\in\mathbb{N}$, bedeuten Teilfolgen von Λ_0 und Λ'_0, Λ''_i Endstücke von Λ_0 bzw. $\Lambda_i, i\in\mathbb{N}$, beide nicht notwendig stets dieselben. Die Banachschen Räume E, $F, E_\iota, F_\iota, \iota\in\Lambda_0$, werden über demselben Körper $\mathbb{K}$ vorausgesetzt. $\|.\|$ bedeutet die Norm in sämtlichen Räumen, eine Unterscheidung ist nicht erforderlich.

Grundlage der Theorie ist folgende Definition:

(1) *Das Paar $(E, (E_\iota))$ heißt diskrete Approximation, wenn folgende Bedingungen erfüllt sind:*

(A 1) *Es gibt eine Abbildung $R\colon E\to\prod E_\iota$ mit*

$$\bigwedge_{u\in E}\ \bigvee_{(R_\iota u)\in\Pi E_\iota} Ru=(R_\iota u),$$

(A 2)
$$\bigwedge_{u\in E} \|R_\iota u\|\to\|u\| \qquad (\iota\in\Lambda_0),$$

(A 3)
$$\bigwedge_{u,v\in E}\ \bigwedge_{\alpha,\beta\in\mathbb{K}} \|\alpha R_\iota u+\beta R_\iota v-R_\iota(\alpha u+\beta v)\|\to 0 \qquad (\iota\in\Lambda).$$

In einer diskreten Approximation $(E, (E_\iota))$ wird folgender Konvergenzbegriff erklärt:

(2) *Die Folge* $(u_\iota) \in \prod E_\iota$ *heißt diskret stark konvergent gegen* $u \in E$, $u_\iota \to u$ $(\iota \in \Lambda_0)$, *wenn* $\|R_\iota u - u_\iota\| \to 0$ $(\iota \in \Lambda_0)$.

Die Spezifizierung „diskret stark" werden wir in der Regel weglassen.

Die Definitionen (1) und (2) sind so getroffen, daß diese Konvergenz insbesondere linear und stetig ist. Diese und andere Eigenschaften werden in [18] ausführlich nachgewiesen. Dort findet man auch zahlreiche Beispiele für diskrete Approximationen. Wir erwähnen nur, daß die übliche starke Konvergenz in Banachschen Räumen unter diese Theorie fällt. Wenn P_ι, $\iota \in \mathbb{N}$, stetige Projektionen in E mit $P_\iota \to I$ $(\iota \in \Lambda)$ sind, dann ist $(E, (P_\iota E))$ mit $R u := (P_\iota u)$ ein weiteres Beispiel.

Im Anschluß an die Definition der Konvergenz von Folgen von Elementen kann die von Folgen von Operatorscharen auf natürliche Weise vorgenommen werden. $(E, (E_\iota))$, $(F, (F_\iota))$ seien im folgenden stets diskrete Approximationen, die Operatoren $A(z)$, $A_\iota(z)$, $\iota \in \Lambda_0$, seien wie in 1.(1) erklärt und die Zahlen z_ι, $z_0 \in \mathbb{K}$.

(3) *Die Folge* $(A_\iota(z_\iota)) \in \prod L(E_\iota, F_\iota)$ *heißt diskret stark konvergent gegen* $A(z_0) \in L(E, F)$, $A_\iota(z_\iota) \to A(z_0)(\iota \in \Lambda)$, *wenn folgende Bedingung erfüllt ist:*

$$\bigwedge_{u \in D(A)} \ \bigwedge_{(u_\iota) \in \Pi D(A_\iota)} u_\iota \to u\,(\iota \in \Lambda_0) \succ A_\iota(z_\iota)\, u_\iota \to A(z_0)\, u(\iota \in \Lambda_0).$$

Mit den folgenden drei Definitionen beschreiben wir Eigenschaften von Folgen $(A_\iota(z_\iota))$ bzw. den Zusammenhang zwischen einer solchen Folge und einer Operatorschar $A(z)$. Sind die Operatoren parameterunabhängig, dann fallen unsere Definitionen mit denen von Stummel [18] zusammen.

(4) *Die Folge* $(A_\iota(z_\iota)) \in \prod B(E_\iota, F_\iota)$ *heißt stabil, wenn* $\alpha \geqq 0$ *existiert mit* $\|A_\iota(z_\iota)\| \leqq \alpha$ *für* $\iota \in \Lambda$.

Die Stabilität von $(A_\iota(z_\iota))$ für eine Folge (z_ι) zieht offenbar die für jede beschränkte Zahlenfolge $y_\iota \in \mathbb{K}$, $\iota \in \Lambda$, nach sich.

(5) *Die Folge* $(A_\iota(z_\iota)) \in \prod L(E_\iota, F_\iota)$ *heißt invers stabil, wenn gilt:*

$$\bigvee_{\gamma > 0} \ \bigvee_{\Lambda' \subset \Lambda} \iota \in \Lambda' \succ \bigwedge_{u \in D(A_\iota)} \gamma\, \|u\| \leqq \|A_\iota(z_\iota)\, u\|.$$

(6) *Das Paar* $(A(z), (A_\iota(z))) \in L(E, F) \times \prod L(E_\iota, F_\iota)$ *heißt konsistent in einer Menge* $M \subset \mathbb{K}$, *wenn gilt:*

$$\bigwedge_{u \in D(A)} \ \bigvee_{(u_\iota) \in \Pi D(A_\iota)} \ \bigwedge_{z_0 \in M} \ \bigwedge_{(z_\iota) \in \Pi M} u_\iota \to u \wedge (z_\iota \to z_0 \succ A_\iota(z_\iota)\, u_\iota \to A(z_0)\, u) \qquad (\iota \in \Lambda_0).$$

Die Konsistenz von $(z L_1 - L_2, (z L_{1\iota} - L_{2\iota}))$ zieht die von $(L_1, (L_{1\iota}))$ und $(L_2, (L_{2\iota}))$ nach sich. Wenn die Folge $(L_{1\iota}) \in \prod B(E_\iota, F_\iota)$ stabil ist, genügt es, (6) für $z_\iota = z_0 = z \in M$, $\iota \in \Lambda$, zu definieren.

Eine Rechtfertigung für diese Definitionen gibt der folgende wichtige Satz, der im wesentlichen in [18] bewiesen wurde.

(7) *Seien* $A(z) \in L(E, F)$ *und* $(A_\iota(z_\iota)) \in \prod B(E_\iota, F_\iota)$. *Dann sind die Konsistenz des Paares* $(A(z), (A_\iota(z)))$ *und die Stabilität von* $(A_\iota(z_\iota))$ *notwendig und hinreichend*

für die Konvergenz von $A_\iota(z_\iota)\to A(z)(\iota\in\Lambda_0)$. *Aus ihr folgt, daß* $A(z)$ *auf* $D(A)$ *beschränkt ist.*

Nun vereinbaren wir noch, daß wir die Menge der Zahlen $z\in\mathbb{K}$, für die $(A_\iota(z))$ die inverse Stabilitätsbedingung (5) und die Forderung $R(A_\iota(z))=F_\iota$, $\iota\in\Lambda'$, erfüllt, als *Beschränktheitsmenge* $B^-(A_\iota)$ bezeichnen. Die Struktur von $B^-(T_\iota)$ für Folgen $(T_\iota(z))$, deren Elemente im folgenden Sinne gleichmäßig der Bedingung (B) unterliegen, erörtern wir im nachstehenden Satz.

(B_ι) *Seien* $(L_\iota, C_\iota)\in L(E_\iota, F_\iota)\times C(E_\iota, F_\iota)$ *mit* $D(C_\iota)\subset D(L_\iota)$ *für* $\iota\in\Lambda_0$. *Die in* $\mathbb{K}$ *durch* $T_\iota(z){:}=zL_\iota-C_\iota$, $\iota\in\Lambda_0$, *definierten Operatorscharen besitzen die folgenden Eigenschaften:*

(8) $$\bigwedge_{\iota\in\Lambda} T_\iota(z)\in C(E_\iota, F_\iota), \qquad \bigwedge_{z\in B^-(T_\iota)}\ \bigvee_{\Lambda'(z)\subset\Lambda}\ \bigvee_{\alpha(z)\geq 0}\ \bigwedge_{\iota\in\Lambda'(z)} \|L_\iota T_\iota^{-1}(z)\|\leqq\alpha(z).$$

Nun behaupten wir:

(9) *Die Folge von Paaren* $((L_\iota, C_\iota))$ *erfülle die Bedingung* (B_ι). *Wenn die Beschränktheitsmenge* $B^-(T_\iota)$ *nichtleer ist, dann ist sie offen und zu jeder kompakten Teilmenge* $K\cap B^-(T_\iota)$ *gibt es ein Endstück* Λ'' *und eine Zahl* $\gamma>0$ *mit*

(10) $$\bigwedge_{\iota\in\Lambda''} K\subset\rho(T_\iota), \qquad \bigwedge_{z\in K}\ \bigwedge_{\iota\in\Lambda''}\|T_\iota^{-1}(z)\|\leqq\gamma^{-1}.$$

Beweis. Da die Existenz von $z_0\in B^-(T_\iota)$ vorausgesetzt ist, gibt es ein Endstück $\Lambda'(z_0)$ und eine Zahl $\gamma(z_0)>0$ mit $z_0\in\rho(T_\iota)$ und $\|T_\iota^{-1}(z_0)\|\leqq\gamma^{-1}(z_0)$ für $\iota\in\Lambda'(z_0)$. Aufgrund von (8) gilt ferner $\|L_\iota T_\iota^{-1}(z_0)\|\leqq\alpha(z_0)$ für $\Lambda''(z_0)$. Also existiert zu jedem $\varepsilon<1$ eine Zahl $\delta(z_0)>0$, so daß für $\iota\in\Lambda''(z_0)$ gilt

$$|z-z_0|\,\|L_\iota T_\iota^{-1}(z_0)\|<\varepsilon$$

für alle z mit $|z-z_0|<\delta(z_0)$. Mit Hilfe des Lemmas von Banach [21] haben wir dann, daß $R(T_\iota(z))=F_\iota$ sowie

$$\|T_\iota^{-1}(z)\|\leqq(1-\varepsilon)^{-1}\,\|T_\iota^{-1}(z_0)\|\leqq((1-\varepsilon)\,\gamma(z_0))^{-1}$$

für alle $|z-z_0|<\delta(z_0)$ und alle $\iota\in\Lambda''(z_0)$. Also zieht $z_0\in B^-(T_\iota)$ nach sich, daß $|z-z_0|<\delta(z_0)$ zu $B^-(T_\iota)$ gehört. Mithin ist $B^-(T_\iota)$ offen.

Wenn $K\subset B^-(T_\iota)$ eine kompakte Menge ist, gibt es zu jedem Punkt $z\in K$ einen Kreis der beschriebenen Art. Nach dem Satz von Heine-Borel existiert eine endliche Überdeckung von K, z. B. durch die Kreise $k_1,\dots,k_t$ um Punkte $z_1,\dots,z_t\in K$. Setzt man

$$\gamma^{-1}{:}=\max_{j=1,\dots,k}((1-\varepsilon)\,\gamma(z_j))^{-1}, \qquad \Lambda''{:}=\bigcap_{j=1}^{k}\Lambda''(z_j),$$

so ist alles bewiesen.

Nach dem Muster von 1.(5) geben wir wieder eine hinreichende Bedingung für die Erfüllung von (B_ι).

(11) *Für alle* $\iota\in\Lambda$ *seien die Paare* $(L_\iota, C_\iota)\in L(E_\iota, F_\iota)\times C(E_\iota, F_\iota)$ *definiert mit* $D(C_\iota)\subset D(L_\iota)$ *und folgender Eigenschaft: Es existiere ein Endstück* $\Lambda'\subset\Lambda$,

so daß

(12) $$\bigwedge_{\varepsilon>0}\ \bigwedge_{\iota\in\Lambda'}\ \bigvee_{k(\varepsilon)>0}\ \bigwedge_{u\in D(C_\iota)} \|L_\iota u\| \leqq \varepsilon\,\|C_\iota u\| + k(\varepsilon)\,\|u\|.$$

Dann besitzt die Folge $((L_\iota, C_\iota))$ *die Eigenschaften* (8).

Beweis. In Anbetracht des Beweises zu 1.(5) brauchen wir nur zu zeigen: Für alle $\iota\in\Lambda'$, $v\in F_\iota$ und beliebiges $z\in B^-(T_\iota)$ haben wir aufgrund von (12) und der Definition von $B^-(T_\iota)$

$$\|L_\iota T_\iota^{-1}(z)\,v\| \leqq \varepsilon(|z|\,\|L_\iota T_\iota^{-1}(z)\,v\| + \|v\|) + k(\varepsilon)\,\gamma^{-1}(z)\,\|v\|.$$

Für $\varepsilon\,|z| \leqq p < 1$ und alle $\iota\in\Lambda'$ folgt hieraus

$$\|L_\iota T_\iota^{-1}(z)\,v\| \leqq \big((1-p)^{-1}(\varepsilon + k(\varepsilon)\,\gamma^{-1}(z))\big)\,\|v\|.$$

In Hilbertschen Räumen H, H_ι, $\iota\in\Lambda_0$, kann man auch eine diskrete schwache Konvergenz von Folgen von Elementen definieren [18]:

(13) *Sei* (H, H_ι) *eine diskrete Approximation mit Hilbertschen Räumen* H, H_ι, $\iota\in\Lambda_0$. *Die Folge* $(u_\iota)\in\prod H_\iota$ *konvergiert diskret schwach gegen* $u\in H$, $u_\iota \rightharpoonup u\,(\iota\in\Lambda_0)$, *wenn*

$$\bigwedge_{(v_\iota)\in\prod H_\iota} v_\iota \to v\,(\iota\in\Lambda_0) \succ (u_\iota, v_\iota) \to (u, v)\,(\iota\in\Lambda_0).$$

3. Approximationsreguläre Operatoren

Zunächst definieren wir die Paare von Operatoren $(A(z), (A_\iota(z)))$, die in diesem Zusammenhang die wesentliche Rolle spielen.

(1) *Es seien* $A(z)\in L(E, F)$ *und* $(A_\iota(z))\in\prod L(E_\iota, F_\iota)$ *für* $z\in\mathbb{K}$. *Dann heißt das Paar* $(A(z), (A_\iota(z)))$ *approximationsregulär (a-regulär) in* $M\subset\mathbb{K}$, *wenn die folgende Bedingung erfüllt ist: Für jede Zahlenfolge* $z_\iota\in M$ *und jede beschränkte Folge* $(u_\iota)\in\prod D(A_\iota)$ *folgt aus*

$$A_\iota(z_\iota)\,u_\iota \to v\in F \quad (\iota\in\Lambda), \qquad z_\iota \to z_0\in M \quad (\iota\in\Lambda),$$

daß eine Teilfolge $\Lambda_1\subset\Lambda$ *und ein Element* $u\in D(A)$ *existieren mit*

$$u_\iota \to u \quad (\iota\in\Lambda_1) \quad \text{und} \quad A(z_0)\,u = v.$$

Wenn $A_\iota(z) := z\,L_\iota - C_\iota$, $\iota\in\Lambda_0$, mit $(L_\iota)\in\prod B(E_\iota, F_\iota)$ und stabil, dann kann in (1) stets $z_\iota = z_0 = z\in M$ gewählt werden.

Einige Beispiele und Kriterien für a-reguläre Paare werden wir im letzten Abschnitt dieser Arbeit erörtern.

Ein öfters zu Hilfe genommener Hilfssatz lautet:

(2) *Sei* $K\subset\mathbb{K}$ *eine kompakte Menge, das Paar* $(A(z), (A_\iota(z)))$ *sei a-regulär in* K. *Wenn* $A(z)$ *für alle* $z\in K$ *injektiv ist, dann erfüllt* $(A_\iota(z))$ *die inverse Stabilitätsbedingung* 2.(5) *gleichmäßig in* K.

Beweis. Wenn die Behauptung falsch wäre, d.h. eine Teilfolge $\Lambda\subset\Lambda_0$ und für $\iota\in\Lambda$ Elemente $u_\iota\in D(A_\iota)$ mit $\|u_\iota\| = 1$ sowie Zahlen $z_\iota\in K$ existierten mit

(3) $$A_\iota(z_\iota)\,u_\iota \to 0 \quad (\iota\in\Lambda),$$

dann folgte aus der Kompaktheit von K, daß es $z_0 \in K$ mit $z_\iota \to z_0$ $(\iota \in \Lambda_1)$ gibt. Mit Hilfe der a-Regularität von $(A(z), (A_\iota(z)))$ könnte aus (3) die Existenz von $u \in D(A)$ mit $\|u\| = 1$ und $A(z_0)\, u = 0$ erschlossen werden. Dies widerspräche jedoch $N(A(z_0)) = \{0\}$.

Die folgende Bedingung is von grundlegender Bedeutung für unsere Theorie, insbesondere für den Satz (4). Sie geht auf Grigorieff [10] zurück.

(F_ι) *Zu jeder kompakten Menge $K \subset \mathbb{K}$ existiert ein Endstück Λ' von Λ_0, so daß für jedes $z \in K$ und alle $\iota \in \Lambda'$ aus $N(A_\iota(z)) = \{0\}$ folgt $R(A_\iota(z)) = F_\iota$.*

Sie ist erfüllt, wenn z.B. $A_\iota(z)$ für $\iota \in \Lambda'$ Fredholm-Operatoren vom Index Null in einem Gebiet $G \in \mathscr{U}(K)$ sind.

Die wichtigsten Eigenschaften a-regulärer Paare im Zusammenhang mit der Lösung inhomogener Aufgaben fassen wir im folgenden Äquivalenzsatz zusammen.

(4) *Sei $K \subset \mathbb{K}$ eine kompakte Menge und $A(z) \in L(E, F)$, $(A_\iota(z)) \in \prod L(E_\iota, F_\iota)$ für alle $z \in K$. Die Folge $(A_\iota(z))$ erfülle die Bedingung (F_ι), ferner gelte für alle $z_0, z_\iota \in K$, $\iota \in \Lambda$, mit $z_\iota \to z_0 (\iota \in \Lambda)$:*

(5) $$(A_\iota(z_\iota) - A_\iota(z_0)) \to 0 \qquad (\iota \in \Lambda).$$

Dann sind die beiden folgenden Aussagen äquivalent: (i) *Für alle $z \in K$ ist $A(z)$ injektiv und das Paar $(A(z), (A_\iota(z)))$ ist in K a-regulär.* (ii) *Für alle $z \in K$ ist $A(z)$ surjektiv, die Folge $(A_\iota(z))$ erfüllt die inverse Stabilitätsbedingung und das Paar $(A(z), (A_\iota(z)))$ ist konsistent in K.*

Beweis. (i)$\succ$(ii). Aus (2) und (F_ι) entnehmen wir $K \subset B^-(T_\iota)$. Daraus erschließt man hernach die Beschränktheit der Folge $(u_\iota) \in \prod D(A_\iota)$ mit $A_\iota(z_\iota)\, u_\iota = v_\iota$, $\iota \in \Lambda'$. Nun erbringt die a-Regularität von $(A(z), (A_\iota(z)))$ in K die Existenz von $u \in D(A)$ mit $A(z_0)\, u = v$. Wenn $u_\iota \to u (\iota \in \Lambda_1')$ falsch wäre, gäbe es $w \in D(A)$ mit $u_\iota \to w (\iota \in \Lambda_2')$, $A(z_0)\, w = v$ und $w \neq u$, dies folgt ebenfalls aus der a-Regularität von $(A(z), (A_\iota(z)))$. Damit läge wegen $A(z_0) u = v$ ein Widerspruch zur Injektivität von $A(z_0)$ vor. Also haben wir für alle Folgen $z_\iota \in K$, $\iota \in \Lambda$, mit $z_\iota \to z_0$ $(\iota \in \Lambda)$:

$$A_\iota^{-1}(z_\iota) \to A^{-1}(z_0) \qquad (\iota \in \Lambda_1').$$

Mit 2.(7) folgt $A^{-1}(z_0) \in B(F, E)$, also die Surjektivität von $A(z)$ in K. Die Konsistenz haben wir schon mitbewiesen.

(ii)$\succ$(i). Die Konsistenz des Paares $(A(z), (A_\iota(z)))$ in K bedeutet die Existenz von $(u_\iota) \in \prod D(A_\iota)$ zu jedem $u \in D(A)$ und von $z_\iota \in K$, $\iota \in \Lambda$, zu jedem $z_0 \in K$ mit $u_\iota \to u$, $z_\iota \to z_0$, $A_\iota(z_\iota)\, u_\iota \to A(z_0)\, u (\iota \in \Lambda)$. Zusammen mit der inversen Stabilitätsbedingung haben wir

(6) $$\gamma \|u_\iota\| \leqq \|A_\iota(z_\iota)\, u_\iota\|, \qquad \iota \in \Lambda_0',$$

der Grenzübergang erbringt die Injektivität von $A(z_0)$ für jedes $z_0 \in K$. Zum Nachweis der a-Regularität des Paares $(A(z), (A_\iota(z)))$ wählen wir eine beschränkte Folge $(w_\iota) \in \prod D(A_\iota)$ und Zahlen $y_\iota \in K$, für die $y_\iota \to z_0$, $v_\iota := A_\iota(y_\iota)\, w_\iota \to v \in F$ $(\iota \in \Lambda)$ konvergiere. Da $A(z_0)$ bijektiv ist, gibt es $u \in D(A)$ mit $A(z_0)\, u = v$. Zu zeigen bleibt $w_\iota \to u(\iota\in\Lambda_1)$. Wie in (6) haben wir mit Konsistenz- und inverser Stabilitätsbedingung

$$\begin{aligned}\gamma \|w_\iota - u_\iota\| &\leqq \|A_\iota(y_\iota)(w_\iota - u_\iota)\| \\ &\leqq \|A_\iota(y_\iota)\, w_\iota - A_\iota(z_\iota)\, u_\iota\| + \|(A_\iota(z_\iota) - A_\iota(y_\iota))\, u_\iota\| \\ &\leqq \|A_\iota(y_\iota)\, w_\iota - A_\iota(z_\iota)\, u_\iota\| + \|(A_\iota(z_\iota) - A_\iota(z_0))\, u_\iota\| \\ &\quad + \|(A_\iota(z_0) - A_\iota(y_\iota))\, u_\iota\| \to 0 \qquad (\iota\in\Lambda_1').\end{aligned}$$

Dabei haben wir (5) benützt.

Bemerkung. Für den Schluß von (ii) nach (i) benötigt man (F_ι) nicht, für den von (i) nach (ii) ist (5) nicht erforderlich. Wenn $A_\iota(z) := z L_\iota - C_\iota$, $\iota\in\Lambda$, mit $(L_\iota)\in B(E_\iota, F_\iota)$ und stabil, dann ist (5) trivialerweise erfüllt. Verzichten kann man auf diese Bedingung, wenn man in (4). (ii) noch fordert: Für alle $z_\iota\in K$, $\iota\in\Lambda$, mit $z_\iota\to z_0\,(\iota\in\Lambda)$ gelte $A_\iota^{-1}(z_\iota)\to A^{-1}(z_0)\,(\iota\in\Lambda')$. Dann beweist man $w_\iota\to u\,(\iota\in\Lambda_1)$ am Ende des vorstehenden Beweises durch

$$w_\iota = A_\iota^{-1}(z_\iota)\, v_\iota \to A^{-1}(z_0)\, v = u \qquad (\iota\in\Lambda_1').$$

Für inhomogene Aufgaben $A(z)\,u = v$ haben wir in Satz (4) einen konstruktiven Existenzsatz mit Hilfe der approximierenden Probleme $A_\iota(z_\iota)\, u_\iota = v_\iota$, $\iota\in\Lambda_0$, gewonnen. Gleichzeitig wurde die Lösbarkeit und die Konvergenz der Lösungen der Näherungsaufgaben erschlossen. Die wesentliche Rolle bei solchen Untersuchungen spielen offenbar die in (1) definierten Paare $(A(z), (A_\iota(z)))$.

Folgendes Korollar haben wir in (4) bereits mitbewiesen:

(7) *Sei $K\subset\mathbb{K}$ eine kompakte Menge, die Folge $(A_\iota(z))$ unterliege der Bedingung (F_ι) und* (4). (i) *sei erfüllt. Dann ist K Teilmenge von $\rho(A)\cap B^-(A_\iota)$ und für alle Folgen $z_\iota\in K$ mit $z_\iota\to z_0(\iota\in\Lambda)$ konvergiert $A_\iota^{-1}(z_\iota)\to A^{-1}(z_0)(\iota\in\Lambda')$.*

Auf der Basis von (7) werden wir nachher die Existenz und Konvergenz von Folgen von Resolventenintegralen untersuchen. Dabei brauchen wir gewisse Stetigkeitseigenschaften von $A^{-1}(.)$, $(A_\iota^{-1}(.))$, die sich nicht alleine aus der a-Regularität ergeben. Sie sind erfüllt, wenn eine Resolventengleichung wie 1. (9) gilt. Daher gehen wir zu den Paaren gemäß (B) bzw. (B_ι) über.

(8) *Die Folge $(T_\iota(z))$ erfülle die Bedingung (B_ι), die Menge $B^-(T_\iota)$ sei nichtleer und das Paar $(T(z), (T_\iota(z)))$ sei konsistent in $B^-(T_\iota)$. Dann ist die Bedingung* (5) *in $B^-(T_\iota)$ erfüllt.*

Beweis. Aufgrund der Konsistenz haben wir für alle $u\in D(T)$ und alle $z_0\in B^-(T_\iota)$ eine Folge $(u_\iota)\in\prod D(T_\iota)$ und $z_\iota\in B^-(T_\iota)$ mit $u_\iota\to u$, $z_\iota\to z_0$, $T_\iota(z_\iota)\to T(z_0)\,(\iota\in\Lambda)$. Nun haben wir mit 2. (9)

$$\begin{aligned}\|(T_\iota(z_\iota) - T_\iota(z_0))\, u_\iota\| &= |z_\iota - z_0|\, \|L_\iota u_\iota\| \\ &\leqq |z_\iota - z_0|\, \|L_\iota T_\iota^{-1}(z_\iota)\|\, \|T_\iota(z_\iota)\, u_\iota\| \leqq \alpha\, |z_\iota - z_0|\, \|T_\iota(z_\iota)\, u_\iota\| \to 0 \qquad (\iota\in\Lambda_1').\end{aligned}$$

Zur Anwendung von 2. (9) muß dabei eine kompakte Menge $K\subset B^-(T_\iota)$ mit z_0, $z_\iota\in K$, $\iota\in\Lambda'$, gewählt werden.

(9) *Wenn die Folge $(T_\iota(z))$ die Bedingung (B_ι) erfüllt und das Paar $(T(z),(T_\iota(z)))$ in einer Menge $M\subset\mathbb{K}$ a-regulär und konsistent ist, dann gilt (B) für $T(z)$. Außerdem gibt es zu jeder kompakten Teilmenge $K\subset\rho(T)\cap M$ ein Endstück Λ', so daß für alle $z_\iota\in K$ mit $z_\iota\to z_0$ $(\iota\in\Lambda)$ gilt:*

$$L_\iota T_\iota^{-1}(z_\iota)\to LT^{-1}(z_0)\qquad(\iota\in\Lambda'). \tag{10}$$

Beweis. Falls $M\cap\rho(T)\neq\emptyset$, gibt es wegen (7) ein Endstück Λ', so daß für jede kompakte Teilmenge K von $M\cap\rho(T)$ auch $K\subset B^-(T_\iota)$ gilt. Die Stabilität von $(L_\iota T_\iota^{-1}(z_\iota))$, $z_\iota\in K$, gewinnt man nach Durchsicht des Beweises von 1.(5) mit Hilfe von 2.(9). Die Konvergenz von $z_\iota\to z_0$ $(\iota\in\Lambda)$ folgt aus der Kompaktheit von K. Nun erbringen die aufgrund (7) vorliegende Konvergenz von $T_\iota^{-1}(z_\iota)\to T^{-1}(z_0)$ $(\iota\in\Lambda')$ und die Konsistenz von $(L,(L_\iota))$ mit Satz 2.(7) die Behauptung (10) und damit auch die Beschränktheit von $LT^{-1}(z_0)$ auf F.

Im folgenden Satz setzen wir $\mathbb{K}=\mathbb{C}$ voraus.

(11) *Die Voraussetzungen von (8) seien erfüllt, K sei eine kompakte Teilmenge von $\rho(T)\cap B^-(T_\iota)\neq\emptyset$. Dann existiert zu jeder rektifizierbaren Jordankurve $\Gamma\subset\mathring{K}$ und jeder auf Γ gleichgradig stetigen komplexen Funktionenfolge (f_ι), die gegen eine auf Γ stetige komplexe Funktion f gleichmäßig konvergiert, ein Endstück Λ', so daß die folgenden Resolventenintegrale existieren und konvergieren:*

$$\int_\Gamma f_\iota(z)\,L_\iota T_\iota^{-1}(z)\,dz\to\int_\Gamma f(z)\,LT^{-1}(z)\,dz\qquad(\iota\in\Lambda'). \tag{12}$$

Beweis. Die Kurve Γ (die mehr als nur einen Punkt enthalte) ist kompakte Teilmenge von $B^-(T_\iota)$, also existieren nach 2.(9) ein Endstück Λ' und $\gamma>0$, so daß $\Gamma\subset\rho(T_\iota)$ und $\|T_\iota^{-1}(z)\|\leqq\gamma^{-1}$ für $z\in\Gamma$, wenn $\iota\in\Lambda'$. Weiter haben wir die Existenz von α mit $\|L_\iota T_\iota^{-1}(z)\|\leqq\alpha$ für $z\in\Gamma$ und $\iota\in\Lambda'$. Auf Γ ist die Folge (f_ι) gleichmäßig beschränkt, z.B. durch $|f_\iota(z)|\leqq\Phi$, $\iota\in\Lambda$. Nun können wir abschätzen:

$$\begin{aligned}&\|f_\iota(z_1)\,L_\iota T_\iota^{-1}(z_1)-f_\iota(z_2)\,L_\iota T_\iota^{-1}(z_2)\|\\&\qquad\leqq|f_\iota(z_1)-f_\iota(z_2)|\;\|L_\iota T_\iota^{-1}(z_1)\|+|f_\iota(z_2)|\;\|L_\iota T_\iota^{-1}(z_1)-L_\iota T_\iota^{-1}(z_2)\|\\&\qquad\leqq\alpha\,|f_\iota(z_1)-f_\iota(z_2)|+\Phi\,\alpha^2\,|z_1-z_2|.\end{aligned}$$

Dabei wurden 1.(9) und 2.(9) zu Hilfe genommen. Damit haben wir die gleichmäßige gleichgradige Stetigkeit der Folge $(f_\iota(.)\,L_\iota T_\iota^{-1}(.))$ auf Γ. Dies zieht zunächst einmal die Existenz der Integrale in (12) nach sich. Mit (7) und (10) gewinnt man noch für alle z:

$$f_\iota(z)\,L_\iota T_\iota^{-1}(z)\to f(z)\,LT^{-1}(z)\qquad(\iota\in\Lambda').$$

Nun kann man die Behauptung mit Hilfe eines Lemmas von Stummel [19] erschließen:

(13) *Seien φ,φ_ι stetige Funktionen auf einer rektifizierbaren Jordan-Kurve Γ mit Werten in Banachschen Räumen Y, Y_ι für $\iota\in\Lambda_0$. Ferner sei die Folge (φ_ι)*

gleichmäßig gleichgradig stetig auf Γ, *für jedes* $z \in \Gamma$ *konvergiere* $\varphi_\iota(z) \to \varphi(z)$ $(\iota \in \Lambda)$ *und die Integrale* $\int_\Gamma \varphi(z)\,dz$, $\int_\Gamma \varphi_\iota(z)\,dz$ *existieren. Dann konvergieren sie auch:* $\int_\Gamma \varphi_\iota(z)\,dz \to \int_\Gamma \varphi(z)\,dz$ $(\iota \in \Lambda)$.

4. Approximation der Eigenwerte und -vektoren

In diesem Abschnitt wird $\mathbb{K} = \mathbb{C}$ gefordert. Die allgemein benötigten Voraussetzungen fassen wir in der folgenden Bedingung (V) zusammen. Dabei müssen wir uns, da wir gewisse Regularitätseigenschaften der Resolventenfunktionen brauchen, von Anfang an auf (B) bzw. (B_ι) stützen.

(V) *Die Folge* $((L_\iota, C_\iota))$ *erfülle die Bedingungen* (B_ι) *und* (F_ι). *In einer kompakten Menge* $K \subset \mathbb{C}$ *sei das Paar* $(T(z), (T_\iota(z)))$ *a-regulär und konsistent. Schließlich seien* $K \cap \rho(T)$ *und* $K \cap B^-(T_\iota)$ *nichtleere Mengen.*

Zunächst beweisen wir die Konvergenz von Spektralmengen.

(1) *Sei* (V) *erfüllt und in* $\mathring{K}$ *eine Spektralmenge* σ *von* T *enthalten. Dann existiert zu jedem Cauchy-Gebiet* Δ *zu* σ *mit* $\bar{\Delta} \subset \mathring{K}$ *ein Endstück* $\Lambda'(\Delta)$, *so daß für alle* $\iota \in \Lambda'(\Delta)$ *die Umgebung* Δ *auch Cauchy-Gebiet zu einer nichttrivialen Spektralmenge von* T_ι *und* $K \setminus \Delta$ *in* $\rho(T_\iota)$ *enthalten ist.*

Beweis. Sei Δ_1 Cauchy-Gebiet zu σ mit $\bar{\Delta}_1 \subset \Delta$. Dann ist $K \setminus \Delta_1$ eine kompakte Teilmenge von $\rho(T)$. Aus 3.(7) folgt die Existenz eines Endstücks Λ' mit $K \setminus \Delta \subset K \setminus \Delta_1 \subset \rho(T_\iota)$ für $\iota \in \Lambda'$. Nun nehmen wir an, der erste Teil der Behauptung sei falsch. Dann existierte Λ_1, so daß $\Delta \subset \rho(T_\iota)$, d.h. $T_\iota^{-1}(.)$ wäre für $\iota \in \Lambda_1$ holomorph in Δ. Insbesondere läge die Holomorphie von $T_\iota^{-1}(.)$ auch in $\bar{\Delta}_1 \subset \Delta$ vor. Das Maximumprinzip [5] würde erbringen, daß die aufgrund 2.(10) für $\iota \in \Lambda_1'$ in $K \setminus \Delta_1$ geltende Abschätzung $\|T_\iota^{-1}(z)\| \leqq \gamma^{-1}$ auch in Δ_1 gälte. Zusammen mit der Konsistenz ergäbe dies für ein beliebiges $z_0 \in \sigma$

$$\gamma \|u_\iota\| \leqq \|T_\iota(z_0)\, u_\iota\|, \qquad \iota \in \Lambda_1',$$

nach dem Grenzübergang hätte man die Injektivität von $T(z_0)$. Vorausgesetzt war jedoch $z_0 \in \sigma$.

Von nun an beschäftigen wir uns speziell mit Operatoren und Folgen, die nicht nur (V) erfüllen, sondern außerdem für alle z aus einem Gebiet $G \in \mathscr{U}(K)$ sämtlich Semifredholm-Operatoren aus $\Phi^l(E, F)$ bzw. $\Phi^l(E_\iota, F_\iota)$, $\iota \in \Lambda_0$, sind. Nach 1.(24) enthält K in diesem Falle höchstens endlich viele Eigenwerte von T bzw. T_ι, die alle Pole der Resolventenfunktionen sind.

Unter diesen Voraussetzungen erhalten wir die folgenden beiden Konvergenzsätze.

(2) *Wenn* $z_\iota \in \mathring{K}$ *Eigenwerte von* T_ι *mit* $z_\iota \to z_0$ $(\iota \in \Lambda)$ *und* $u_\iota \in D(T_\iota)$ *mit* $\|u_\iota\| = 1$, $\iota \in \Lambda$, *zugehörige Eigenvektoren sind, dann ist* z_0 *Eigenwert von* T *und es existieren* $\Lambda_1 \subset \Lambda$ *und* $u_0 \in D(T)$ *mit*

$$u_\iota \to u_0 \quad (\iota \in \Lambda_1), \quad \|u_0\| = 1, \quad T(z_0)\, u_0 = 0.$$

Diese Behauptung ergibt sich unmittelbar aus der a-Regularität von $(T(z), (T_\iota(z)))$. Umgekehrt haben wir den Satz:

(3) *Wenn $z_0 \in \mathring{K}$ ein Eigenwert von T ist, dann gibt es eine Teilfolge Λ_1 und Eigenwerte $z_\iota \in \mathring{K}$ von T_ι, $\iota \in \Lambda$, mit $z_\iota \to z_0$ $(\iota \in \Lambda)$. Ferner gilt*

(4)
$$\overline{\lim_{\iota \in \Lambda}} \dim N(T_\iota(z_\iota)) \leqq \dim N(T(z_0)).$$

Beweis. Der erste Teil ist eine Folge von Satz (1). Die Behauptung (4) beweisen wir indirekt [10]. Es existiere Λ mit

(5)
$$d := \dim N(T(z_0)) < \dim N(T_\iota(z_\iota)).$$

Die z_ι sind Eigenwerte mit zugehörigen Elementen $u_\iota^{(1)} \in N(T_\iota(z_\iota))$, $\|u_\iota^{(1)}\| = 1$, $\iota \in \Lambda$. Aufgrund der a-Regularität von $(T(z), (T_\iota(z)))$ existieren $\Lambda_1 \subset \Lambda$ und $u^{(1)} \in N(T(z_0))$ mit $u_\iota^{(1)} \to u^{(1)}$ $(\iota \in \Lambda_1)$ und $\|u^{(1)}\| = 1$. Wenn $d > 0$, dann gibt es $u_\iota^{(2)} \in N(T_\iota(z_\iota))$, $\iota \in \Lambda_1$, mit

$$u_\iota^{(2)} \to u^{(2)} \quad (\iota \in \Lambda_2), \qquad \|u^{(2)}\| = 1, \qquad \|u^{(2)} - \operatorname{span}(u^{(1)})\| \geqq \tfrac{1}{2}.$$

Fortsetzung dieses Konstruktionsverfahrens führt zu Elementen $u^{(j)} \in N(T(z_0))$, $j = 1, \ldots, d+1$, die folgende Bedingung erfüllen:

$$\|u^{(j)}\| = 1, \qquad \|u^{(j)} - \operatorname{span}(u^{(1)}, \ldots, u^{(j-1)})\| \geqq \tfrac{1}{2}, \quad j = 1, \ldots, d+1.$$

Nach dem Rieszschen Lemma [21] sind $z^{(1)}, \ldots, u^{(d+1)}$ linear unabhängig. Die Annahme (5) führt also auf einen Widerspruch.

Wir gehen nun über zu der Frage, ob die in (3) bewiesene Konvergenz der Eigenwerte der Vielfachheit nach erfolgt.

(6) *Sei $z_0 \in \mathring{K}$ Eigenwert von T und Δ mit $\bar{\Delta} \subset \mathring{K}$ ein Cauchy-Gebiet zu $\{z_0\}$. Dann existiert ein Endstück Λ', so daß gilt:*

(7)
$$m := \dim M(z_0, T) \leqq \underline{\lim}_{\iota \in \Lambda'} \dim M(\sigma_\iota, T_\iota).$$

Beweis. Nach Satz (1) existiert ein Endstück Λ', so daß Δ für alle $\iota \in \Lambda'$ eine nichttriviale Spektralmenge σ_ι von T_ι enthält, die aufgrund 1.(24) aus endlich vielen Polen von $T_\iota^{-1}(.)$ besteht. Die Existenz und Konvergenz der Projektionen $F(\sigma_\iota, T_\iota) \to F(z_0, T)$ $(\iota \in \Lambda')$ erhalten wir, wenn wir ein Cauchy-Gebiet Δ_1 zu $\{z_0\}$ wählen mit $\bar{\Delta} \subset \Delta_1$, $\bar{\Delta}_1 \subset K$. Dann haben wir mit 3.(7), daß $K \setminus \Delta \subset \rho(T) \cap B^-(T_\iota)$. Nun wende man Satz 3.(11) mit $\Gamma := \partial \Delta_1$ an.

Aufgrund von 1.(23) können wir nun eine Basis $w_1, \ldots, w_m$ von $R(F(z_0, T))$ auswählen. Nach 2.(1) existieren dann Folgen $(R_\iota w_j) \in \prod F_\iota$ mit $R_\iota w_j \to w_j$ $(\iota \in \Lambda)$ und $F(\sigma_\iota, T_\iota) R_\iota w_j \in M(\sigma_\iota, T_\iota)$ für $\iota \in \Lambda'$ und $j = 1, \ldots, m$. Die Elemente $F(\sigma_\iota, T_\iota) R_\iota w_j$, $j = 1, \ldots, m$, sind für alle $\iota \in \Lambda''$ linear unabhängig. Wäre dies nicht wahr, so existierten Λ_1'' und Zahlen c_j^ι, $j = 1, \ldots, m$, mit

$$\sum_{j=1}^{m} (c_j^\iota)^2 \leqq 1, \qquad \sum_{j=1}^{m} c_j^\iota F(\sigma_\iota, T_\iota) R_\iota w_j = 0, \quad \iota \in \Lambda_1''.$$

Weiter gäbe es eine Teilfolge Λ_2'' und c_j, $j=1,\ldots,m$, mit

$$\sum_{j=1}^{m}(c_j)^2 \leqq 1, \quad c_j^\iota \to c_j \quad (\iota\in\Lambda_2''),$$

$$0=\sum_{j=1}^{m} c_j^\iota F(\sigma_\iota, T_\iota) R_\iota w_j \to \sum_{j=1}^{m} c_j F(z_0, T) w_j \quad (\iota\in\Lambda_2'').$$

Da $w_j\in R(F(z_0,T))$ haben wir $w_j=F(z_0,T)w_j$, $j=1,\ldots,m$, also $\sum_{j=1}^{m} c_j w_j=0$. Dies widerspräche der Voraussetzung, daß $w_1,\ldots,w_m$ Basis von $R(F(z_0,T))$ ist.

In einem wichtigen Spezialfall haben wir das folgende weitergehende Resultat:

(8) *Wenn in* (6) *die Folge* $(F(\sigma_\iota, T_\iota))$ *diskret kompakt ist, dann existiert* Λ' *derart, daß für alle* $\iota\in\Lambda'$ *gilt:* $\dim M(\sigma_\iota, T_\iota)=m$.

Beweis. Existierte eine Teilfolge Λ mit $\dim M(\sigma_\iota, T_\iota)>m$ für $\iota\in\Lambda$, so hätten wir eine Folge $(v_\iota)\in\prod R(F(\sigma_\iota, T_\iota))$ mit

(9) $\|v_\iota\|=1, \quad \|v_\iota-\operatorname{span}(F(\sigma_\iota, T_\iota)R_\iota w_1,\ldots,F(\sigma_\iota, T_\iota)R_\iota w_m)\|\geqq\frac{1}{2}, \quad \iota\in\Lambda.$

Weil $v_\iota=F(\sigma_\iota, T_\iota)v_\iota$, $\iota\in\Delta$, und $(F(\sigma_\iota, T_\iota))$ diskret kompakt ist, erhielten wir durch Grenzübergang die Existenz einer Teilfolge Λ_1 und eines Elements w mit $v_\iota=F(\sigma_\iota, T_\iota)v_\iota\to w(\iota\in\Lambda_1)$. Aufgrund der Konvergenz von $F(\sigma_\iota, T_\iota)\to F(z_0,T)$ $(\iota\in\Lambda)$ wäre also $w=F(z_0,T)w$, d.h. $w\in R(F(z_0,T))$. Wegen (9) folgte im Grenzübergang, daß

$$\|w\|=1, \quad \|w-\operatorname{span}(w_1,\ldots,w_m)\|\geqq\tfrac{1}{2}.$$

Das widerspräche der Tatsache, daß $w_1,\ldots,w_m$ Basis von $R(F(z_0,T))$ ist.

Daran anschließend erhalten wir eine Aussage über die Konvergenz von Folgen von Hauptvektoren.

(10) *Seien* $z_\iota\in\mathring{K}$ *Eigenwerte von* T_ι *mit* $z_\iota\to z_0$ $(\iota\in\Lambda)$. *Jede Folge* (u_ι) *von Hauptvektoren* u_ι *von* T_ι *zu* z_ι, $\iota\in\Lambda$, *besitzt eine gegen einen Hauptvektor* u *von* T *zu* z_0 *konvergente Teilfolge genau dann, wenn die Folge* $(F(\sigma_\iota, T_\iota))$ *diskret kompakt ist.*

Beweis. Im folgenden ist $y\in\mathring{K}\setminus\Delta_1$ zu wählen. Wenn $(F(\sigma_\iota, T_\iota))$ diskret kompakt ist, dann existieren Λ_1 und $v\in F$ mit

$$v_\iota=F(\sigma_\iota, T_\iota)v_\iota\to v \quad (\iota\in\Lambda_1),$$

falls $(v_\iota)\in\prod M(\varphi(z_\iota), V_\iota)$. Aufgrund der Konvergenz von $F(\sigma_\iota, T_\iota)\to F(z_0,T)$ $(\iota\in\Lambda)$ haben wir dann $v\in M(\varphi(z_0), V)$. Die Hinlänglichkeit folgt nun aus 1.(21) und der Konvergenz von $T_\iota^{-1}(y)\to T^{-1}(y)$ $(\iota\in\Lambda_1)$.

Zum Beweis der Notwendigkeit wählen wir $(v_\iota)\in\prod M(\varphi(z_\iota), V_\iota)$. Aufgrund 1.(21) haben wir dann $u_\iota=T_\iota^{-1}(y)v_\iota\in M(z_\iota, T_\iota)$ für $\iota\in\Lambda$, es existieren also Λ_1 und $u\in M(z_0,T)$ mit

$$u_\iota=T_\iota^{-1}(y)v_\iota\to u \quad (\iota\in\Lambda_1).$$

Hieraus erschließt man sofort die Behauptung.

Eine über (4) hinausgehende, mit (6) oder sogar (8) vergleichbare Aussage über die Konvergenz der geometrischen Vielfachheit nach besteht im allgemeinen nicht. Michlin zeigte nämlich [8], daß die geometrische Vielfachheit von σ_ι nicht unterhalb-, Grigorieff [9], daß sie nicht oberhalbstetig ist.

Eine letzte Bemerkung gilt dem Verhalten des Anstiegs gegen die Approximation.

(11) *Unter den Voraussetzungen von* (6) *und* (8) *gilt*

$$\varliminf_{\iota\in\Lambda'} \sum_{z_\iota\in\sigma_\iota} \alpha(z_\iota) \geqq \alpha(z_0).$$

Beweis. Jeder Eigenwert von V bzw. V_ι ist ein Pol der entsprechenden Resolventenfunktion mit dem Anstieg als Ordnung. Für alle $\iota\in\Lambda'$ ist dann [5]

$$\frac{1}{2\pi i}\int_C \prod_{z_\iota\in\sigma_\iota} (z-z_\iota)^{\alpha(z_\iota)} L_\iota T_\iota^{-1}(z)\,dz=0, \tag{12}$$

denn der Integrand ist für jedes $\iota\in\Lambda'$ holomorph in $\bar{\Delta}$. Aus (8) erschließt man weiter die Existenz von $q\in\mathbb{N}$ und Λ_1', so daß $q=\sum_{z_\iota\in\sigma_\iota}\alpha(z_\iota)$, $\iota\in\Lambda_1'$. Nun ziehen (12) und 3.(12)

$$\frac{1}{2\pi i}\int_C (z-z_0)^q\, LT^{-1}(z)\,dz=0$$

und damit $q\geqq\alpha(z_0)$ nach sich, denn z_0 ist Pol der Ordnung $\alpha(z_0)$ von $T^{-1}(.)$.

5. Kriterien und Beispiele

Zunächst erinnern wir daran, daß $A(z)$, $B(z)$, ... stets durch Paare linearer Operatoren im Sinne von 1.(1) definiert sind. Man bemerkt nun sofort:

(1) *Wenn* $(E,(E_\iota))$ *eine diskrete Approximation ist, dann haben wir die a-Regularität von* $(z\,I,(z\,I_\iota))$.

Dies folgt unmittelbar aus 2.(1) und 2.(2).

Fast ebenso leicht ist einzusehen:

(2) *Für alle* $z\in M\subset\mathbb{K}$ *seien* $B(z)\in B(E,F)$, $A(z)\in B(F,G)$, $(B_\iota(z))\in\prod B(E_\iota,F_\iota)$, $(A_\iota(z))\in\prod B(F_\iota,G_\iota)$ *und* $(B_\iota(z))$ *stabil. Wenn* $(A(z),(A_\iota(z)))$ *und* $(B(z),(B_\iota(z)))$ *a-reguläre Paare in* M *sind, dann besitzt auch* $(A(z)\,B(z),(A_\iota(z)\,B_\iota(z)))$ *diese Eigenschaft.*

Beweis. Für eine beschränkte Folge $(v_\iota)\in\prod E_\iota$ und $z_\iota\in M$, $\iota\in\Lambda$, gelte

$$z_\iota\to z_0\in M\quad(\iota\in\Lambda),\qquad A_\iota(z_\iota)\,B_\iota(z_\iota)\,v_\iota\to w\in G\quad(\iota\in\Lambda).$$

Da auch $(B_\iota(z_\iota)\,v_\iota)$ beschränkt ist, existieren nach 3.(1) eine Teilfolge $\Lambda_1\subset\Lambda$ und $u\in F$ mit $B_\iota(z_\iota)\,v_\iota\to u$ $(\iota\in\Lambda_1)$ und $A(z_0)\,u=w$. Wiederholte Anwendung von 3.(1) erbringt die Existenz von $\Lambda_2\subset\Lambda_1$ und $v\in E$ mit $v_\iota\to v$ $(\iota\in\Lambda_2)$ und $B(z_0)\,v=u$, d.h., $A(z_0)\,B(z_0)\,v=w$.

Die a-Regularität eines Paares ist eine gegen gewisse Störungen stabile Eigenschaft. Wir zeigen dafür zwei Beispiele, für den Fall parameterunabhängiger Operatoren wurden sie in [10] angeführt.

(3) *Das Paar* $(A(z), (A_\iota(z)))$ *sei a-regulär in* $M \subset \mathbb{K}$, *die Folge* $(B_\iota(z_\iota)) \in \prod B(E_\iota, F_\iota)$ *sei für* $z_\iota \in M$, $\iota \in \Lambda$, *diskret kompakt, ferner gelte* $B_\iota(z_\iota) \to B(z_0)$ $(\iota \in \Lambda)$ *für* $z_\iota \in M$ *mit* $z_\iota \to z_0 \in M$ $(\iota \in \Lambda)$. *Dann ist auch* $(A(z)+B(z), (A_\iota(z)+B_\iota(z)))$ *a-regulär in* M.

Beweis. Für die beschränkte Folge $(u_\iota) \in \prod D(A_\iota)$ und für $z_\iota \in M$, $\iota \in \Lambda$, gelte $z_\iota \to z_0 \in M$ $(\iota \in \Lambda)$ und $(A_\iota(z_\iota)+B_\iota(z_\iota))\, u_\iota \to v$ $(\iota \in \Lambda)$. Weil $(B_\iota(z_\iota))$ diskret kompakt ist, existieren $\Lambda_1 \subset \Lambda$ und $w \in F$ mit $A_\iota(z_\iota)\, u_\iota \to w+v$ $(\iota \in \Lambda_1)$. Nun nützt man die a-Regularität von $(A(z), (A_\iota(z)))$ aus: Es existieren demnach $\Lambda_2 \subset \Lambda_1$ und $u \in D(A)$ mit $u_\iota \to u$ $(\iota \in \Lambda_2)$ und $A(z_0)\, u = w+v$. Die Konvergenz von $B_\iota(z_\iota) \to B(z_0)$ $(\iota \in \Lambda)$ erbringt nun die Behauptung.

Die geschilderte Situation tritt z.B. auf, wenn kompakte Integraloperatoren mit Hilfe einer konvergenten Folge von Quadraturformeln diskretisiert werden. In den Arbeiten [1, 2, 3] wurde stets $E = E_\iota = C(\Omega)$, $\iota \in \Lambda$, gewählt, wobei $C(\Omega)$ den Raum der stetigen Funktionen auf einer kompakten Mannigfaltigkeit Ω im $\mathbb{R}^n$ bedeutet. Weiter ist zu setzen $(A(z), (A_\iota(z))) = (z\, I, (z\, I_\iota))$ und mit einer endlichen Teilmenge $\Omega_\iota \subset \Omega$:

$$(B(z)\, u)(x) := \int_\Omega K(x, y)\, u(y)\, dy, \quad x \in \Omega;$$

$$(B_\iota(z)\, u)(x) := \sum_{y \in \Omega_\iota} K(x, y)\, u(y)\, \mu_\iota(y), \quad x \in \Omega, \ \iota \in \Lambda.$$

Wegen der Lösung von Eigenwertaufgaben für Integraloperatoren in L^p-Räumen verweisen wir auf [11].

(4) *Wenn* $A(z) \in B(E, F)$ *und die stabile Folge* $(A_\iota(z)) \in \prod B(E_\iota, F_\iota)$ *die Voraussetzungen und die Bedingung* (i) *von Satz* 3.(4) *in* $K \subset \mathbb{K}$ *erfüllen, ferner für* $B(z) \in B(E, F)$ *und* $(B_\iota(z)) \in \prod B(E_\iota, F_\iota)$ *gilt* $B_\iota(z_\iota) \to B(z_0)$ $(\iota \in \Lambda)$ *für* $z_\iota \in K$ *mit* $z_\iota \to z_0 \in K$ $(\iota \in \Lambda)$ *und* $\sup_{\iota \in \Lambda''} \|B_\iota(z)\| < (\sup_{\iota \in \Lambda'} \|A_\iota^{-1}(z)\|)^{-1}$, *dann haben wir:* $A(z)+B(z)$ *ist injektiv für* $z \in K$ *und* $(A(z)+B(z), (A_\iota(z)+B_\iota(z)))$ *ist konsistent und a-regulär in* K.

Beweis. Die Konsistenz von $(A(z)+B(z), (A_\iota(z)+B_\iota(z)))$ erschließt man aus der Konvergenz $A_\iota(z_\iota) \to A(z)$ $(\iota \in \Lambda)$, die ihrerseits Konsequenz von 2.(7) und 3.(4) ist. Daß $\Lambda' \subset \Lambda$ und $\gamma > 0$ mit $\sup_{\iota \in \Lambda'} \|A_\iota^{-1}(z)\| = \gamma^{-1}$ für alle $z \in K$ existieren, entnimmt man 3.(4). Nun schätzen wir mit Hilfe von $\sup_{\iota \in \Lambda'} \|B_\iota(z)\| < \gamma$ ab:

$$\|(A_\iota(z)+B_\iota(z))\, u\| \geqq \gamma\, \|u\| - \|B_\iota(z)\, u\| \geqq \gamma(1-q)\, \|u\|$$

mit $q < 1$ für $u \in E_\iota$, $\iota \in \Lambda''$, und erhalten damit die inverse Stabilität von $(A_\iota(z)+B_\iota(z))$. Die Bijektivität von $A(z)$ für $z \in K$ und die Konvergenz $A_\iota^{-1}(z_\iota) \to A^{-1}(z_0)$ $(\iota \in \Lambda')$ für $z_\iota \in K$ mit $z_\iota \to z_0$ $(\iota \in \Lambda)$ erbringt wieder 3.(4). Damit führen wir zunächst in

$$\|A_\iota^{-1}(z)\, B_\iota(z)\| \leqq p < \gamma^{-1}\, \gamma = 1 \tag{5}$$

den Grenzübergang mit dem Resultat $\|A^{-1}(z)\,B(z)\|<1$ durch. Das Lemma von Banach [21] sichert nun die Bijektivität von $I+A^{-1}(z)\,B(z)$ und damit die von $A(z)+B(z)=A(z)\big(I+A^{-1}(z)\,B(z)\big)$, die Behauptung folgt, weil 3.(4).(ii) hinreichend für 3.(4).(i) ist.

Man kann jetzt (1), (3) und (4) zusammenfassen, wir tun dies in einem Spezialfall:

(6) *Seien* $E=F$ *und* $E_\iota=F_\iota$, $\iota\in\Lambda$, *ferner sei* $T(z):=zI-S-V$ *und* $T_\iota(z):=zI_\iota-S_\iota-V_\iota$, $\iota\in\Lambda$ *und* $z\in\mathbb{K}$, *mit* $\|S_\iota\|\leqq q$, $\iota\in\Lambda$, $S_\iota\to S\,(\iota\in\Lambda)$ *und* $(V_\iota)\in\prod B(E_\iota,E_\iota)$ *diskret kompakt*, $V_\iota\to V\in B(E,E)$ $(\iota\in\Lambda)$. *Dann ist* $\big(zI-S-V,(zI_\iota-S_\iota-V_\iota)\big)$ *für* $|z|>q$ *a-regulär und die Folge* $\big(F(\sigma_\iota,T_\iota)\big)$ *diskret kompakt.*

Beweis. Setze in (4) für $\big(B(z),(B_\iota(z))\big)$ das Paar $\big(-S,(-S_\iota)\big)$ und für $\big(A(z),(A_\iota(z))\big)$ das Paar $\big(zI,(zI_\iota)\big)$, anschließend in (3) für $\big(A(z),(A_\iota(z))\big)$ das Paar $\big(zI-S,(zI_\iota-S_\iota)\big)$ und für $\big(B(z),(B_\iota(z))\big)$ das Paar $\big(-V,(-V_\iota)\big)$ ein. Die diskrete Kompaktheit von (V_ι) zieht die von $\big(F(\sigma_\iota,T_\iota)\big)$ nach sich.

Hier kann man die Resultate von Bruhn-Wendland [4] und die aus [13] unserer Theorie unterordnen. Zu setzen ist $E=F$ und für $\iota\in\Lambda$: $S_\iota:=P_\iota SP_\iota$ sowie $V_\iota:=P_\iota VP_\iota$ bzw. (V_ι) diskret kompakt und V_ι, V vollstetig mit $V_\iota\to V\,(\iota\in\Lambda)$. (P_ι) bedeutet dabei eine gegen die Identität konvergente Folge von Projektionsoperatoren. An dieser Stelle ist auch bemerkenswert, daß bei der in (6) geschilderten Situation unter Verwendung der Projektionsmethode unsere a-regulären Operatoren und die A-proper mappings von Petryshyn [16] zusammenfallen.

Stummel hat in [19] folgende Paare gründlich studiert: $T(z):=A-zB$ mit $A\in B(E,F)$, B vollstetig und $T_\iota(z):=A_\iota-zB_\iota$ mit $A_\iota\in B(E_\iota,F_\iota)$, B_ι vollstetig und (B_ι) diskret kompakt, $\iota\in\Lambda$. Er gibt für diese Klasse den folgenden Satz:

(7) *Sei* $\rho(T)\neq\emptyset$ *und* $B^-(T_\iota)\neq\emptyset$. *Dann ist die Konsistenz der Paare* $\big(A,(A_\iota)\big)$ *und* $\big(B,(B_\iota)\big)$ *hinreichend und für eine stabile Folge* (A_ι) *auch notwendig für* $B^-(T_\iota)=\rho(T)$ *und*

$$T_\iota^{-1}(z)\to T^{-1}(z)\qquad(\iota\in\Lambda),\ z\in\rho(T).$$

Wegen 3.(7) ist dann das oben definierte Paar $\big(T(z),(T_\iota(z))\big)$ a-regulär in jeder kompakten Teilmenge von $\rho(T)$. Da ferner nur Eigenwerte ohne Häufungspunkt im Endlichen auftreten, ist dies Paar in jeder kompakten Teilmenge von $\mathbb{K}$ a-regulär. In der Arbeit [19] werden zahlreiche Arbeiten zusammengestellt, deren Resultate sich den für die dort untersuchten speziellen a-regulären Paare geltenden unterordnen lassen, z.B. Anselone-Palmer [1], Atkinson [2], Brakhage [3], Grigorieff [8]. Sämtliche in diesen Arbeiten untersuchten Konstellationen von beschränkten Operatoren und zugehörigen Näherungsverfahren ordnen sich der von uns behandelten Situation unter.

Wir wenden uns nun noch unbeschränkten Operatoren zu.

(8) *Sei* $(B_\iota)\in\prod B(E_\iota,F_\iota)$, $(C_\iota)\in\prod C(E_\iota,F_\iota)$. *Die Folge* (C_ι) *erfülle die inverse Stabilitätsbedingung und die Folge* $(C_\iota^{-1})\in\prod B(F_\iota,E_\iota)$ *sei diskret kompakt. Ferner sei das Paar* $\big(C,(C_\iota)\big)$ *konsistent und besitze die Eigenschaft*

(9) $$\bigwedge_{u\in E}\ \bigwedge_{(u_\iota)\in\prod D(C_\iota)} u_\iota\to u\,(\iota\in\Lambda),\ Cu_\iota\to v\in F\,(\iota\in\Lambda)\succ u\in D(C),\ Cu=v.$$

Schließlich sei (B_ι) *stabil und das Paar* $(B,(B_\iota))$ *konsistent. Dann ist das Paar* $(zB-C,(zB_\iota-C_\iota))$ *a-regulär in jeder kompakten Teilmenge* $K\subset\mathbb{K}$.

Beweis. Zunächst erbringt die Konsistenz von $(C,(C_\iota))$ zusammen mit der inversen Stabilität von (C_ι) die Injektivität von C. Nun wählen wir eine beschränkte Folge $(u_\iota)\in\prod D(C_\iota)$ mit $C_\iota u_\iota\to v\in F$ $(\iota\in\Lambda)$. Die diskrete Kompaktheit von (C_ι^{-1}) hat zur Folge, daß $\Lambda_1\subset\Lambda$ und $u\in E$ existieren mit $u_\iota=C_\iota^{-1}v_\iota\to u$ $(\iota\in\Lambda')$. Mit (9) haben wir nun die a-Regularität von $(C,(C_\iota))$. Da (C_ι) die Bedingung (F_ι) erfüllt, folgt aus 3.(4) die Surjektivität von C und $C_\iota^{-1}\to C^{-1}$ $(\iota\in\Lambda')$. Nun können wir das Paar $((zBC^{-1}-I_F)C,((zB_\iota C_\iota^{-1}-I_{F_\iota})C_\iota))$ betrachten. Offensichtlich ist $(B_\iota C_\iota^{-1})$ diskret kompakt und es gilt $B_\iota C_\iota^{-1}\to BC^{-1}$ $(\iota\in\Lambda')$. Beides folgt mit der aus Satz 2.(7) zu schließenden Konvergenz von $B_\iota\to B$ $(\iota\in\Lambda)$. Also erschließt man mit den Sätzen (1) und (3) die a-Regularität von $(zBC^{-1}-I_F,(zB_\iota C_\iota^{-1}-I_{F_\iota}))$. Die Behauptung erhält man endlich mit Satz (2).

Ein Beispiel für die in (8) geschilderte Situation findet man in der Arbeit [15] von Knauer. Dort wird die Eigenwertaufgabe $zI-T$ für einen m-sektoriellen Operator $T\in C(H,H)$ in einem separablen Hilbertschen Raum H mit dem Galerkin-Verfahren gelöst. Man muß, um das Resultat dieser Arbeit hier einzuordnen, $E=F=H$ und $E_\iota=F_\iota=P_\iota H$, $\iota\in\mathbb{N}$, setzen, wobei (P_ι) Projektionen in H mit $P_\iota\to I$ $(\iota\in\Lambda)$ sind. Dann sei weiter $B=I$, $B_\iota=P_\iota$, $\iota\in\Lambda_0$, C sei ein m-sektorieller, abgeschlossener und dicht definierter Operator in H mit vollstetiger Inverser C^{-1} und es seien $C_\iota:=P_\iota CP_\iota$, $\iota\in\mathbb{N}$. Nun zeigt Knauer in [15]: Wenn $Q_\iota:=P_\iota|CE_\iota$: $CE_\iota\to E_\iota$, so existiert Q_ι^{-1} auf E_ι und damit sind C_ι und $T^{-1}Q_\iota^{-1}P_\iota$, $\iota\in\Lambda$, zueinander invers. Die diskrete Kompaktheit von (C_ι^{-1}) folgt nun offensichtlich aus der Vollstetigkeit von C^{-1}. Die Bedingung (9) folgt aus den Eigenschaften von (P_ι). Die Stabilität und Konsistenz von (P_ι) ist wegen der Konvergenz gegeben. Damit sind alle Voraussetzungen von (8) erfüllt.

Bei der Untersuchung von Paaren abgeschlossener Operatoren $(T(z),(T_\iota(z)))$ auf ihre a-Regularität hin ist man häufig auf A-priori-Abschätzungen angewiesen. Dazu setzen wir voraus, daß E, E_ι, $\iota\in\Lambda_0$, Hilbertsche Räume sind und jede beschränkte Folge $(u_\iota)\in\prod E_\iota$ eine schwach konvergente Teilfolge enthält: $u_\iota\rightharpoonup u\in E$ $(\iota\in\Lambda)$. Für den Fall injektiver, parameterunabhängiger Operatoren wurde dies Kriterium von Grigorieff [10] ausgesprochen.

(10) *Sei* $K\subset\mathbb{K}$ *eine kompakte Menge und in einem Gebiet* $G\in\mathscr{U}(K)$ *seien* $T(.)\in\Phi^l(E,F)$, $T_\iota(.)\in\Phi^l(E_\iota,F_\iota)$, $\iota\in\Lambda_0$. *Die Folge* $(T_\iota(z))$ *erfülle die Bedingung* (B_ι) *und besitze für alle* $z_\iota\in K$ *mit* $z_\iota\to z_0$ $(\iota\in\Lambda)$ *die Eigenschaft*

$$(11)\qquad \bigwedge_{u\in E}\ \bigwedge_{(u_\iota)\in\prod D(T_\iota)} u_\iota\rightharpoonup u,\ T_\iota(z_\iota)u_\iota\to v\ (\iota\in\Lambda)\succ u\in D(T),\ T(z_0)u=v.$$

Ferner sei das Paar $(T(z),(T_\iota(z)))$ *konsistent in* K. *Schließlich sei eine Folge* $(K_\iota)\in B(E_\iota,F_\iota)$ *gegeben mit*

$$(12)\qquad \bigwedge_{(u_\iota)\in\prod E_\iota} u_\iota\rightharpoonup 0\ (\iota\in\Lambda)\succ\|K_\iota u_\iota\|\to 0\ (\iota\in\Lambda),$$

so daß folgende Abschätzung gilt:

(13) $$\bigwedge_{\iota\in\Lambda'}\ \bigwedge_{u\in D(T_\iota)}\ \bigwedge_{z_\iota\in K} \gamma\,\|u\| \leqq \|T_\iota(z_\iota)\,u\| + \|K_\iota u\|\,.$$

Dann ist das Paar $(T(z), (T_\iota(z)))$ *a-regulär in* K.

Beweis. Wir nehmen zunächst an, für eine Folge $(u_\iota)\in\prod D(T_\iota)$, $\|u_\iota\|=1$, $\iota\in\Lambda$, und für $z_\iota\in K$, $\iota\in\Lambda$, gelte:

(14) $$T_\iota(z_\iota)\,u_\iota\to 0 \qquad (\iota\in\Lambda).$$

Die Konvergenz $z_\iota\to z_0\in K$ $(\iota\in\Lambda)$ und $u_\iota\rightharpoonup u$ $(\iota\in\Lambda)$ kann aufgrund der Voraussetzungen angenommen werden. Mit (11) haben wir also $u\in D(T)$, $T(z_0)\,u=0$.

Wenn nun $z_0\in\rho(T)$, folgt $u=0$, d.h., zusammen mit (12) steht (14) im Widerspruch zu (13). Für $z_0\in\rho(T)$ haben wir damit die inverse Stabilitätsbedingung für $(T_\iota(z_\iota))$ nachgewiesen. Zusammen mit der Surjektivität von $T(z_0)$ und 3.(8) erbringt nun 3.(4) die a-Regularität in kompakten Teilmengen von $K\cap\rho(T)$.

Wenn hingegen $T(z_0)$ nicht injektiv ist, müssen wir noch zeigen, daß $u_\iota\to u$ $(\iota\in\Lambda)$. Dazu benützen wir die Konsistenzbedingung: Zu u und z_0 gibt es $(v_\iota)\in\prod D(T_\iota)$ und $y_\iota\in K$, $\iota\in\Lambda_0$, mit $v_\iota\to u$, $y_\iota\to z_0$, $T_\iota(y_\iota)\,v_\iota\to 0$ $(\iota\in\Lambda_0)$. Nun erbringt (13):

$$\gamma\,\|u_\iota-v_\iota\| \leqq \|T_\iota(z_\iota)(u_\iota-v_\iota)\| + \|K_\iota(u_\iota-v_\iota)\|$$
$$\leqq \|T_\iota(z_\iota)\,u_\iota - T_\iota(y_\iota)\,v_\iota\| + \|(T_\iota(y_\iota)-T_\iota(z_\iota))\,v_\iota\| + \|K_\iota(u_\iota-v_\iota)\| \to 0 \quad (\iota\in\Lambda).$$

Dabei haben wir (14), 3.(8) und die Tatsache verwendet, daß mit $v_\iota\to u$ $(\iota\in\Lambda)$ auch $v_\iota\rightharpoonup u$ $(\iota\in\Lambda)$, d.h. $(u_\iota-v_\iota)\rightharpoonup 0$ $(\iota\in\Lambda)$ konvergiert.

A-priori Abschätzungen wie (13) hat man z.B. in Form von Koerzivitätsungleichungen bei der Differenzenapproximation oder Störung von gewöhnlichen Differentialoperatoren. Die Arbeiten [7] bzw. [9] von Grigorieff behandeln diese Problemkreise. Im Zusammenhang mit dem Dirichletschen Problem für nichtelliptische partielle Differentialoperatoren gibt Simader [17] Abschätzungen vom Typus (13) an, die die Untersuchung von Störungen dieser Operatoren bei simultaner Störung von Grundgebiet und Daten zulassen. Im Gegensatz zu [7, 9], wo T, T_ι, $\iota\in\Lambda$, beschränkte Operatoren sind, gelten die Abschätzungen aus [17] für einen abschließbaren Operator. Dieser Fall und der, daß man Probleme wie das in Abschnitt 1 im Anschluß an die Bedingung (B) gegebene Beispiel nicht als stetige Abbildung von $W^{m,2}(\mathbb{R}^+)$ in $L^2(\mathbb{R}^+)$ (wie in [7, 9]), sondern direkt in $L^2(\mathbb{R}^+)$ hinsichtlich Störungen und Approximationen untersuchen möchte, sind Beispiele für Operatoren, die erst mit der hier vorgelegten Theorie behandelt werden können.

Literatur

1. Anselone, P. M., Palmer, T. W.: Spectral analysis of collectively compact, strongly convergent operator sequences. Pacific J. Math. **25**, 423–431 (1968).
2. Atkinson, K. E.: The numerical solution of the eigenvalue problem for compact integral operators. Trans. Amer. Math. Soc. **129**, 458–467 (1967).

3. Brakhage, H.: Über die numerische Behandlung von Integralgleichungen nach der Quadraturformelmethode. Numerische Math. **2**, 183–196 (1960).
4. Bruhn, G., Wendland, W.: Über die näherungsweise Lösung von linearen Funktionalgleichungen. Funktionalanalysis, Approximationstheorie, Numerische Mathematik, S. 136–164. Basel-Stuttgart: Birkhäuser 1967.
5. Dieudonne, J.: Elements d'analyse. Tome I. Paris: Gauthier-Villars 1968.
6. Goldberg, S.: Unbounded linear operators with applications. New York: McGraw-Hill 1966.
7. Grigorieff, R. D.: Approximation von Eigenwertproblemen und Gleichungen zweiter Art in Hilbertschen Räumen. Math. Ann. **183**, 45–77 (1969).
8. — Über die Lösung regulärer und koerzitiver Rand- und Eigenwertaufgaben mit dem Galerkinverfahren. Manuscripta math. **1**, 385–411 (1969).
9. — Die Konvergenz des Rand- und Eigenwertproblems linearer gewöhnlicher Differenzengleichungen. Numerische Math. **15**, 13–48 (1970).
10. — Zur Theorie linearer approximationsregulärer Operatoren. Manuskript.
11. — Jeggle, H.: Quadraturformelmethoden zur Lösung nichtlinearer Eigenwertaufgaben bei Integralgleichungen. Erscheint in Z. Angew. Math. Mech. **52** (1972).
12. — — Diskrete Approximation nichtlinearer Eigenwertaufgaben. Manuskript.
13. Jeggle, H.: Bemerkungen zur näherungsweisen Lösung von Operatorgleichungen der mathemathischen Physik. Meth. Verf. math. Phys. **3**, 1–28 (1970).
14. — Diskrete Approximation von Eigenwertproblemen. Habilitationsschrift. Darmstadt 1971.
15. Knauer, B.: Über das Galerkinverfahren bei Eigenwertproblemen unbeschränkter Operatoren. Math. Z. **123**, 113–121 (1971).
16. Petryshyn, W. V.: On projectional solvability and the Fredholm alternative for equations involving linear A-proper mappings. Arch. Rat. Mech. Analysis **30**, 270–284 (1968).
17. Simader, C. G.: Zum Dirichletproblem für nichtelliptische partielle Differentialgleichungen. Vortrag in Oberwolfach, 1. 3. 1971.
18. Stummel, F.: Diskrete Konvergenz linearer Operatoren I. Math. Ann. **190**, 45–92 (1970).
19. — Diskrete Konvergenz linearer Operatoren II. Erscheint in der Math. Z.
20. — Diskrete Konvergenz linearer Operatoren III. Manuskript.
21. Taylor, A. E.: Introduction to functional analysis. New York: Wiley 1958.

Prof. Dr. Hansgeorg Jeggle
Fachbereich Mathematik
der Technischen Hochschule
D-6100 Darmstadt, Hochschulstraße 1

(Eingegangen am 7. Juli 1971)

Math. Z. 124, 343–352 (1972)

Bounding the Rank of Certain Permutation Groups

PETER J. CAMERON

If G is a finite transitive permutation group on a set Ω, the *rank* of G is the number of orbits of the stabiliser G_α of a point $\alpha \in \Omega$. If $\Gamma_1(\alpha), \ldots, \Gamma_r(\alpha)$ are the orbits of G_α, define the *subrank* of G to be $\max\limits_{1 \leqq i \leqq r} \operatorname{rank}(G_\alpha^{\Gamma_i(\alpha)})$. The following result is a consequence of a theorem of Manning [6]:

Theorem 1. *If G is a primitive permutation group with subrank* 2, *then one of the following holds:*

(i) *G has rank* 2 *(i.e. G is triply transitive).*

(ii) *G is a Frobenius group with Frobenius complement of order* 2.

This and other evidence suggests the conjecture that if G is a primitive permutation group with given subrank m, then either the rank of G is bounded or G_α has non-trivial regular orbits. It may even be true that there are only finitely many isomorphism classes of primitive permutation groups G with subrank m and rank greater than m in which G_α has no regular orbits, for any given m. In this paper I verify this conjecture for $m=3$.

Theorem 2. *If G is a primitive permutation group with subrank* 3, *then one of the following holds:*

(i) *G has rank at most* 3.

(ii) *G is a Frobenius group with Frobenius complement of order* 3.

(iii) *$G \cong [V_{16}]\, D_{10}$, with degree* 16 *and rank* 4.

(Here $[V_{16}]\, D_{10}$ denotes a split extension of an elementary abelian group of order 16 by a dihedral group of order 10.)

The proof given here depends on the graph-theoretic notation of Sims, Higman, and others, and elementary facts about permutation characters. These ideas are discussed briefly below.

A number of interesting groups satisfy the hypotheses and conclusion (i) of Theorem 2. These include the affine symplectic and orthogonal groups in even dimensions over $GF(2)$, the projective special linear groups (except $PSL(2, 2^a)$), and simple groups discovered by Higman and Sims, McLaughlin, and Conway.

§ 1. Notation

All groups considered are finite and act on finite sets. For the notation and theory of finite permutation groups, see Wielandt [10].

$[H]K$ denotes a split extension of the group H by the group K, i.e. a group G with subgroups H and K such that $HK=G$, $H\cap K=1$, and H is normal in G. It acts as a permutation group on the elements of H, having H as a regular normal subgroup; the stabiliser of the identity is K.

Z_n denotes the cyclic group of order n; V_q, the elementary abelian group of order q (where q is a prime power); D_{2n}, the dihedral group of order $2n$; $PSL(n,q)=X/Z(X)$, where X is the group of $n\times n$ matrices over the field $GF(q)$ having determinant 1, and $Z(X)$, the centre of X, consists of the scalar matrices in X.

If φ and ψ are (complex) characters of a group G, $(\varphi,\psi)_G$ denotes their inner product. 1 will be used to denote the integer, or the identity element or the principal character of any relevant group, depending on the context.

§ 2. Graph Theory

If G is a transitive permutation group on Ω, and Δ is a subset of the cartesian product $\Omega\times\Omega$ which is fixed by G (acting in the natural way on $\Omega\times\Omega$), then $\Delta(\alpha)=\{\beta\in\Omega\,|\,(\alpha,\beta)\in\Delta\}$ is a subset of Ω fixed by G_α; Δ thus defines an "orbital" function. This procedure sets up a 1-1 correspondence between G-orbits in $\Omega\times\Omega$ and G_α-orbits in Ω. (The number of such orbits is the rank of G.) $\Delta^*=\{(\beta,\alpha)\,|\,(\alpha,\beta)\in\Delta\}$ is the orbital (or subset of $\Omega\times\Omega$) paired with Δ; Δ is self-paired if $\Delta=\Delta^*$. Note that $|\Delta(\alpha)|=|\Delta^*(\alpha)|=|\Delta|/|\Omega|$. If Γ and Δ are two fixed sets of G in $\Omega\times\Omega$, let $\Gamma\circ\Delta$ denote the set $\{(\alpha,\beta)\,|\,\exists\gamma\in\Omega$ such that $(\alpha,\gamma)\in\Gamma$, $(\gamma,\beta)\in\Delta;\ \alpha\neq\beta\}$. This is also a fixed set of G. The diagonal $\{(\alpha,\alpha)\,|\,\alpha\in\Omega\}$ is a single "trivial" G-orbit. Suborbits of G are G_α-orbits in Ω; when there is no confusion, I shall use the term for G-orbits in $\Omega\times\Omega$ also.

If Γ is a G-orbit in $\Omega\times\Omega$ which is not the diagonal, the Γ-graph is the regular directed graph whose vertex set is Ω and whose edges are precisely the ordered pairs in Γ. It admits G as a flag-transitive group of automorphisms. If Γ is self-paired, the Γ-graph can be regarded as an undirected graph. It is helpful to think of a colour associated with each non-trivial G-orbit in $\Omega\times\Omega$, and to suppose that in the complete directed graph on Ω all edges belonging to the Γ-graph are coloured with the colour associated with Γ. A connected component of any such graph is a block of imprimitivity for G; G is primitive if and only if every such graph is connected.

Further detail can be found in papers of Sims (for example, [9]) or Higman; the "coloured graph" notation is due to Neumann.

§ 3. Character Theory

Let G be a group which acts as a permutation group on Ω, and π the permutation character of G, the integer-valued function on G defined by $\pi(g)=$ number of fixed points of g, for all $g\in G$. The formula

$$\frac{1}{|G|}\sum_{g\in G}\pi(g)=\text{number of orbits of } G,$$

is well-known and easy to prove. In terms of inner products of characters, it takes the form $(\pi, 1)_G$ = number of orbits of G. In particular, if G is transitive, the multiplicity of the principal character as a constituent of π is one.

If G acts as a permutation group on Ω_1 and Ω_2, with permutation characters π_1 and π_2, the number m of G-orbits in $\Omega_1 \times \Omega_2$ is

$$m=(\pi_1\pi_2, 1)_G=(\pi_1, \pi_2)_G,$$

since $\pi_1\pi_2$ is the permutation character of G on $\Omega_1\times\Omega_2$. (Note that if G is transitive on Ω_1 and Ω_2 and $\omega\in\Omega_1$, there is a natural 1-1 correspondence between G-orbits in $\Omega_1\times\Omega_2$ and G_ω-orbits in Ω_2, defined by $\Delta\leftrightarrow\Delta(\omega)=\{\alpha\in\Omega_2\,|\,(\omega,\alpha)\in\Delta\}$.) In particular, if G is a transitive permutation group on Ω with permutation character π, the rank r of G is given by

$$\begin{aligned} r&=(\pi, \pi)_G \\ &=\text{sum of squares of multiplicities of irreducible constituents of } \pi. \end{aligned}$$

Frobenius described an irreducible complex character ψ of a group G as being of the first, second, or third kind according as ψ is the character of a real representation, a real character which is not afforded by any real representation, or a non-real character, respectively. He proved that

$$\frac{1}{|G|}\sum_{g\in G}\psi(g^2)=+1,\ -1,\text{ or } 0$$

according as ψ is of the first, second, or third kind. (See Feit [2], 3.5.) The next lemma is an easy consequence of this result.

Lemma 1. *If G is a transitive permutation group, s the number of self-paired suborbits of G, and n_1, n_2, n_3 the numbers of irreducible constituents of the permutation character of G of the first, second, and third kind respectively (counted with multiplicity), then $s=n_1-n_2$.*

Note that an irreducible character of the second kind must have even multiplicity in π, since π is afforded by a real representation. So if π is "multiplicity-free" (i.e. if every irreducible constituent of π has multiplicity 1), this takes the form

Number of self-paired suborbits = number of real irreducible constituents of π;

Number of non-self-paired suborbits = number of non-real irreducible constituents of π.

Lemma 2. *Let G be a transitive permutation group on Ω, with permutation character π.*

(i) *G is 2-transitive if and only if $\pi=1+\psi$ where ψ is a non-principal irreducible character.*

(ii) *G has rank 3 if and only if $\pi=1+\psi_1+\psi_2$, where ψ_1 and ψ_2 are distinct non-principal irreducible characters.*

(iii) *G is 2-homogeneous (transitive on the set of 2-element subsets of Ω) but not 2-transitive if and only if $\pi = 1 + \psi + \bar{\psi}$, where ψ is a non-real irreducible character and $\bar{\psi}$ its complex conjugate.*

(G is 2-homogeneous but not 2-transitive if and only if G has rank 3 and the non-trivial suborbits are paired with each other. In this case G has odd order.)

Lemma 3. *Let G be a transitive permutation group on Ω with subrank m. Let α be a point of Ω, and $\Gamma(\alpha)$ and $\Delta(\alpha)$ two G_α-orbits (not necessarily distinct). Take a point $\gamma \in \Gamma(\alpha)$. The number of $G_{\alpha\gamma}$-orbits in $\Delta(\alpha)$ is at most m; if this number is m, then $|\Gamma(\alpha)| = |\Delta(\alpha)|$, $G_\alpha^{\Gamma(\alpha)}$ and $G_\alpha^{\Delta(\alpha)}$ have rank m, and G_α has the same permutation character on $\Gamma(\alpha)$ and $\Delta(\alpha)$.*

Proof. Let r_1 and r_2 be the ranks of $G_\alpha^{\Gamma(\alpha)}$ and $G_\alpha^{\Delta(\alpha)}$ respectively, and u the number of orbits of $G_{\alpha\gamma}$ in $\Delta(\alpha)$. Suppose G_α has permutation characters $\pi_1 = \sum e_i \psi_i$ and $\pi_2 = \sum f_i \psi_i$ on $\Gamma(\alpha)$ and $\Delta(\alpha)$ respectively, where the ψ_i are distinct irreducible characters and e_i and f_i are non-negative integers.

Then
$$r_1 = \sum e_i^2 \leqq m,$$
$$r_2 = \sum f_i^2 \leqq m,$$
and
$$\begin{aligned} u^2 &= \left(\sum e_i f_i\right)^2 \\ &\leqq \left(\sum e_i^2\right)\left(\sum f_i^2\right) \quad \text{(Cauchy's inequality)} \\ &= r_1 r_2 \\ &\leqq m^2, \end{aligned}$$
so
$$u \leqq m.$$

If equality holds, then $r_1 = r_2 = m$ and $e_i = f_i$, so $\pi_1 = \pi_2$.

§ 4. Lemmata

Lemma 4 (Manning [6]). *If G is primitive on Ω and G_α is doubly transitive on $\Gamma(\alpha)$, with $|\Gamma(\alpha)| = d > 2$, then either G is triply transitive or G_α has an orbit $\Delta(\alpha)$ with $|\Delta(\alpha)| > d$, $|\Delta(\alpha)|$ divides $d(d-1)$.*

This result, together with the fact that a primitive group with a subdegree 2 is a Frobenius group (see [10], Theorem 18.7), proves Theorem 1.

Lemma 5. *If G is primitive and not regular on Ω, $\Delta(\alpha)$ is a suborbit of maximal size, $\Gamma(\alpha)$ is any non-trivial suborbit, and $\gamma \in \Gamma(\alpha)$, then $G_{\alpha\gamma}$ is not transitive on $\Delta(\alpha)$, and so the permutation characters of G_α on $\Gamma(\alpha)$ and $\Delta(\alpha)$ have a non-principal irreducible constituent in common.*

Proof. If $G_{\alpha\gamma}$ is transitive on $\Delta(\alpha)$, then the maximality of $|\Delta(\alpha)|$ implies that $\Delta(\alpha)$ is a single G_γ-orbit. Since G_α and G_γ generate G ([10], Proposition 8.7), $\Delta(\alpha)$ is a fixed set of G, which is false.

Lemma 6. *If G has rank 3 on Ω, then any connected component of either of the graphs associated with G which is not the whole of Ω is a clique.*

Proof. If G is an imprimitive rank 3 permutation group, then the stabiliser of a block is doubly transitive on the points of the block.

From now on, assume that G is a primitive permutation group on Ω which has subrank 3.

Lemma 7. *If G_α is 2-homogeneous but not 2-transitive on a suborbit $\Gamma(\alpha)$, then G is $\frac{3}{2}$-transitive, and G_α is 2-homogeneous on every nontrivial suborbit.*

Proof. By Lemma 3, G_α has permutation character $\pi = 1 + \psi + \bar{\psi}$ on $\Gamma(\alpha)$, where ψ is a non-real irreducible character and $\bar{\psi}$ its complex conjugate. A rational character which contains either ψ or $\bar{\psi}$ contains both. Applying Lemma 5 twice, G_α is 2-homogeneous on the largest suborbit and then on every non-trivial suborbit; the permutation characters on all non-trivial suborbits are equal.

Lemma 8. *For any suborbit Γ, $\Gamma \circ \Gamma^*$ is the union of at most 2 suborbits.*

Proof. G has at most 2 orbits on triples $(\alpha, \gamma_1, \gamma_2)$, with $\gamma_1, \gamma_2 \in \Gamma^*(\alpha)$ and $\gamma_1 \neq \gamma_2$.

Lemma 9. *If $\Gamma = \Gamma^*$ and $\Delta \subseteq \Gamma \circ \Gamma$ $(\Delta \neq \Gamma)$, then for any $\delta \in \Delta(\alpha)$, the set $\Gamma(\alpha) \cap \Gamma(\delta)$ is not a union of non-trivial connected components of the restriction of the Γ-graph to $\Gamma(\alpha)$.*

Proof. Suppose the conclusion is false. If $\Gamma(\alpha) \cap \Gamma(\delta) = \Gamma(\alpha)$, then $\Gamma(\delta) = \Gamma(\alpha)$, contradicting the primitivity of G. Otherwise, the connected components are proper subsets of $\Gamma(\alpha)$, and so by Lemma 6 they are cliques. We can choose γ_1 and γ_2 in $\Gamma(\alpha)$ so that $(\gamma_1, \gamma_2) \in \Delta$; then $\alpha \in \Gamma(\gamma_1) \cap \Gamma(\gamma_2)$, and α thus lies in a clique of the Γ-graph in $\Gamma(\gamma_1) \cap \Gamma(\gamma_2)$. So there is a point in $\Gamma(\gamma_1) \cap \Gamma(\gamma_2) \cap \Gamma(\alpha)$. Thus γ_1 and γ_2 are joined by a path of length 2 in the restriction of the Γ-graph to $\Gamma(\alpha)$. Since connected components of this graph are cliques, $(\gamma_1, \gamma_2) \in \Gamma$, contrary to assumption.

Lemma 10. *If $\Gamma = \Gamma^*$ and $\Gamma \circ \Gamma = \Gamma \cup \Delta$, then $\Gamma \cup \Delta \subseteq \Gamma \circ \Delta$.*

Proof. $\Delta \subseteq \Gamma \circ \Gamma$, so $\Gamma \subseteq \Gamma \circ \Delta$. Suppose $\Delta \nsubseteq \Gamma \circ \Delta$. Take $\delta \in \Delta(\alpha)$, and let γ be a point in $\Gamma(\alpha) \cap \Gamma(\delta)$. Then $|\Gamma(\gamma) \cap \Gamma(\delta)| = |\Gamma(\gamma) \cap \Gamma(\alpha)| > 0$ and $\Gamma(\gamma) \cap \Gamma(\delta)$ contains no points of $\Delta(\alpha)$ (since $\Gamma(\delta) \cap \Delta(\alpha)$ is empty); so $\Gamma(\gamma) \cap \Gamma(\delta) \subseteq \Gamma(\alpha)$, and $\Gamma(\gamma) \cap \Gamma(\delta) = \Gamma(\gamma) \cap \Gamma(\alpha)$. It follows that $\Gamma(\alpha) \cap \Gamma(\delta)$ is a union of connected components of the Γ-graph in $\Gamma(\alpha)$, contradicting Lemma 9.

§ 5. Small Subdegrees

As before, G is a primitive permutation group with subrank 3. G cannot have a subdegree 2, since primitive groups with a subdegree 2 are Frobenius groups and have subrank 2 ([10], Theorem 18.7).

Suppose that G has no subdegree greater than 5. By Lemma 4, G_α has rank 3 on all the non-trivial suborbits. Since 3, 4, and 5 are pairwise coprime, [10] Theorem 17.5 implies that G is $\frac{3}{2}$-transitive.

(i) If $|\Gamma(\alpha)|=3$ and $G_\alpha^{\Gamma(\alpha)}\cong Z_3$, then $G_\alpha\cong Z_3$ (Lemma 7) and G is a Frobenius group satisfying conclusion (ii) of the theorem.

(ii) Suppose $|\Gamma(\alpha)|=4$ and $G_\alpha^{\Gamma(\alpha)}\cong D_8$. Then G is in the class of groups determined by Sims [9 II]. Inspection of his list shows that G_α acts faithfully on $\Gamma(\alpha)$ and G has an elementary abelian normal subgroup V. (The other groups have subdegrees greater than 4.) Regarding V as a vector space over $GF(p)$ and identifying Ω with V, G_0 acts as an irreducible linear group on V. It is easy to see that V has dimension 2, and that if $G_0=\langle a,b\,|\,a^4=b^2=(ab)^2=1\rangle$ then $V=\langle x,y\rangle$ where $x^a=y, y^a=x^{-1}, x^b=x, y^b=y^{-1}$. (Note that $p\neq 2$.) If $j\in GF(p)-\{0,+1,-1\}$, $x\,y^j$ is in a G_0-orbit of length 8. So we must have $p=3$, whence $|\Omega|=9$, $G\cong[V_9]\,D_8$, and G has rank 3.

(iii) Suppose $|\Gamma(\alpha)|=5$ and $G_\alpha^{\Gamma(\alpha)}\cong D_{10}$. Then G is a $Z(r-1,5)$ group in the notation of McDermott [7]. (G_α acts primitively on every non-trivial suborbit, so it acts faithfully on $\Gamma(\alpha)$, by a theorem of Manning [5].) By [7], Lemma 2.3 (noting that we do not need McDermott's assumption that G is simple, since we already know G contains an involution), $r|4$, so $r=2$ or 4. If $r=2$ then $|\Omega|=6$, $|G|=60$, and $G\cong PSL(2,5)$; G is 2-transitive. If $r=4$ then $|\Omega|=16$, $|G|=160$. G is soluble and so has a regular normal subgroup; thus $G\cong[V_{16}]\,D_{10}$. This is case (iii) of the theorem.

§ 6. The Self-Paired Case

Now we may assume that G is primitive on Ω with subrank 3, and that G_α has an orbit $E(\alpha)$ with $|E(\alpha)|\geqq 6$. There are two cases:

A. All suborbits of $G_\alpha^{E(\alpha)}$ are self-paired.

B. $G_\alpha^{E(\alpha)}$ is 2-homogeneous.

Assume case A occurs. The complete graph on $E(\alpha)$ is the union of the restrictions to $E(\alpha)$ of at most two undirected graphs associated with G. By Ramsey's theorem (M. Hall [3], Section 6.1) at least one of these must contain a triangle. So there is a suborbit Γ such that Γ is self-paired and $\Gamma\subseteq\Gamma\circ\Gamma$. If $\Gamma=\Gamma\circ\Gamma$ then $\{\alpha\}\cup\Gamma(\alpha)$ is a connected component of the Γ-graph, so $\Omega=\{\alpha\}\cup\Gamma(\alpha)$ and G is 2-transitive. Otherwise, by Lemma 8, $\Gamma\circ\Gamma=\Gamma\cup\Delta$ for some self-paired suborbit Δ, and, by Lemma 10, $\Gamma\cup\Delta\subseteq\Gamma\circ\Delta$. If $\Gamma\circ\Delta=\Gamma\cup\Delta$, then $\{\alpha\}\cup\Gamma(\alpha)\cup\Delta(\alpha)$ is a connected component of the Γ-graph, so G has rank 3. Suppose this is not the case; then, by Lemma 3, $\Gamma\circ\Delta$ contains one more suborbit, $|\Gamma(\alpha)|=|\Delta(\alpha)|$, and G_α has the same permutation character on $\Gamma(\alpha)$ and $\Delta(\alpha)$.

Suppose the parameters of the Γ-graph are as shown in Fig. 1. $\gamma\in\Gamma(\alpha)$. Circles represent $G_{\alpha\gamma}$-orbits, numbers in circles are the sizes of these orbits, and other numbers represent numbers of edges joining a point in one orbit to points in another orbit.

Any point in $\Delta(\alpha)$ is joined to l points in $\Gamma(\alpha)$ (since any point in $\Gamma(\alpha)$ is joined to l points in $\Delta(\alpha)$, and $|\Gamma(\alpha)|=|\Delta(\alpha)|$). Since $\Gamma\circ\Delta\neq\Gamma\cup\Delta$, there are points of $\Delta(\alpha)$ not joined to γ by a path of length 1 or 2. Suppose B is the set

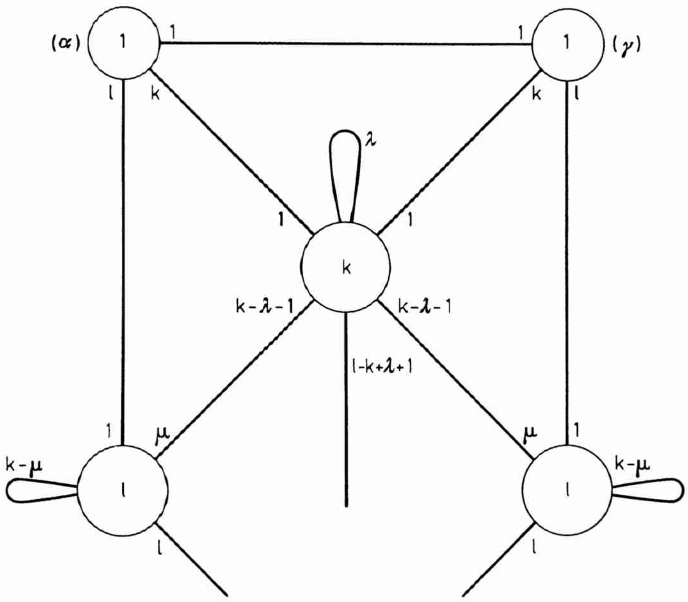


Fig. 1

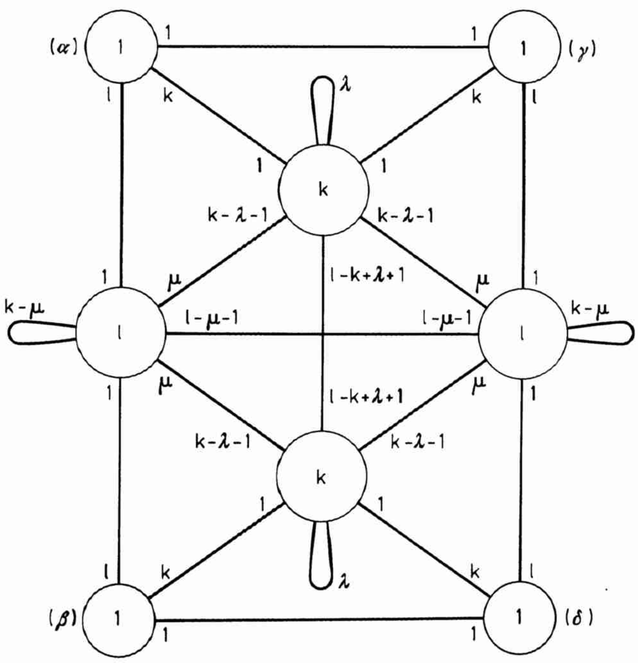


Fig. 2

of such points, and $|B|=b$. If $\beta\in B$, $\Gamma(\alpha)\cap\Gamma(\beta)=\Gamma(\alpha)\cap\Delta(\gamma)$, and this set has size l. By Lemma 9, the Γ-graph in $\Gamma(\alpha)$ is connected.

(i) Suppose $b=1$. Then the graph is as follows.

There is a unique point δ in $\Gamma(\beta)$ but not in $\Gamma(\alpha)$ or $\Delta(\alpha)$. Also, the Γ-graph in $\Delta(\alpha)$ is connected. If $\eta\in\Gamma(\beta)\cap\Delta(\alpha)$, then there are k points in $\Gamma(\beta)\cap\Gamma(\eta)$. Of these, $k-\lambda-1$ lie in $\Gamma(\alpha)\cap\Delta(\gamma)$, and λ lie in $\Gamma(\beta)\cap\Delta(\alpha)$, so there is just one more, which must be δ. Since the Γ-graph in $\Delta(\alpha)$ is connected, δ is joined to every point of $\Delta(\alpha)$; so $\Delta(\alpha)=\Gamma(\delta)$, contradicting the primitivity of G.

(ii) Suppose $b>1$. For any $\beta_1,\beta_2\in B$, $\Gamma(\alpha)\cap\Gamma(\beta_1)=\Gamma(\alpha)\cap\Gamma(\beta_2)$. So B is a block for $G_\alpha^{\Delta(\alpha)}$. Now the set-wise stabiliser of $\Gamma(\alpha)\cap\Delta(\gamma)$ (in G_α) contains the set-wise stabiliser of B; so $\Gamma(\alpha)\cap\Delta(\gamma)$ is a union of proper blocks for $G_\alpha^{\Gamma(\alpha)}$. By Lemma 9, these blocks are not cliques in the Γ-graph; so $\{\gamma\}\cup(\Delta(\gamma)\cap\Gamma(\alpha))$ is a clique in the Δ-graph, i.e. $\Gamma(\alpha)\cap\Delta(\gamma)$ is a void graph in the Γ-graph. Thus there are no figures of the shape shown in Fig. 3 in the Γ-graph. This condition, however, implies that there are no paths of length 2 in the Γ-graph in $\Gamma(\alpha)$, which is clearly false, since this graph is connected and has more than 2 vertices.

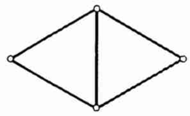


Fig. 3

§ 7. The 2-Homogeneous Case

Suppose case B of the previous section occurs. Lemma 7 shows that G is $\frac{3}{2}$-transitive and G_α is 2-homogeneous on every non-trivial suborbit. Let v be the common subdegree; then $v\equiv 3 \pmod 4$ and $v\geq 7$. Let Γ be any non-trivial suborbit.

If $\Gamma\circ\Gamma^*$ is a single suborbit Δ, then $|\Gamma(\alpha)\cap\Gamma(\delta)|=v-1$ for $\delta\in\Delta(\alpha)$. So for any $\delta_1,\delta_2\in\Delta(\alpha)$, $\Gamma(\alpha)\cap\Gamma(\delta_1)\cap\Gamma(\delta_2)$ is non-empty, and so $\delta_2\in\Delta(\delta_1)$. Thus $\{\alpha\}\cup\Delta(\alpha)$ is a connected component of the Δ-graph, and G is 2-transitive. (In this case, theorems of Bender [1] and Huppert [4] show that either

(i) G has a normal subgroup isomorphic to $PSL(2,q)$, for some prime power $q\equiv 3 \pmod 4$, acting on the projective line over $GF(q)$ or (ii) G is a soluble group of degree 8 and order 168.)

Otherwise, $\Gamma\circ\Gamma^*$ is the union of two suborbits, which are easily seen to be paired with each other, say Δ and Δ^*. For $\delta\in\Delta(\alpha)$, $|\Gamma(\alpha)\cap\Gamma(\delta)|=\frac{1}{2}(v-1)$, and the translates under G_α of $\Gamma(\alpha)\cap\Gamma(\delta)$ form the blocks of a symmetric $(v,\frac{1}{2}(v-1),\frac{1}{4}(v-3))$ block design on the points of $\Gamma(\alpha)$. (See Hall [3], Chapter 10.) Any two blocks of such a design have $\frac{1}{4}(v-3)$ common points; so $\delta_2\in(\Gamma\circ\Gamma^*)(\delta_1)$ for any $\delta_1,\delta_2\in\Delta(\alpha)$. Similarly, the points of $\Delta^*(\alpha)$ give rise to a design on the points of $\Gamma(\alpha)$, and as before $\eta_2\in(\Gamma\circ\Gamma^*)(\eta_1)$ for any $\eta_1,\eta_2\in\Delta^*(\alpha)$.

Now take $\delta \in \Delta(\alpha)$. If $\eta \in (\Gamma \circ \Gamma^*)(\delta)$ for any $\eta \in \Delta^*(\alpha)$, then $\{\alpha\} \cup (\Gamma \circ \Gamma^*)(\alpha)$ is a connected component of the Δ-graph, and G has rank 3. (Now $|\Omega| = 2v+1$ is odd and $|G_\alpha|$ is odd, so $|G|$ is odd and G is soluble; Passman's classification of $\frac{3}{2}$-transitive soluble groups [8] shows that there are no groups of this form at all.) So we may now assume that there is a point $\eta \in \Delta^*(\alpha)$ such that

$$\eta \notin (\Gamma \circ \Gamma^*)(\delta).$$

The block in the Δ^*-design defined by η is disjoint from the block in the Δ-design defined by δ. There is exactly one such point η, for if two blocks were disjoint from a given block, they would have at least $\frac{1}{2}(v-3)$ common points.

Now $|\Delta(\delta) \cap \Delta(\alpha)| = |\Delta^*(\delta) \cap \Delta(\alpha)| = |\Delta(\eta) \cap \Delta(\alpha)| = \frac{1}{2}(v-1)$, and so

$$|\Delta^*(\eta) \cap \Delta(\alpha)| = \tfrac{1}{2}(v-1)$$

also, for the union of the last two sets mentioned is $\Delta(\alpha) - \{\delta\}$. Thus only one point of $\Delta(\delta)$ is not in $(\Gamma \circ \Gamma^*)(\alpha)$. The parameters of the Δ-graph are shown in Fig. 4.

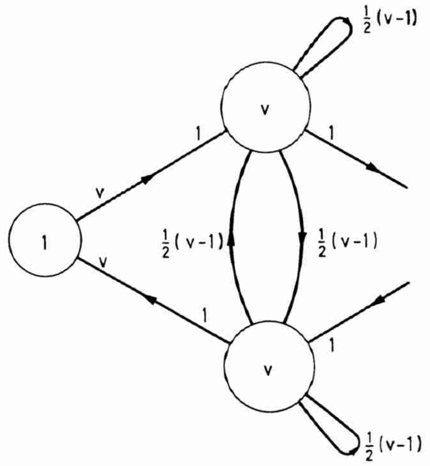


Fig. 4

The translates of $\Delta(\delta) \cap \Delta(\alpha)$ under G_α give rise to a $\left(v, \frac{1}{2}(v-1), \frac{1}{4}(v-3)\right)$ design on $\Delta(\alpha)$, and the translates of $\Delta(\delta) \cap \Delta^*(\alpha)$ give rise to a design with the same parameters in $\Delta^*(\alpha)$. So for $\delta_1, \delta_2 \in \Delta(\alpha)$ we have

$$|\Delta(\delta_1) \cap \Delta(\delta_2) \cap \Delta(\alpha)| = \tfrac{1}{4}(v-3)$$

and

$$|\Delta(\delta_1) \cap \Delta(\delta_2) \cap \Delta^*(\alpha)| = \tfrac{1}{4}(v-3).$$

$|\Delta(\delta_1) \cap \Delta(\delta_2)| = \frac{1}{2}(v-1)$, so $\Delta(\delta_1) \cap \Delta(\delta_2)$ contains one further point ω. Then $\omega \in \Delta(\delta)$ for all $\delta \in \Delta(\alpha)$, and so $\Delta(\alpha) = \Delta^*(\omega)$, contradicting the primitivity of G.

This completes the proof of Theorem 2.

Acknowledgment. This paper was prepared while the author held a Rhodes Scholarship at Oxford University. The main theorem, however, was proved while travelling on a German train, on a grant from the Mathematisches Forschungsinstitut, Oberwolfach. The author wishes to express special gratitude to Professor D. G. Higman, without whose interest in the original conjecture the present theorem would not have been proved.

References

1. Bender, H.: Endliche zweifach transitive Permutationsgruppen, deren Involutionen keine Fixpunkte haben. Math. Z. **104**, 175–204 (1968).
2. Feit, W.: Characters of finite groups. New York-Amsterdam: W. A. Benjamin Inc. 1967.
3. Hall, M., Jr.: Combinatorial theory. Waltham, Mass.-Toronto-London: Blaisdell Publishing Co. 1967.
4. Huppert, B.: Zweifach transitive, auflösbare Permutationsgruppen. Math. Z. **68**, 126–150 (1957).
5. Manning, W. A.: Simply transitive primitive groups. Trans. Amer. Math. Soc. **29**, 815–825 (1927).
6. — A theorem concerning simply transitive primitive groups. Bull. Amer. Math. Soc. **35**, 330–332 (1929).
7. McDermott, J. P. J.: Characterisations of some 3/2-transitive groups. Math. Z. **120**, 204–210 (1971).
8. Passman, D. S.: Exceptional 3/2-transitive permutation groups. Pacific J. Math. **29**, 669–713 (1969).
9. Sims, C. C.: Graphs and finite permutation groups, I, II. Math. Z. **95**, 76–86 (1967); **103**, 276–281 (1968).
10. Wielandt, H.: Finite permutation groups. New York-London: Academic Pr. 1964.

Dr. P. J. Cameron
Mathematical Institute
24–29 St. Giles'
Oxford
England

(Received July 21, 1971)

Math. Z. 124, 353–360 (1972)

Fixed Point Subgroups that Contain Centralizers of Involutions

GEORGE GLAUBERMAN

1. Introduction and Notation

The purpose of this paper is to prove the following property of automorphism groups of finite groups:

Theorem 1. *Let G be a finite group and A be a subgroup of the automorphism group of G such that G and A have relatively prime orders. Let $C(A)$ be the fixed point subgroup of G with respect to A.*

Suppose $C(A)$ contains the centralizer in G of some element $\tau \in G$ of order two. Let N be the subgroup of $C(A)$ generated by the elements $t^{-1}\tau^{-1}t\tau$ for all $t \in C(A)$. Then:

(i) *N is a normal subgroup of G, and if $A \neq 1$, $\tau \notin N$.*

Let $\bar{G} = G/N$, $\bar{C} = C(A)/N$, and let $\bar{\tau}$ be the element τN of $\bar{G}$.

(ii) *A acts in an induced manner on $\bar{G}$; this action is faithful, and $\bar{C}$ is the fixed point subgroup of $\bar{G}$ with respect to A.*

(iii) *$\bar{\tau} \in \bar{C}$, and $\bar{C}$ is the centralizer of $\bar{\tau}$ in $\bar{G}$; if $A \neq 1$, $\bar{\tau}$ has order two.*

(iv) *The group (G, A) generated by the elements $g^{-1}g^{\alpha}$, $g \in G$, $\alpha \in A$, is a normal nilpotent subgroup of G which has odd order.*

Let C be a finite group containing an element τ of order two in its center. In [1], Brauer and Fowler prove that, up to isomorphism, there exist only finitely simple groups G such that $C \subseteq G$ and C is the centralizer of τ in G. In subsequent papers various authors have considered specific choices of C and have determined all the corresponding groups G. Thus it appears that the structure of a finite non-cyclic simple group (which must have even order, by the Feit-Thompson Theorem [3]) is intimately related to the structure of the centralizers of its elements of order two. The following corollary of Theorem 1 illustrates this relation.

Corollary 1. *Suppose G is a finite non-cyclic simple group or, more generally, a finite group having no non-identity normal subgroup of odd order. Let A be a subgroup of the automorphism group of G such that A and G have relatively prime orders. Let $C(A)$ be the fixed point subgroup of G with respect to A. If $C(A)$ contains some element τ of order two in G, then A acts faithfully on $C_G(\tau)$. Thus if*

$$C_G(\tau) \subseteq C(A) \subseteq G,$$

then $G = C(A)$, i.e., $A = 1$.

If G contains only one conjugate class of elements of order two, then by Lemma 4, $C(A)$ must contain at least one element of order two. Hence Corollary 1 applies in this case. This situation occurs in many of the known simple groups, for example, in the simple groups of orders 60 and 168.

By using a semi-direct product we obtain the following result, which is analogous to Corollary 1 and also to Thompson's "$P \times Q$ Lemma" on p-groups [8, Theorem 5.3.4].

Corollary 2. *Assume the hypothesis of Corollary 1. Suppose A has odd order, and suppose τ is an outer automorphism of G that centralizes A and has order two. Then A acts faithfully on the fixed point subgroup of G with respect to τ.*

These results raise several questions. We do not know whether we may take τ of odd prime order, or indeed of arbitrary order, in Corollaries 1 and 2. As one can see from the proof of Theorem 6 of [7], (G, A) need not be nilpotent in Theorem 1 if $|A|=2$ and τ has odd prime order. Moreover, we cannot replace τ by an elementary group of order four in Corollary 1. Let G be the linear fractional group $LF(2, 5^7)$ of degree two over the field F of 5^7 elements. Let A be a cyclic group of automorphisms of G that corresponds to the group of all field automorphisms of F. Then A has order seven and $C(A)$ is a subgroup of G of the form $LF(2, 5)$. Suppose S is a Sylow 2-subgroup of $C(A)$. Then S is an elementary group of order four; in fact, S is a Sylow 2-subgroup of G. However, $C_G(S) \subseteq N_G(S) \subseteq C(A)$, although G has order relatively prime to seven.

Notation. Let G be an arbitrary (multiplicative) group. We denote the identity element and identity subgroup of G by 1. Suppose $g, h, J, K, \ldots$ are elements or subsets of G. We let $\langle g, h, J, K, \ldots\rangle$ be the subgroup of G which they generate; if J is the empty set, define $\langle J\rangle = 1$. If $g, h, J, K, \ldots$ are elements of G, let $\{g, h, J, K, \ldots\}$ be the subset of G they constitute; also, define $g^h = h^{-1} g h$ and $(g, h) = g^{-1} h^{-1} g h$.

Suppose J and K are subsets of G. We write $J \subseteq K$ or $K \supseteq J$ ($J \subset K$ or $K \supset J$) if J is a (proper) subset of K. We define

$$(J, K) = \langle (g, h) \colon g \in J, h \in K\rangle.$$

If J is a group, let $N_J(K)$ and $C_J(K)$ be respectively the normalizer and centralizer of K in J. Let $Z(G)$ be the center of G.

Let us assume G is a finite group; all groups considered in this paper are finite. For every subset J of G we let $|J|$ be the number of elements in J. If J is a subgroup of G we let $[G:J]$ be the index of J in G. An element of G of order two is called an *involution.*

We say that a (finite) group A is a group of *operators*, or *operator group*, on G if, for every pair of elements $g \in G$, $\alpha \in A$, there is defined a unique element g^α of G such that

$$(g h)^\alpha = g^\alpha h^\alpha, \quad g^{\alpha\beta} = (g^\alpha)^\beta, \quad g^1 = g,$$

for arbitrary $g, h \in G$ and $\alpha, \beta \in A$. If J is a subset of G and $\alpha \in A$ we let $J^\alpha = \{g^\alpha \colon g \in J\}$. We say that α (or A) *fixes*, or *normalizes*, an element or subset J of

G if $J^\alpha = J$ (or $J^\alpha = J$, for all $\alpha \in A$). A fixed element will also be called a *fixed point*; the set of all such elements is the *fixed point subgroup* of G with respect to α (or A). We may clearly consider A to be an operator group on any subgroup J of G which A fixes; if J is a normal subgroup of G we may also consider A as an operator group on G/J, since A permutes the cosets of J. We define

$$(G, A) = \langle g^{-1} g^\alpha : g \in G, \alpha \in A \rangle .$$

If we consider G and A as subgroups of their semi-direct product ([9], pp. 88–90) then this definition coincides with the definition given above. We shall consider groups of automorphisms as groups of operators. Note that G is a group of operators on itself by the definition $g^h = h^{-1} g h$ given above.

Theorem 1 is the main result of a doctoral dissertation submitted to the University of Wisconsin. The author wishes to thank his thesis advisor, Professor R.H. Bruck, for his generous suggestions, encouragement, and advice over a period of several years. During a year's visit to Wisconsin, Professor H. Wielandt also gave many helpful and stimulating suggestions that were eventually incorporated in this work. In addition, the author is indebted to Professor J. Thompson for problems and ideas that led to this investigation. Finally, he thanks the citizens of the United States of America for their support via a National Science Foundation Graduate Fellowship during the preparation of the dissertation.

2. Some Preliminary Lemmas

Lemma 1. *Let ρ and σ be involutions in a group G and let $J = \langle \rho, \sigma \rangle$. Let $\nu = \rho\sigma$. Then $\langle \nu \rangle$ is a subgroup of index two in J and $\rho^{-1} \nu \rho = \sigma^{-1} \nu \sigma = \nu^{-1}$. If ν has odd order, say, $2n-1$, then $\rho = \nu^n \sigma \nu^{-n}$; thus ρ is conjugate to σ in J.*

Proof. First, $\rho^{-1} \nu \rho = \rho^{-1} (\rho \sigma) \rho = \sigma \rho = \sigma^{-1} \rho^{-1} = \nu^{-1} = \sigma^{-1} \rho^{-1} = \sigma^{-1} \rho = \sigma^{-1} (\rho \sigma) \sigma = \sigma^{-1} \nu \sigma$. Thus $\langle \nu \rangle$ is a normal subgroup of J. Now $J = \langle \rho\sigma, \sigma \rangle = \langle \nu, \sigma \rangle$. Hence the index of $\langle \nu \rangle$ in J is at most two. If $\langle \nu \rangle = J$, then ρ and σ coincide with the unique element of order two in $\langle \nu \rangle$, but then $\langle \nu \rangle = \langle \rho^2 \rangle = 1 \neq J$.

Suppose ν has odd order $2n-1$. Then

$$\nu^n \sigma \nu^{-n} = \nu^n (\sigma^{-1} \nu^{-n} \sigma) \sigma = \nu^n (\sigma^{-1} \nu \sigma)^{-n} \sigma = \nu^n (\nu^{-1})^{-n} \sigma = \nu^{2n} \sigma = \nu \sigma = \rho \sigma^2 = \rho .$$

Lemma 2. *Let $x_1, x_2, \ldots, x_n$ and a be real numbers such that $x_1 + x_2 + \cdots + x_n = n a$. Then $x_1^2 + x_2^2 + \cdots + x_n^2 \geqq n a^2$.*

Proof. Let $y_i = x_i - a$, $i = 1, 2, \ldots, n$. Then

$$y_1 + y_2 + \cdots + y_n = (x_1 + x_2 + \cdots + x_n) - n a = 0 .$$

Hence

$$\sum x_i^2 = \sum (a + y_i)^2 = \sum a^2 + 2a \sum y_i + \sum y_i^2 = n a^2 + \sum y_i^2 \geqq n a^2 .$$

In Lemmas 3 through 10 we assume that A is an arbitrary operator group on an arbitrary group G such that G and A have relatively prime orders; let $C(A)$ be the fixed point subgroup of G with respect to A.

Lemma 3 ([4], Corollary 2 of Theorem 1). *Let J be a normal subgroup of G that is fixed by A. Consider A as an operator group on G/J. The fixed point subgroup of G/J with respect to A is $C(A)\,J/J$.*

Lemma 4 ([4], Corollaries of Theorem 4). *If A fixes a conjugate class K of elements or subgroups of G, then A fixes at least one member of K. In particular, for every prime p, A fixes a Sylow p-subgroup of G.*

Lemma 5 ([4], Corollaries of Theorem 3). *If two elements or subsets of G are fixed by A and are conjugate in G, then they are conjugate by an element of $C(A)$. For every non-empty subset S of $C(A)$, $N_G(S)=N_{C(A)}(S)\,C_G(S)$.*

Lemma 6 ([10], Hilfssatz 4.2). *Let P be a p-subgroup of G that is fixed by A. Then P is contained in a Sylow p-subgroup of G that is fixed by A.*

Lemma 7. *Let p be a prime. Suppose there exists a p-subgroup P of $C(A)$ such that $C_G(P)\subseteq C(A)$. Then $C(A)$ contains every Sylow p-subgroup of G which contains P and is fixed by A. Moreover, if J is a subgroup of G fixed by A, then $J\cap C(A)$ contains a Sylow p-subgroup of J.*

Proof. Let S be a Sylow p-subgroup of G such that $P\subseteq S$. Let $S_0=S\cap C(A)$. Since $P\subseteq S_0$, $C_S(S_0)\subseteq C_S(P)\subseteq S\cap C(A)=S_0$. Hence by Lemma 5, $N_S(S_0)=N_{S_0}(S_0)\,C_S(S_0)=S_0$. By Sylow's Theorem, S_0 is a Sylow p-subgroup of S, i.e., $S=S_0\subseteq C(A)$.

Now let J be a subgroup of G that is fixed by A. By Lemma 4, A fixes some Sylow p-subgroup P of J. Take S as in the above paragraph. By Sylow's Theorem, P is conjugate to some subgroup R of S. By Lemma 5, there exists $t\in C(A)$ such that $P=t^{-1}R\,t$. Hence $P\subseteq t^{-1}S\,t\subseteq t^{-1}\,C(A)\,t=C(A)$.

Lemma 8. (a) *(G, A) is a normal subgroup of G which is fixed by A; $G=C(A)(G, A)$; and $(G, A)=((G, A), A)$.*

(b) *If G is Abelian then $G=(G, A)\times C(A)$.*

Proof. (a) This is proven in the proof of Corollary 3 of Theorem 1 of [4].

(b) This is proven in [8], p. 177, for the case in which G is a p-group. However, the proof is valid in the general case.

Lemma 9. *Let N be a normal subgroup of G contained in $C(A)$. Then (G, A) centralizes N.*

Proof. Let $S=N$ in Lemma 5.

Lemma 10. *Let M be the set of all $t\in C(A)$ such that*

$$[C_G(t)\colon C_{C(A)}(t)]=[G\colon C(A)].$$

Then M is the largest normal subgroup of G that is contained in $C(A)$. Every normal subgroup of $C(A)$ which is contained in M is a normal subgroup of G.

Proof. Let N be the largest normal subgroup of G that is contained in $C(A)$. Let $t\in M$. Then

$$\begin{aligned}[G\colon C_G(t)]&=[G\colon C_G(t)]\,[C_G(t)\colon C_{C(A)}(t)]/[G\colon C(A)]\\&=[G\colon C_{C(A)}(t)]/[G\colon C(A)]=[C(A)\colon C_{C(A)}(t)].\end{aligned}$$

Thus the number of conjugates of t in G is the same as the number of conjugates of t in $C(A)$. So both sets of conjugates coincide. Clearly they generate a normal subgroup of G that is contained in $C(A)$. Hence $t\in N$. Since t is arbitrary, $M\subseteq N$.

Conversely, let $t\in N$. Then all the conjugates of t in G lie in $C(A)$. By Lemma 5 these elements are conjugate to t within $C(A)$. Hence $[G: C_G(t)]=[C(A): C_{C(A)}(t)]$, and so $[G: C(A)]=[C_G(t): C_{C(A)}(t)]$. Thus $N\subseteq M$. So $N=M$.

Let L be a normal subgroup of $C(A)$ that is contained in M. Then the conjugates of L in G are all contained in M and therefore in $C(A)$. By Lemma 5 all of these conjugates are equal to L. Hence L is a normal subgroup of G.

3. Proof of Parts (i) to (iii) of Theorem 1

Using the lemmas in Section 2, we are able to prove parts (i), (ii), and (iii) of Theorem 1 by a relatively brief counting argument. The theorem is trivial if $A=1$; we assume $A\neq 1$. Let K and K^* be respectively the conjugate class of τ in G and in $C(A)$. By Lemma 5, $K^*=K\cap C(A)$. We let P (respectively, P^*) be the set of all ordered pairs (ρ,σ) such that $\rho\sigma\in C(A)$ and $\rho,\sigma\in K$ (respectively, $\rho,\sigma\in K^*$). For $t\in C(A)$, define h_t (respectively, h_t^*) to be the number of pairs (ρ,σ) in P (respectively, in P^*) such that $\rho\sigma=t$. We let I (respectively, I^*) be the set of all elements t in $C(A)$ such that $h_t\neq 0$ (respectively, $h_t^*\neq 0$). Finally, let $h=[C(A): C_{C(A)}(\tau)]$. Note that $I^*\subseteq I$.

Proposition 1. *$|P|=\sum_{t\in I} h_t$ and $|P^*|=\sum_{t\in I^*} h_t^*$.*

Proof. We merely enumerate the pairs (ρ,σ) by noting the number of pairs with a given product $\rho\sigma$.

Proposition 2. *Let $t\in I$. Then $t\in I^*$ and $h_t=[C_G(t): C_{C(A)}(t)]\, h_t^*$.*

Proof. Assume $t\in I$. Let J be the set of all g in G such that $g^{-1}tg$ equals t or t^{-1}. Clearly J is a subgroup of G which is fixed by A and contains $C_G(t)$. Suppose $(\rho,\sigma)\in P$ and $\rho\sigma=t$. By Lemma 1, $\rho^{-1}t\rho=\sigma^{-1}t\sigma=t^{-1}$. Hence $\rho,\sigma\in J$.

In Lemma 7 take $p=2$ and $P=\langle\tau\rangle$; we see that $J\cap C(A)$ contains some Sylow 2-subgroup S of J. Since $C_G(t)$ is a normal subgroup of index one or two in J (one if $t^2=1$), $J=C_G(t)\,S$.

Suppose $(\rho,\sigma)\in P$, $u\in C_G(t)$, and $\rho\sigma=t$. Then $(u^{-1}\rho u, u^{-1}\sigma u)\in P$ and $(u^{-1}\rho u)(u^{-1}\sigma u)=u^{-1}tu=t$. Thus $C_G(t)$ permutes by conjugation the set, say, $P(t)$, of all $(\rho,\sigma)\in P$ such that $\rho\sigma=t$. Let D be an orbit of $P(t)$ with respect to $C_G(t)$. Suppose $(\rho,\sigma)\in D$. Now, some conjugate of ρ lies in S. Take $g\in J$ such that $g^{-1}\rho g\in S$. Take $k\in C_G(t)$ and $s\in S$ such that $g=ks$. Then $s^{-1}k^{-1}\rho ks\in S$ and thus $k^{-1}\rho k\in S$. Let $\rho^*=k^{-1}\rho k$ and $\sigma^*=k^{-1}\sigma k$. Then $(\rho^*,\sigma^*)\in D$, $\rho^*\in C(A)$, and $\sigma^*=\rho^*(\rho^*\sigma^*)=\rho^* t\in C(A)$. Thus $(\rho^*,\sigma^*)\in D\cap P^*$. Hence $t\in I^*$.

Let L be the set of all elements of $C_G(t)$ which fix (ρ^*,σ^*), i.e., the set of all $g\in C_G(t)$ such that $g^{-1}\rho^*g=\rho^*$, $g^{-1}\sigma^*g=\sigma^*$. From the elementary theory of permutation groups (e.g., [9], Corollary 5.2.1, p. 56), $|D|=[C_G(t): L]$. Since $\rho^*\in K^*$, there exists $u\in C(A)$ such that $\rho^*=u^{-1}\tau u$. Then

$$L\subseteq C_G(\rho^*)=u^{-1}C_G(\tau)\,u\subseteq u^{-1}C(A)\,u=C(A).$$

Hence $L\subseteq C_{C(A)}(t)$.

Let M be the set of all $g \in C_G(t)$ such that $(g^{-1} \rho^* g, g^{-1} \sigma^* g) \in D \cap P^*$. Since $C_G(t)$ is transitive on D, $|M|/|L| = |D \cap P^*|$. Clearly $C_{C(A)}(t) \subseteq M$. Conversely, suppose $g \in M$. By Lemma 5 there exists $u \in C(A)$ such that $g^{-1} \rho^* g = u^{-1} \rho^* u$. Hence $g u^{-1} \in C_G(\rho^*) \subseteq C(A)$, and $g \in C(A) u = C(A)$. Thus $M = C_{C(A)}(t)$. This shows that

$$|D| = [C_G(t): L] = [C_G(t): C_{C(A)}(t)] [C_{C(A)}(t): L]$$
$$= [C_G(t): C_{C(A)}(t)] \, |D \cap P^*|.$$

Summing over all D, we obtain

$$h_t = |P(t)| = \sum |D| = \sum [C_G(t): C_{C(A)}(t)] \, |D \cap P^*|$$
$$= [C_G(t): C_{C(A)}(t)] \sum |D \cap P^*| = [C_G(t): C_{C(A)}(t)] \, h_t^*.$$

Recall that $h = [C(A): C_G(\tau)]$.

Proposition 3. $|P| \geqq [G: C(A)] \, h^2$ and $|P^*| = h^2$.

Proof. Trivially, $\rho \sigma \in C(A)$ if $\rho, \sigma \in K^*$. Hence P^* contains all pairs (ρ, σ) for $\rho, \sigma \in K^*$, and so $|P^*| = |K^*|^2 = h^2$.

If ρ, $\sigma \in K$, the following are equivalent: $\rho \sigma \in C(A)$; $C(A) \rho \sigma = C(A)$; $C(A) \rho = C(A) \sigma$. Thus $(\rho, \sigma) \in P$ if and only if ρ and σ lie in the same left coset of $C(A)$ in G. Let $n = [G: C(A)]$ and let $C(A) g_1$, $C(A) g_2, \ldots, C(A) g_n$ be the distinct left cosets of $C(A)$ in G. Define $x_i = |C(A) g_i \cap K|$, $i = 1, 2, \ldots, n$. Then $x_1 + x_2 + \cdots + x_n = |K|$ and $|P| = x_1^2 + x_2^2 + \cdots + x_n^2$. Now, $|K| = [G: C_T(\tau)] = [G: T][T: C_T(\tau)] = n h$. By Lemma 2, $|P| \geqq n h^2 = [G: C(A)] \, h^2$.

Let N be the subgroup of G defined in Theorem 1.

Proposition 4. *N is a normal subgroup of G and $\tau \notin N$.*

Proof. We first verify that N is a normal subgroup of $C(A)$. Let $t, u \in C(A)$. Then $u^{-1}(t^{-1} \tau^{-1} t \tau) u = ((t u)^{-1} \tau^{-1} (t u) \tau)(u^{-1} \tau^{-1} u \tau)^{-1}$. Thus $u^{-1} N u \subseteq N$. Similarly, $u N u^{-1} = (u^{-1})^{-1} N u^{-1} \subseteq N$, so $N \subseteq u^{-1} N u$. Hence $u^{-1} N u = N$. Since u is arbitrary, N is a normal subgroup of $C(A)$.

By Propositions 1, 2, and 3,

$$\sum_{t \in I^*} [G: C(A)] \, h_t^* = [G: C(A)] |P^*| = [G: C(A)] \, h^2 \leqq |P| = \sum_{t \in I^*} h_t$$
$$= \sum_{t \in I^*} [C_G(t): C_{C(A)}(t)] \, h_t^*.$$

Thus

$$\sum_{t \in I^*} \big([G: C(A)] - [C_G(t): C_{C(A)}(t)]\big) \, h_t^* \leqq 0.$$

Since each term in this summation is nonnegative, each term must be zero. But $h_t^* > 0$ for all $t \in I^*$. Hence for each $t \in I^*$, $[G: C(A)] = [C_G(t): C_{C(A)}(t)]$. Let M be the largest normal subgroup of G contained in $C(A)$. By Lemma 10, $I^* \subseteq M$. Likewise, $\tau \notin M$ because $1 = [C_G(\tau): C_{C(A)}(\tau)] < [G: C(A)]$. Now if $t \in C(A)$, then $t^{-1} \tau^{-1} t \tau = (t^{-1} \tau t) \tau \in I^*$. Hence $N \subseteq M$. By Lemma 10, N is a normal subgroup of G. Since $\tau \notin M$, $\tau \notin N$.

Thus part (i) of Theorem 1 has been proved. Lemma 3, applied to A and to each cyclic subgroup of A, yields (ii). For (iii), consider $\bar{\tau}$. By (i), $\bar{\tau} \neq 1$. Hence $\bar{\tau}$ has order two. If $t \in C(A)$ then $(tN)^{-1}(\tau N)^{-1}(tN)(\tau N) = (t^{-1}\tau^{-1} t \tau) N = N$. Hence $C_{\bar{G}}(\bar{\tau}) \supseteq \bar{C}$. Conversely, suppose $g \in G$ and $gN \in C_{\bar{G}}(\bar{\tau})$. Then $g^{-1}\tau^{-1} g \tau \in N$. Hence $g^{-1}\tau g = g^{-1}\tau^{-1} g \in N\tau \subseteq C(A)$. By Lemma 5 there exists $t \in C(A)$ such that $g^{-1}\tau g = t^{-1}\tau t$. Now $g t^{-1} \in C_G(t) \subseteq C(A)$, and $g \in C(A)\,t = C(A)$. Thus $gN \in \bar{C}$. Hence $C_{\bar{G}}(\bar{\tau}) = \bar{C}$.

4. Proof of Theorem 1 (iv)

For every finite group G, let $K(G)$ be the unique maximal normal subgroup of odd order in G and let $Z^*(G)$ be the subgroup of G that contains $K(G)$ and satisfies

$$Z^*(G)/K(G) = Z(G/K(G)).$$

Define $F(G)$ to be the unique maximal nilpotent normal subgroup of G ([8], p. 218). To prove part (iv) of Theorem 1, we use three results.

Theorem 4.1. *Let S be a Sylow 2-subgroup of a finite group G. Suppose τ is an involution in S that does not lie in $Z^*(G)$. Then there exists $g \in G$ such that $g^{-1}\tau g \in S$ and $g^{-1}\tau g \neq \tau$.*

Theorem 4.2 [3]. *Every finite group of odd order is solvable.*

Theorem 4.3. *Let G be a finite solvable group. Let r and s be two primes that do not divide $|G|$. Suppose that r is odd, A is a cyclic r-group, B is a group of order s, $A \times B$ is an operator group on G, and every fixed point of G under B is a fixed point of G under A. Then $(G, A) \subseteq F(G)$.*

Theorem 4.1 is Theorem 2 of [6]. Theorem 4.3 is a special case of Theorem 6 of [7].

Now we prove part (iv) of Theorem 1 by induction on $|G|$. Take G, A, and N as in the hypothesis of the theorem. We may assume that $A \neq 1$. By Lemma 8, (G, A) is a normal subgroup of G. First, suppose $N \neq 1$. By parts (i) and (iii) of Theorem 1 and the induction hypothesis, $(G, A)\,N/N$ is a nilpotent group of odd order. Let $M = N \cap (G, A)$; then $(G, A)/M$ is a nilpotent group of odd order. By Lemma 9, $M \subseteq Z((G, A))$. So the upper central series of (G, A) reaches (G, A), and (G, A) is nilpotent. By Lemma 8, A fixes (G, A). Let S be the Sylow 2-subgroup of (G, A) and let R be the product of the Sylow subgroups of (G, A) of odd order. Then A fixes R and S, and $(G, A) = R \times S$. Note that $S \subseteq M \subseteq N \subseteq C(A)$. By Lemma 8,

$$(G, A) = ((G, A), A) \subseteq (R, A) \times (S, A) = (R, A) \subseteq R.$$

Thus (G, A) has odd order, as desired.

Now suppose $N = 1$. Then $\tau \in Z(C(A))$ and $C(A) = C_G(\tau)$. By Lemma 7, $[G\colon C(A)]$ is odd, so τ lies in a Sylow 2-subgroup S of G such that $S \subseteq C(A)$. Suppose $g \in G$ and $g^{-1}\tau g \in S$. By Lemma 5, there exists $t \in C(A)$ such that $t^{-1}\tau t = g^{-1}\tau g$. Since $\tau \in Z(C(A))$, $g^{-1}\tau g = \tau$. Hence by Theorem 4.1, $\tau \in Z^*(G)$.

Let $K=K(G)$, $R=\langle\tau\rangle$, and $M=RK$. Since $\tau\in Z^*(G)$, M is a normal subgroup of G and R is a Sylow 2-subgroup of M.

Now let g be an arbitrary element of G. Then $g^{-1}Rg$ is a Sylow 2-subgroup of M. Take $h\in M$ such that $g^{-1}Rg=h^{-1}Rh$. Take $r\in R$ and $k\in K$ such that $h=rk$. Then $g^{-1}Rg=k^{-1}Rk$. Therefore, $gk^{-1}\in N_G(R)=C_G(\tau)=C(A)$. Let $t=gk^{-1}$ and let $\alpha\in A$. Then $g=tk$ and $g^{-1}g^{\alpha}=k^{-1}t^{-1}t^{\alpha}k^{\alpha}=k^{-1}k^{\alpha}$. Thus $(G,A)=(K,A)$. Therefore, (G,A) is contained in K and has odd order.

By Theorem 4.2, K is solvable. Let β be the automorphism of K given by $g\to\tau^{-1}g\tau$, $g\in G$. Let $B=\langle\beta\rangle$. Since $A\neq 1$ we have $(K,A)=(G,A)\neq 1$ and $K\nsubseteq C(A)=C_G(\tau)$. So $|B|=2$. Since $|A|$ and $|G|$ are relatively prime, $|A|$ is odd. For every odd prime r and every cyclic r-subgroup A_1 of A, Theorem 4.3 yields that $(K,A_1)\subseteq F(K)$ and therefore that A_1 operates trivially on $K/F(K)$. However, A is generated by the set of all such subgroups A_1. Consequently, A operates trivially on $K/F(K)$. Hence

$$(G,A)=(K,A)\subseteq F(K),$$

and (G,A) is nilpotent. This completes the proof of part (iv) of Theorem 1 and thus of Theorem 1.

Remark. At the time the original version of this paper was written, Theorem 4.1 was not yet known. Part (iv) of Theorem 1 was obtained for $N=1$ by using a special case of Theorem 4.1. The latter was proved by defining a loop operation on the set $\{g^{-1}g^{\tau}|g\in G\}$ and using the methods of [5].

References

1. Brauer, R., Fowler, K.A.: On groups of even order. Ann. of Math. **62**, 565–583 (1955).
2. Curtis, C.W., Reiner, I.: Representation theory of finite groups and associative algebras. New York-London: John Wiley & Sons 1962.
3. Feit, W., Thompson, J.G.: Solvability of groups of odd order. Pacific J. Math. **13**, 775–1029 (1963).
4. Glauberman, G.: Fixed points in groups with operator groups. Math. Z. **84**, 120–125 (1964).
5. — On loops of odd order. J. Algebra **1** 374–396 (1964).
6. — Central elements in core-free groups. J. Algebra **4**, 403–420 (1966).
7. — Correspondences of characters for relatively prime operator groups. Canadian J. Math. **20**, 1465–1488 (1968).
8. Gorenstein, D.: Finite groups. New York: Harper & Row 1968.
9. Hall, M.: The theory of groups. New York: Macmillan 1959.
10. Wielandt, H.: Beziehungen zwischen den Fixpunktzahlen von Automorphismengruppen einer endlichen Gruppe. Math. Z. **73**, 146–158 (1960).

Prof. G. Glauberman
Department of Mathematics
University of Chicago
Chicago 37, Illinois
U.S.A.

(Received August 16, 1971)

Math. Z. 124, 361–370 (1972)

On the Riesz Uniqueness Theorem for Functions of Nearly Bounded Characteristics

JIROKICHI NAGATOMO and NIRO YANAGIHARA

Let $f(z)$ be a holomorphic function in the unit disk D: $|z|<1$. If

$$\sup_{0 \leq r<1} \int_0^{2\pi} \log^+ |f(r e^{i\theta})| \, d\theta < \infty, \tag{1}$$

$f(z)$ is said to be a function *of bounded characteristic*, following Nevanlinna [5]. The class of holomorphic functions of bounded characteristic is denoted by N. Properties of functions of the class N are well studied. Among them we mention here two famous theorems:

Theorem F (Fatou-Nevanlinna). *Suppose $f(z) \in N$, then $f(z)$ has finite radial limits almost everywhere on the circumference $|z|=1$.*

We denote the unit circumference $|z|=1$ by C.

Theorem R (F. and M. Riesz). *Suppose $f(z) \in N$. If there is a set $E \subset C$, meas $(E)>0$, such that $f(z)$ has radial limit 0 at each point of E, then $f(z)$ must be identically equal to 0.*

By the ambiguous point theorem of Bagemihl [1] we can restate this theorem in the following equivalent form:

Theorem R*. *Suppose $f(z) \in N$. If there is a set $E \subset C$, meas $(E)>0$, such that $f(z)$ has asymptotic value 0 at each point of E, then $f(z)$ must be identically equal to 0.*

We say $f(z)$ has asymptotic value α at a point $\zeta \in C$ if there is a path W: $z=z(t)$, $0 \leq t<1$, in D such that

$$\lim_{t \to 1} z(t)=\zeta, \quad \lim_{t \to 1} f(z(t))=\alpha.$$

On the other hand, suppose $\chi(t)$ be a real function of t, $0 \leq t<\infty$, such that $\chi(t)$ is non-negative, continuous and non-decreasing, and

$$\chi(t)=o(t) \quad \text{as } t \to \infty. \tag{2}$$

When a holomorphic function $f(z)$ in D satisfies the condition

$$\sup_{0 \leq r<1} \int_0^{2\pi} \chi[\log^+ |f(r e^{i\theta})|] \, d\theta < \infty, \tag{3}$$

we say that $f(z)$ belongs to the class N^χ. If a function $f(z)$ belongs to N^χ for some function $\chi(t)$, which obeys the condition (2), we say that $f(z)$ belongs to the class N^{near} of holomorphic functions *of nearly bounded characteristic.*

For functions of N^{near}, the above theorem F is not valid. In fact, Paley and Zygmund [6] showed that : *Given any function $\chi(t)$ satisfying* (2), *there is a function $f(z)$ which belongs to N^χ and has radial limits almost nowhere on C.* See also Kahane [3, p. 109].

In this note we wish to show that also the Riesz uniqueness theorem in the form of theorem R^* is not valid for functions of N^{near}. That is, we prove the

Theorem 1. *Given any function $\chi(t)$ satisfying* (2), *there is a function $f(z)\in N^\chi$, not identically* 0, *such that $f(z)$ has asymptotic value* 0 *at almost every point of C.*

By the way, we remark that the invalidity of Theorems F and R for functions of slow growth were proved by MacLane [4] and Barth-Schneider [2], respectively.

First we prove the following

Theorem 2. *Given any function $\chi(t)$ satisfying* (2), *there is a holomorphic function $g(z)=u(z)+i\,v(z)$ in D such that*

$$\sup_{0\leqq r<1}\int_0^{2\pi}\chi[u^+(r\,e^{i\theta})]\,d\theta<\infty, \tag{4}$$

and $u(z)$ has asymptotic value $-\infty$ at almost every point of C. (We write $u^+(z)$ for $\max(u(z),0)$.)

Proof. $\chi(t)$ may be supposed to be such that

$$\chi(t_1+t_2)\leqq\chi(t_1)+\chi(t_2),\qquad 0\leqq t_1,\ t_2<\infty, \tag{5}$$

and

$$\chi(t)/t\downarrow 0\qquad \text{as } t\to\infty. \tag{5'}$$

We take a sequence $\{\alpha_n\}$, $\alpha_n>0$, $\alpha_n\downarrow 0$, such that

$$\sum\alpha_n<1/60, \tag{6}$$

$$\sum\alpha_n\,\chi(1/\alpha_n)=B<\infty, \tag{7}$$

$$\sum\frac{\chi[\chi(1/\alpha_n)]}{\chi(1/\alpha_n)}=A<\infty. \tag{8}$$

Let $\omega_n^*(z)$ be a harmonic function in D with boundary values

$$\begin{aligned}\omega_n^*(e^{i\theta})&=m_n, && 0\leqq|\theta|<\alpha_n\,\pi\\ &=-k_n, && \alpha_n\,\pi<|\theta|<(\alpha_n+\alpha_n^2)\,\pi\\ &=-1, && (\alpha_n+\alpha_n^2)\,\pi<|\theta|\leqq\pi,\end{aligned}$$

where m_n and k_n are positive numbers such that

$$m_n \alpha_n = M \ (=\text{const independent of } n, M>1),$$
$$k_n \alpha_n^2 = m_n \alpha_n - 1 + (\alpha_n + \alpha_n^2)$$
$$= M - 1 + \alpha_n(1+\alpha_n).$$

Then

$$\omega_n^*(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega_n^*(e^{i\theta})\, d\theta = 0.$$

Hence the niveau line L_n: $\omega_n^*(z)=0$ goes through the origin. We need the following

Lemma. *L_n does not penetrate into the left half-disk $|z|<1$, $\operatorname{Im}[z]<0$, supposed that α_1 is sufficiently small.*

Proof. It suffices to show that $\omega_n^*(z) \leqq 0$ on the imaginary axis. For simplicity we omit suffix n in this proof.

Let $C_1 = \{e^{i\theta}; |\theta| \leqq \alpha\pi\}$, $C_2 = \{e^{i\theta}; |\theta| \leqq (\alpha+\alpha^2)\pi\}$, $C_3 = C \setminus C_2$. We denote by $\omega'(z)$, $\omega''(z)$ and $\omega'''(z)$ the harmonic measures of C_1, C_2 and C_3 respectively. We put

$$\omega^*(z) = (m+k)\,\omega' - k\,\omega'' - \omega''',$$

where $m, k>0$ are constants such that

$$k\alpha^2 = m\alpha - 1 + (\alpha+\alpha^2). \tag{9}$$

We will do if we can show that

$$\omega^*(ir) \leqq 0, \qquad 0 \leqq r < 1, \tag{10}$$

supposed that α is sufficiently small.

We write, for a point z, $|z|<1$,

$$\theta'(z) = \arg \frac{e^{i\alpha\pi} - z}{e^{-i\alpha\pi} - z}, \qquad \theta''(z) = \arg \frac{e^{i(\alpha+\alpha^2)\pi} - z}{e^{-i(\alpha+\alpha^2)\pi} - z}.$$

Then

$$\omega'(z) = \frac{\theta'(z) - \alpha\pi}{\pi}, \qquad \omega''(z) = \frac{\theta''(z) - (\alpha+\alpha^2)\pi}{\pi},$$

$$\omega'''(z) = 1 - \omega''(z).$$

Thus

$$\omega^*(z) = \frac{1}{\pi}\{(m+k)\,\theta'(z) - (k-1)\,\theta''(z)\} - 2.$$

By easy calculation we can see that for $z = ir$

$$\theta'(ir) = \tan^{-1} \frac{\sin 2\alpha\pi}{r^2 + \cos 2\alpha\pi},$$

$$\theta''(ir) = \tan^{-1} \frac{\sin 2\beta\pi}{r^2 + \cos 2\beta\pi}, \qquad \beta = \alpha + \alpha^2.$$

Thus

$$\frac{d}{dr}\,\omega^*(i\,r) = -2r\left\{\frac{(m+k)\sin 2\alpha\pi}{r^4+2r^2\cos 2\alpha\pi+1} - \frac{(k-1)\sin 2\beta\pi}{r^4+2r^2\cos 2\beta\pi+1}\right\},$$

from which we get without difficulty

$$\frac{d}{dr}\,\omega^*(i\,r) = \frac{-4r}{r^4+2r^2\cos 2\alpha\pi+1}\,\{1+O(\alpha^2)\},$$

where O is independent of r. Hence, if α is taken small enough, we have $(d/dr)\cdot\omega^*(i\,r)\leqq 0$ and (10) follows. Q.E.D.

We assume α_1 small enough so as to meet this lemma. Let $\phi_n^*(z)$ be the holomorphic function in D such that $\operatorname{Re}[\phi_n^*(z)]=\omega_n^*(z)$ and $\phi_n^*(0)=0$.

We choose sequences of constants $\{\varepsilon_n\}$, $\{\sigma_n\}$, $\{\eta_n'\}$, $\{\mu_n\}$, $\{\eta_n\}$, as well as of sets $\{e_n^*\}$, $\{e_n\}$, $\{E_n^*\}$, $\{E_n\}$, as follows:

$\{\varepsilon_n\}$: $\varepsilon_n>0$, $\varepsilon_n\downarrow 0$, and $\sum\varepsilon_n=\varepsilon<\infty$.

$\{\sigma_n\}$: $\sigma_n>0$ are numbers such that

$$|\phi_n^*(z)|\leqq\varepsilon_n \quad \text{for } |z|\leqq\sigma_n.$$

$\{\eta_n'\}$: $\eta_n'>0$ are numbers such that

$$\omega_n^*(re^{i\theta})\leqq -1/2 \quad \text{for } 2\alpha_n\leqq|\theta|\leqq\pi,\ \eta_n'\leqq r\leqq 1.$$

$\{\mu_n\}$: μ_n are positive integers such that

$$(1/2)^{\mu_n}\leqq\sigma_n,$$

$\{\eta_n\}$: $\eta_n=(\eta_n')^{1/\mu_n}>0$.

We put

$$\phi_n(z)=\phi_n^*(z^{\mu_n}). \tag{11}$$

$\{e_n^*\}$, $\{e_n\}$: Let C_n^* and C_n be subarcs of C such that

$$C_n^*=\{e^{i\theta};\,2\alpha_n\leqq|\theta|\leqq\pi\},$$
$$C_n=\{e^{i\theta};\,3\alpha_n\leqq|\theta|\leqq\pi\},$$

then we set

$$e_n^*=\{z;\,|z|=1,\,z^{\mu_n}\in C_n^*\},$$
$$e_n=\{z;\,|z|=1,\,z^{\mu_n}\in C_n\}.$$

$\{E_n^*\}$, $\{E_n\}$: These are subsets of the closed disk $|z|\leqq 1$,

$$E_n^*=\{r\,e^{i\theta};\,e^{i\theta}\in e_n^*,\,\eta_n\leqq r\leqq 1\},$$
$$E_n=\{r\,e^{i\theta};\,e^{i\theta}\in e_n,\,\eta_n\leqq r\leqq 1\}.$$

Then

$$|\phi_n(z)|\leqq\varepsilon_n \quad \text{if} \quad |z|\leqq 1/2, \tag{12}$$

$$\omega_n(z)=\operatorname{Re}[\phi_n(z)]\leqq -1/2 \quad \text{if} \quad z\in E_n^*. \tag{13}$$

Now we determine sequences $\{\rho_n\}$ of positive numbers and $\{\lambda_n\}$ of positive integers inductively in the following way:

Put

$$\rho_1 = 1/2, \qquad \lambda_1 = 1.$$

Let $0 < \rho_2 < 1$ be a number such that

$$\rho_2^{\lambda_1} \geqq \eta_1, \qquad \rho_2 \geqq \rho_1, \qquad \rho_2 \geqq 1 - 1/2,$$

and λ_2 be a positive integer such that

$$\begin{aligned} &\rho_2^{\lambda_2} \leqq 1/2, \\ &\frac{1}{\lambda_2} \frac{\pi}{2} \leqq \frac{\alpha_1}{\lambda_1} \pi, \\ &\lambda_2/\lambda_1 \quad \text{is an integer.} \end{aligned}$$

Suppose ρ_n, λ_n are determined. Then let $\rho_{n+1} (0 < \rho_{n+1} < 1)$ be a number such that

$$\rho_{n+1}^{\lambda_n} \geqq \eta_n, \qquad \rho_{n+1} \geqq \rho_n, \qquad \rho_{n+1} \geqq 1 - \frac{1}{n+1}, \tag{14}$$

and λ_{n+1} be a positive integer such that

$$\begin{aligned} &\rho_{n+1}^{\lambda_{n+1}} \leqq 1/2, \\ &\frac{1}{\lambda_{n+1}} \frac{\pi}{2} \leqq \min \left\{ \frac{\alpha_n}{\lambda_n}, \frac{\alpha_{n-1}}{\lambda_{n-1}}, \ldots, \frac{\alpha_1}{\lambda_1} \right\}, \\ &\lambda_{n+1}/\lambda_n \quad \text{is an integer.} \end{aligned} \tag{15}$$

We put

$$g(z) = \sum \phi_n(z^{\lambda_n}). \tag{16}$$

The sum on the right-hand side converges uniformly in $|z| \leqq \rho$ for every $\rho < 1$, whence $g(z)$ is holomorphic in D.

Let $\{s_n^*\}$, $\{s_n\}$, and $\{S_n^*\}$, $\{S_n\}$ be sequences of sets such that

$$\begin{aligned} &s_n^* = \{z \in C;\ z^{\lambda_n} \in e_n^*\}, \\ &s_n = \{z \in C;\ z^{\lambda_n} \in e_n\} \subset s_n^* \end{aligned}$$

and

$$\begin{aligned} &S_n^* = \{z = r\,e^{i\theta};\ e^{i\theta} \in s_n^*,\ \rho_n \leqq r \leqq 1\}, \\ &S_n = \{z = r\,e^{i\theta};\ e^{i\theta} \in s_n,\ \rho_n \leqq r \leqq 1\} \subset S_n^*, \end{aligned}$$

and put

$$E = \bigcup_{h=1}^{\infty} \left(\bigcap_{n=h}^{\infty} S_n \right) = \bigcup_{h=1}^{\infty} \left(\bigcap_{n=h}^{\infty} s_n \right) \subset C.$$

We have

$$\operatorname{meas}(E)=2\pi, \tag{17}$$

since for every $\eta>0$ we take an h_0, $6\sum_{n=h_0}^{\infty}\alpha_n\pi<\eta$, then

$$\begin{aligned}\operatorname{meas}(E)&\geqq \operatorname{meas}\left(\bigcap_{n=h_0}^{\infty}S_n\right)=\operatorname{meas}\left(C\setminus\bigcup_{n=h_0}^{\infty}(C\setminus S_n)\right)\\&=\operatorname{meas}(C)-\operatorname{meas}\left[\bigcup_{n=h_0}^{\infty}(C\setminus S_n)\right]\\&\geqq 2\pi-\sum_{n=h_0}^{\infty}\operatorname{meas}(C\setminus S_n)\geqq 2\pi-6\sum_{n=h_0}^{\infty}\alpha_n\pi\\&>2\pi-\eta.\end{aligned}$$

We will show that for each $e^{i\theta}\in E$ there is a path $W=W(e^{i\theta})$, which ends at $e^{i\theta}$ and along which $u(z)=\operatorname{Re}[g(z)]$ tends to $-\infty$.

Take a point $e^{i\theta_0}\in E$ and let γ be the unit radius ending at $e^{i\theta_0}$. Let h_1 be the least positive integer such that

$$\gamma\cap\left(\bigcap_{n=h_1}^{\infty}S_n\right)\neq\emptyset.$$

Let h_2 be an integer such that

$$2\varepsilon+2\sum_{n=1}^{h_1-1}m_n\leqq\sum_{n=h_1}^{h_2-1}1.$$

Let $K>0$ be any given number and k be an integer such that

$$\sum_{n=h_2}^{k-1}1\geqq 2K.$$

Then, if $z\in\gamma_m=\gamma\cap\{\rho_{k+m}\leqq|z|\leqq\rho_{k+m+1}\}$ for an $m\geqq 0$,

$$\begin{aligned}u(z)=\operatorname{Re}[g(z)]&=\sum_{n=1}^{h_1-1}\operatorname{Re}[\phi_n(z^{\lambda_n})]+\sum_{n=h_1}^{h_2-1}+\sum_{n=h_2}^{k+m-1}+\operatorname{Re}[\phi_{k+m}(z^{\lambda_{k+m}})]+\sum_{n=k+m+1}^{\infty}\\&\leqq\sum_{n=1}^{h_1-1}m_n+\sum_{n=h_1}^{h_2-1}(-1/2)+\sum_{n=h_2}^{k+m-1}(-1/2)+\operatorname{Re}[\phi_{k+m}(z^{\lambda_{k+m}})]+\sum_{n=k+m+1}^{\infty}\varepsilon_n \qquad (18)\\&\leqq -K-\frac{m}{2}+\operatorname{Re}[\phi_{k+m}(z^{\lambda_{k+m}})].\end{aligned}$$

Let $\gamma^*=\gamma_m^*$ be the image of γ_m by $z^{\lambda\mu}$ ($\lambda=\lambda_{k+m}$, $\mu=\mu_{k+m}$), i.e.,

$$\gamma^*=\{r\,e^{i\lambda\mu\theta_0};\ \sigma^*\leqq r\leqq\eta'^*\}\qquad(\sigma^*=\sigma_{k+m},\ \eta'^*=\eta'_{k+m}).$$

Then

$$\theta_0^* = \lambda\mu\,\theta_0 \in \{3\alpha_{k+m}\pi \leqq |\theta| \leqq \pi\} \qquad \mathrm{mod}\,2\pi.$$

We may suppose $3\alpha_{k+m}\pi \leqq \theta_0^* \leqq \pi$.

Suppose γ^* does not intersect with the niveau line

$$L_{k+m}\colon\ \omega_{k+m}^*(z) = 0,$$

then we put

$$W_m = \gamma_m.$$

Suppose γ^* intersects with L_{k+m}. Let L_{k+m}^* be the part of $L_{k+m} \cap \{|z| \geqq \sigma_{k+m}\}$ lying between the radius $\gamma^\# = \{r\,e^{i\theta_0^*}; 0 \leqq r \leqq 1\}$ and the upper imaginary axis $\{i\,r\colon r \geqq 0\}$. We denote the part of $|z| = \sigma_{k+m}$ lying between $\gamma^\#$ and L_{k+m} by L_{k+m}^{**}. We write

$$L^\# = L_{k+m}^\# = L_{k+m}^* \cup L_{k+m}^{**},$$

and put W_m the component of $\{z; z^{\lambda\mu} \in L^\#\}$ which intersects with γ_m. Take a point $z = r\,e^{i\theta} \in W_m$. We prove that the point $z_h = z^{\lambda_h\mu_h}$ satisfies

$$\eta_h' \leqq |z_h| < 1, \qquad 2\alpha_h\pi \leqq |\arg[z_h]| \leqq \pi, \tag{19}$$

for $h = h_1, h_1+1, h_1+2, \ldots, k+m-1$.

By the choice of θ_0 and θ,

$$3\alpha_h\pi \leqq |\lambda_h\mu_h\theta_0| \leqq \pi \qquad \mathrm{mod}\,2\pi \tag{19$_1$}$$

and

$$0 \leqq \lambda\mu(\theta - \theta_0) < \frac{\pi}{2} \qquad \mathrm{mod}\,2\pi, \tag{19$_2$}$$

hence

$$|\lambda_h\mu_h(\theta - \theta_0)| \leqq \frac{\pi}{2}\,\frac{\lambda_h\mu_h}{\lambda\mu} \leqq \frac{\pi}{2}\,\frac{\lambda_h}{\lambda} \leqq \lambda_h(\alpha_h/\lambda_h)\,\pi = \alpha_h\pi.$$

Therefore we have

$$2\alpha_h\pi \leqq |\lambda_h\mu_h\theta| \leqq \pi \qquad \mathrm{mod}\,2\pi,$$

for $z = r\,e^{i\theta} \in W_m$ and for any h ($h_1 \leqq h \leqq k+m-1$). Hence z_h satisfies (19) and there holds for $z \in W_m$

$$\begin{aligned} &\mathrm{Re}[\phi_h(z^{\lambda_h})] \leqq -1/2, \quad h = h_1,\ h_1+1,\ h_1+2, \ldots, k+m-1,\\ &\mathrm{Re}[\phi_{k=m}(z^{\lambda_{k+m}})] \leqq 0. \end{aligned} \tag{20}$$

Let

$$W = W(e^{i\theta_0}) = (\gamma \cap \{0 \leqq |z| \leqq \rho_{k-1}\}) \cup W_0 \cup W_1 \cup \ldots,$$

then we have, for $z = r\,e^{i\theta} \in W_m$,

$$|\theta - \theta_0| \leqq \frac{\pi}{2}\,\lambda_{k+m}^{-1}\,\mu_{k+m}^{-1}, \qquad u(z) \leqq -K - \frac{m}{2},$$

as seen from (18), (19_2) and (20). Letting $m\to\infty$ we obtain that

$$\text{the path } W \text{ tends to } e^{i\theta_0} \text{ and } u(z)\to-\infty \text{ along } W,$$

as desired.

It remains to show that the function $g(z)$ in (16) satisfies the condition (4). Write

$$\operatorname{Re}[\phi_n(z^{\lambda_n})]=u_n(z)$$

and put

$$\tilde{e}_n=\{\theta;\ -\pi\leqq\theta<\pi,\ e^{i\theta}\notin s_n^*\}.$$

Then

$$\operatorname{meas}(\tilde{e}_n)=4\alpha_n\pi.$$

Take a number ρ, $0<\rho<1$. If $\rho_h\leqq\rho<\rho_{h+1}$,

$$\begin{aligned}u^+(\rho e^{i\theta})&\leqq\sum_{n=1}^{h-1}u_n^+(\rho e^{i\theta})+u_h^+(\rho e^{i\theta})+\sum_{h+1}^{\infty}u_n^+(\rho e^{i\theta})\\&=Q_1+Q_2+Q_3.\end{aligned}$$

$|Q_3|<\varepsilon$ and hence

$$\int_{-\pi}^{\pi}\chi[Q_3(\rho e^{i\theta})]\,d\theta\leqq2\varepsilon\pi. \tag{21}$$

It is easy to see that

$$\int_{-\pi}^{\pi}\chi[u_h^+(\rho e^{i\theta})]\,d\theta\leqq\int_{-\pi}^{\pi}u_h^+(\rho e^{i\theta})\,d\theta\leqq\int_{-\pi}^{\pi}u_h^+(e^{i\theta})\,d\theta\leqq2m_h\alpha_h\pi=2M\pi. \tag{22}$$

For $n=1,\dots,h-1$, $\rho e^{i\theta}\in S_n^*$ if $e^{i\theta}\in s_n^*$, and we have

$$u_n(\rho e^{i\theta})\leqq-1/2,\qquad \text{so } u_n^+(\rho e^{i\theta})=0,$$

since $\rho_n<(\rho_h\leqq)\,\rho$. Hence

$$\begin{aligned}\int_{-\pi}^{\pi}\chi[u_n^+(\rho e^{i\theta})]\,d\theta&=\int_{\tilde{e}_n}\chi[u_n^+(\rho e^{i\theta})]\,d\theta=\int_{\tilde{e}_n\cap U_n}\chi[u_n^+(\rho e^{i\theta})]\,d\theta\\&\quad+\int_{\tilde{e}_n\cap U_n^C}\chi[u_n^+(\rho e^{i\theta})]\,d\theta\\&\quad(U_n=\{\theta;\ u_n^+(\rho e^{i\theta})>\chi(1/\alpha_n)\})\\&\leqq\int_{U_n}\frac{\chi(u_n^+)}{u_n^+}u_n^+\,d\theta+\int_{\tilde{e}_n}\chi\left(\frac{1}{\alpha_n}\right)d\theta\\&\leqq\frac{\chi[\chi(1/\alpha_n)]}{\chi(1/\alpha_n)}2m_n\alpha_n\pi+4\alpha_n\pi\chi(1/\alpha_n).\end{aligned}$$

Thus by the subadditivity (5) of $\chi(t)$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\chi[Q_1]\,d\theta \leqq \sum_{n=1}^{h-1}\int_{-\pi}^{\pi}\chi[u_n^+(\rho\,e^{i\theta})]\,d\theta\,\frac{1}{2\pi}$$
$$\leqq M\sum_{n=1}^{h-1}\frac{\chi[\chi(1/\alpha_n)]}{\chi(1/\alpha_n)}+2\sum_{n=1}^{h-1}\alpha_n\,\chi(1/\alpha_n) \tag{23}$$
$$\leqq MA+2B.$$

Using the subadditivity (5) of $\chi(t)$ again, we have from (21), (22), and (23),

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\chi[u^+(\rho\,e^{i\theta})]\,d\theta \leqq MA+M+2B+\varepsilon,$$

and the proof of our theorem is completed.

Proof of Theorem 1. The function

$$f(z)=\exp[g(z)],$$

where $g(z)$ is the function constructed in Theorem 2, meets our purpose. Q.E.D.

Next we will show that the theorem of Khintchine-Ostrovski [7, p. 83] is not valid for functions of N^{near}.

Theorem 3. *Given any function $\chi(t)$ satisfying* (2), *there is a sequence of functions $\{f_n(z)\}\subset N^\chi$ such that*

(i) *each $f_n(z)$ has asymptotic values almost everywhere on C. The asymptotic value at a point $e^{i\theta}$, if any, is denoted by $f_n(e^{i\theta})$. The boundary function $f_n(e^{i\theta})$ is well defined at almost every point of C, by the assumption and the ambiguous point theorem* [2],

(ii) *$f_n(z)$ satisfies*

$$\int_0^{2\pi}\chi[\log^+|f_n(r\,e^{i\theta})|]\,d\theta \leqq K, \qquad 0\leqq r<1, \tag{24}$$

where K is a constant independent of n,

(iii) *there is a set $F\subset C$,* $\operatorname{meas}(F)=2\pi$, *on which $f_n(e^{i\theta})$ tends to* 0 *uniformly as $n\to\infty$, while the sequence $\{f_n(z)\}$ converges nowhere in D.*

Proof. We put

$$f(z)=\exp[g(z)], \qquad f_n(z)=(-1)^n f(z)+\frac{z}{n},$$

where $g(z)$ is the function constructed in Theorem 2. The sequence $\{f_n(z)\}$ is the requested one. Q.E.D.

References

1. Bagemihl, F.: Curvilinear cluster sets of arbitrary functions. Proc. Nat. Acad. Sci. U.S.A. **41**, 379–382 (1955).
2. Barth, K. F., Schneider, W. J.: On the impossibility of extending the Riesz uniqueness theorem to functions of slow growth. Ann. Acad. Sci. Fennicae, Ser. A-I, no. **432**, 1–9 (1968).
3. Kahane, J.-P.: Some random series of functions. Lexington, Massachusetts: D. C. Heath & Company 1968.
4. MacLane, G. R.: Holomorphic functions of arbitrarily slow growth, without radial limits. Michigan Math. J. **9**, 21–24 (1962).
5. Nevanlinna, R.: Eindeutige analytische Funktionen, zweite Auflage. Berlin-Göttingen-Heidelberg: Springer 1953.
6. Paley, R. E. A. C., Zygmund, A.: A note on analytic functions in the unit circle. Proc. Cambridge Philos. Soc. **28**, 266–272 (1932).
7. Priwalow, I. I.: Randeigenschaften analytischer Funktionen. Berlin: VEB Deutscher Verlag der Wissenschaften 1956.

Jirokichi Nagatomo and Niro Yanagihara
Department of Mathematics
Chiba University
Chiba-shi
Japan

(Received January 15, 1971)